Luciano Amito Lomonaco

Geometria e algebra

Vettori, equazioni e curve elementari

ARACNE editrice S.r.l.

www.aracneeditrice.it
info@aracneeditrice.it

via Raffaele Garofalo, 133/A–B
00173 Roma
(06) 93781065

ISBN 978-88-548-6247-0

I edizione: agosto 2013

a nadia, stefano e viola

Qui non si trattano solo tematichematematichematematichemetàmatematichemetàmetamatematiche

(s. rao)

Indice

Premessa

In questo testo ho provato ad esporre in modo sintetico i principali argomenti di base di Algebra Lineare e di Geometria, cercando di mantenere un livello accettabile di rigore formale. L'opera è indirizzata a studenti del primo anno di corsi di laurea in Matematica, Fisica, Informatica e Ingegneria.

Al lettore è richiesta una certa familiarità con alcuni argomenti di base quali l'insiemistica (insiemi, coppie ordinate, relazioni ed applicazioni) e la costruzione degli insiemi numerici (numeri naturali, interi, razionali, reali, complessi).

Ho scelto di omettere molte dimostrazioni. Quelle principali saranno disponibili, insieme ad altri contenuti aggiuntivi, sul mio sito docente, in area pubblica.

Avrei tanti ringraziamenti da fare. Mi limiterò a menzionare gli amici Maurizio Brunetti, Giovanni Cutolo, Bruno Buonomo, Francesco Della Pietra, Carlo Nitsch.

Luciano A. Lomonaco

Capitolo 1

Strutture algebriche

1.1 Richiami

Si assume che il lettore abbia già una certa familiarità con l'insiemistica (insiemi, sottoinsiemi, relazioni di appartenenza e di inclusione, implicazioni, coppie ordinate) e conosca i seguenti simboli standard:

$$\emptyset\,,\ \in\,,\ \subseteq\,,\ \subset\,,\ \Leftarrow\,,\ \Rightarrow\,,\ \Leftrightarrow\,,\ \forall\,,\ \exists\,,\ \exists!\,,$$

nonché infine gli insiemi numerici più comuni, quali ad esempio i numeri naturali $\mathbb{N}$, gli interi $\mathbb{Z}$, gli interi non negativi $\mathbb{N}_0$, i razionali $\mathbb{Q}$, i reali $\mathbb{R}$, i complessi $\mathbb{C}$, e l'insieme $\mathbb{R}[x]$ dei polinomi a coefficienti reali. Useremo nel seguito i connettivi logici $\wedge$ (*congiunzione*) e $\vee$ (*disgiunzione inclusiva*). I simboli $\mathbb{R}^+$ ed $\mathbb{R}_0^+$ indicheranno i reali positivi ed i reali non negativi.

Consideriamo un insieme non vuoto Y. Indicheremo con $\mathcal{P}(Y)$ il corrispondente insieme delle parti, ovvero

$$\mathcal{P}(Y) = \left\{\, X \;\middle|\; X \subseteq Y \,\right\}\,.$$

Diremo che il sottoinsieme $X \subseteq Y$ è una parte *propria* di Y se non coincide con Y, e cioè se $X \subset Y$.

Siano A, B due insiemi. Costruiamo un nuovo insieme, il prodotto cartesiano di A e B, indicato con il simbolo $A \times B$, costituito da tutte le coppie ordinate con prima componente in A e seconda in B, ovvero

$$A \times B = \left\{\, (a, b) \;\middle|\; a \in A\,,\ b \in B \,\right\}\,.$$

Osserviamo che

$$A = \emptyset \ \vee\ B = \emptyset \ \Rightarrow\ A \times B = \emptyset\,.$$

In modo analogo si può procedere per definire il prodotto cartesiano di più di due insiemi, usando le *terne*, *quadruple*, o, più in generale *n-ple* di elementi.

Ad esempio, se n è un numero naturale e $A_1, \ldots, A_n \neq \emptyset$, definiamo il prodotto cartesiano $A_1 \times A_2 \times \cdots \times A_n$ di tali insiemi ponendo

$$A_1 \times A_2 \times \cdots \times A_n = \left\{ (a_1, a_2, \ldots, a_n) \mid a_1 \in A_1,\ a_2 \in A_2,\ \ldots\ ,\ a_n \in A_n \right\} .$$

Se gli insiemi in questione sono tutti uguali, ad esempio $A_1 = A_2 = \cdots = A_n = A$, scriveremo A^n per indicare tale prodotto, che sarà detto n-ma *potenza cartesiana* di A. Avremo quindi

$$A^n = \underbrace{A \times A \times \cdots \times A}_{n \text{ copie}} = \left\{ (a_1, a_2, \ldots, a_n) \mid a_1, a_2, \ldots, a_n \in A \right\} .$$

Quando $n = 2$ si parla di *quadrato cartesiano*. Se $n = 1$ la potenza cartesiana A^1 è identificata con A. Introduciamo ora la nozione di corrispondenza tra due insiemi, che ci permetterà di sviluppare il concetto di relazione in un insieme e di applicazione tra due insiemi.

1.1 Definizione. *Siano $A, B \neq \emptyset$. Una corrispondenza h di A in B è una coppia $(A \times B, G)$, dove G è un sottoinsieme di $A \times B$. G si dice grafico della corrispondenza. Se la coppia $(x, y) \in G$, si dice che l'elemento y di B corrisponde all'elemento x di A, e si scrive $x\ h\ y$.*

Osserviamo esplicitamente che un elemento x di A può non avere alcun corrispondente in B, oppure averne più d'uno. Definire una corrispondenza h equivale a specificare, per ogni $x \in A$, quali elementi di B corrispondono ad x, ovvero per quali elementi $y \in B$ accade che $x\ h\ y$.

1.2 Esempio. *Sia $A = \mathbb{R}$, $B = \mathbb{R}_0^+$, e definiamo il sottoinsieme*

$$G = \left\{ (x, y) \in \mathbb{R} \times \mathbb{R}_0^+ \mid x^2 = y \right\} .$$

La corrispondenza h di $\mathbb{R}$ in $\mathbb{R}_0^+$ di cui G è il grafico è tale che $(0, 0) \in G$, ovvero $0\ h\ 0$, o ancora 0 corrisponde a 0, ed anche $(1, 1), (-1, 1) \in G$, ovvero $1\ h\ 1$ e $-1\ h\ 1$, o ancora 1 corrisponde sia a 1 che a -1.

1.3 Esempio. *Sia $A = \mathbb{R}$, $B = \mathbb{R}_0^+$, e definiamo il sottoinsieme*

$$G' = \left\{ (x, y) \in \mathbb{R} \times \mathbb{R}_0^+ \mid x = y^2 \right\} .$$

La corrispondenza h' di $\mathbb{R}$ in $\mathbb{R}_0^+$ di cui G è il grafico è tale che $(4, 2) \in G'$, ovvero $4\ h'\ 2$, o ancora 2 corrisponde a 4, ma $-2\ \not{h}'\ y$, ovvero $(-2, y) \notin G'$, per ogni $y \in \mathbb{R}_0^+$. In altri termini, non eiste alcun elemento di $\mathbb{R}_0^+$ che corrisponde all'elemento $-2 \in \mathbb{R}$.

Ci sono due casi particolari molto importanti.

1.4 Definizione. *Sia $A \neq \emptyset$. Una relazione (binaria) h in A è una corrispondenza di A in A stesso.*

1.5 Definizione. *Siano $A, B \neq \emptyset$. Un'applicazione h di A in B è una corrispondenza che gode della seguente proprietà:*

$$\forall\, x \in A\, \exists !\, y \in B \;\big|\; x\, h\, y \;.$$

Scriveremo, in tale situazione, $h(x) = y$ e diremo che y è l'immagine di x (mediante h). Spesso si utilizza il simbolo $h : A \to B$, e gli insiemi A, B sono detti rispettivamente dominio e codominio *di h.*

Studiamo ora alcune questioni riguardanti le relazioni.

1.6 Definizione. *Consideriamo una relazione h nell'insieme non vuoto A.*

- *h si dice riflessiva se per ogni $x \in A$ accade che $x\, h\, x$;*
- *h si dice simmetrica se per ogni $x, x' \in A$ vale la seguente implicazione:*

$$x\, h\, x' \;\Rightarrow\; x'\, h\, x \;;$$

- *h si dice antisimmetrica se per ogni $x, x' \in A$ vale la seguente implicazione:*

$$x\, h\, x' \;\wedge\; x'\, h\, x \;\Rightarrow\; x = x' \;;$$

- *h si dice transitiva se per ogni $x, x', x'' \in A$ vale la seguente implicazione:*

$$x\, h\, x' \;\wedge\; x'\, h\, x'' \;\Rightarrow\; x\, h\, x'' \;.$$

Siamo interessati a due tipi di relazioni.

1.7 Definizione. *La relazione h si dice* relazione d'equivalenza *se è riflessiva, simmetrica e transitiva.*

1.8 Definizione. *La relazione h si dice* relazione d'ordine *se è riflessiva, antisimmetrica e transitiva.*

Vediamo ora qualche questione riguardante le relazioni d'equivalenza.

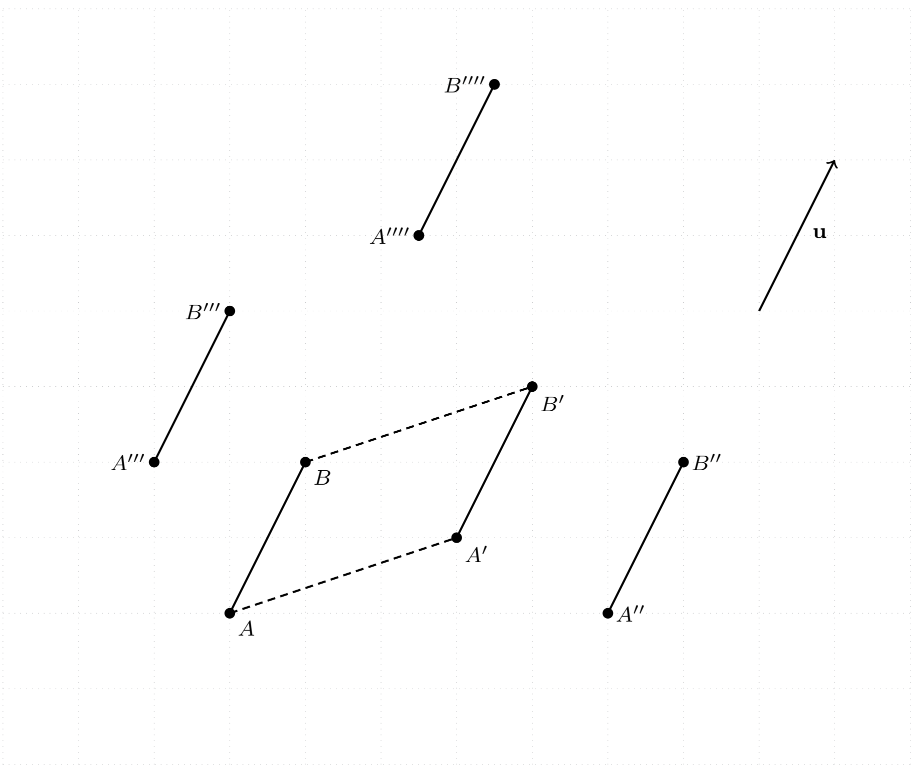
1.9 Definizione. *Sia h una relazione d'equivalenza nell'insieme non vuoto A. Per ogni $x \in A$, si dice classe d'equivalenza di x il sottoinsieme*

$$[x] = \left\{\, x' \in A \;\big|\; x\, h\, x' \,\right\} \;.$$

È chiaro che $x \in [x]$, in quanto $x\ h\ x$. Ogni elemento di $[x]$, e quindi anche x stesso, si dice *rappresentante* della classe. Si vede facilmente che l'unione delle classi d'equivalenza è proprio A, e che se $a, b \in A$ possono presentarsi due casi: $[a] = [b]$ (quando $a\ h\ b$), oppure $[a] \cap [b] = \emptyset$ (quando $a \not h\ b$). Tale situazione si esprime dicendo che le classi d'equivalenza costituiscono una partizione di A. Un simbolo usato di frequente per le relazioni d'equivalenza è $\equiv$.

1.10 DEFINIZIONE. *Sia h una relazione d'equivalenza nell'insieme non vuoto A. L'insieme delle classi d'equivalenza rispetto a tale relazione si dice insieme quoziente di A rispetto ad h, e si denota con il simbolo A/h.*

1.11 ESEMPIO. *Sia $\mathbb{E}^2$ il piano della geometria elementare e consideriamo l'insieme Σ dei segmenti orientati di $\mathbb{E}^2$. Se $\sigma \in \Sigma$ è un segmento di primo estremo A e secondo estremo B, indicheremo tale segmento con il simbolo $\sigma(A, B)$ oppure AB. Sia $\tau = \tau(A', B')$ un altro segmento. Ricordiamo che σ è equipollente a τ (in simboli $\sigma \equiv \tau$) se e solo se il quadrilatero $ABB'A'$ è un parallelogramma (eventualmente degenere). Osserviamo che la relazione di equipollenza è una relazione di equivalenza in Σ. Denotiamo con V^2 l'insieme quoziente $\Sigma/\equiv$. Un elemento di V^2 sarà detto* vettore libero ordinario *ed è una classe di segmenti equipollenti.*



Nella figura è evidenziato che $AB \equiv A'B'$, in quanto $ABB'A'$ è un parallelogramma e i segmenti AB, $A'B'$, $A''B''$, $A'''B'''$, $A''''B''''$ sono tutti equipollenti tra loro. Ad esempio il quadrilatero $ABB''''A''''$ è un parallelogramma degenere. Il vettore libero $\mathbf{u}$ *è la classe di ognuno di tali segmenti orientati, e si ha*

$$\mathbf{u} = \{\, AB, A'B', A''B'', A'''B''', A''''B'''', \dots \,\} .$$

Se ad esempio $\mathbf{u} \in V^2$ e AB è un rappresentante di $\mathbf{u}$, scriveremo $\mathbf{u} = [AB]$. Ricordiamo che invece il segmento AB è talvolta chiamato vettore applicato, *con punto di applicazione A e secondo estremo B e che per ogni vettore libero ordinario $\mathbf{u} \in V^2$ e per ogni $P \in \mathbb{E}^2$ esiste un unico $Q \in \mathbb{E}^2$ tale che $\mathbf{u} = [PQ]$.*

1.12 Esempio. *Consideriamo l'insieme numerico $\mathbb{Z}$ degli interi. In esso definiamo una relazione $\equiv$ ponendo, per ogni $a, b \in \mathbb{Z}$, $a \equiv b \iff a - b$ è un numero pari. Si verifica agevolmente che $\equiv$ è una relazione d'equivalenza, e che le classi di equivalenza sono esattamente due: il sottoinsieme $\mathbb{P}$ dei numeri pari e il sottoinsieme $\mathbb{D}$ dei numeri dispari. Pertanto l'insieme quoziente $\mathbb{Z}/\equiv$ consta di due elementi.*

1.13 Esempio. *Sia A l'insieme dei deputati della Repubblica Italiana in un fissato istante t_0. In A definiamo una relazione $\equiv$ ponendo, per ogni $a, b \in A$, $a \equiv b$ se e solo se i deputati a e b appartengono allo stesso gruppo parlamentare. La relazione $\equiv$ è d'equivalenza, e le classi di equivalenza sono proprio i gruppi parlamentari. Pertanto l'insieme quoziente $A/\equiv$ è l'insieme dei gruppi parlamentari (all'istante t_0).*

1.14 Esempio. *Sia X un insieme non vuoto e sia $n \in \mathbb{N}$, $n > 1$. Consideriamo ora l'insieme $A = X^n$, la potenza cartesiana n-ma dell'insieme X. In A definiamo una relazione σ al modo seguente. Due n-ple sono in relazione se e solo se l'una si ottiene dall'altra mediante una permutazione dell'ordine in cui gli elementi compaiono. Ad esempio, in $\mathbb{R}^3$ (cioè con $X = \mathbb{R}$, e $n = 3$), abbiamo $(3, 5, 7)\ \sigma\ (3, 7, 5)$, o anche $(1, 1, 2)\ \sigma\ (2, 1, 1)$. Osserviamo esplicitamente che una particolare permutazione è quella identica, quella cioè che lascia ogni elemento al suo posto. Quindi, ad esempio, $(3, 5, 7)\ \sigma\ (3, 5, 7)$. Si verifica che σ è una relazione d'equivalenza in X^n. Una classe di equivalenza si dice* orbita *in X^n (rispetto all'azione di permutazione), o, più spesso,* sistema *di elementi di X, di ordine n. In un sistema sono quindi rilevanti le eventuali ripetizioni, ma non è rilevante l'ordine in cui compaiono gli elementi.*

Passiamo ora alle relazioni d'ordine.

1.15 Definizione. *Sia h una relazione d'ordine nell'insieme non vuoto A. Tale relazione si dice* totale *se per ogni $x, x' \in A$ si ha che $x\ h\ x' \ \vee\ x'\ h\ x$.*

1.16 Esempio. *Sia Y un insieme non vuoto e sia $A = \mathcal{P}(Y)$, l'insieme delle parti di X. In A consideriamo la relazione $\subseteq$ di inclusione. È agevole verificare che essa è una relazione d'ordine. Non si tratta però di un ordine totale. Consideriamo ad esempio la seguente situazione. Sia $Y = \{a, b, c, d\}$ e consideriamo i sottoinsiemi $S = \{a, b, c\}$, $T = \{c, d\}$. Abbiamo che $S \nsubseteq T \ \wedge T \nsubseteq S$.*

1.17 Esempio. *Nell'insieme $A = \mathbb{R}$ dei numeri reali, consideriamo la ben nota relazione $\leq$. Essa è una relazione d'ordine e si tratta di un ordine totale. Osserviamo che possiamo definire un'analoga relazione d'ordine su $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$.*

1.18 Esempio. *Nell'insieme $A = \mathbb{R}^2$, consideriamo la relazione $\lhd$ definita come segue. Per ogni $(a,b),(c,d) \in \mathbb{R}^2$ poniamo $(a,b) \lhd (c,d)$ se e solo se $a \leq c$. Tale relazione non è una relazione d'ordine. Infatti sono verificate le proprietà riflessiva e transitiva, ma la proprietà antisimmetrica non è verificata. Ad esempio, si ha che $(1,2) \lhd (1,3)$ e $(1,3) \lhd (1,2)$, ma ovviamente $(1,2) \neq (1,3)$.*

1.19 Esempio. *Nell'insieme $A = \mathbb{R}^2$, consideriamo la relazione $\preceq$ definita come segue. Per ogni $(a,b),(c,d) \in \mathbb{R}^2$ poniamo $(a,b) \preceq (c,d)$ se e solo se $a \leq c \ \wedge \ b \leq d$. Tale relazione è una relazione d'ordine (non totale).*

A livello intuitivo possiamo prendere la relazione $\subseteq$ come prototipo di relazione d'ordine, e $\leq$ come prototipo di relazione d'ordine totale. Nelle relazioni d'ordine totale abbiamo il concetto intuitivo di elemento massimo, o minimo, dell'insieme o di una sua parte, che comunque definiremo in modo preciso. Se la relazione d'ordine non è totale, è più utile introdurre un altro concetto, quello di elemento massimale, o minimale, ricordando che se un elemento è massimo (o minimo), esso è anche massimale (minimale rispettivamente). Usiamo il simbolo $\preceq$ per indicare una generica relazione d'ordine.

1.20 Definizione. *Sia $\preceq$ una relazione d'ordine nell'insieme non vuoto A. Sia inoltre B un sottoinsieme non vuoto di A e consideriamo un suo elemento m. Diremo che m è massimale in B (rispetto alla relazione $\preceq$) se per ogni $x \in B$ si ha che $m \npreceq x$. L'elemento m sarà invece detto minimale in B se per ogni $x \in B$ si ha che $x \npreceq m$.*

1.21 Definizione. *Sia $\preceq$ una relazione d'ordine nell'insieme non vuoto A. Sia inoltre B un sottoinsieme non vuoto di A e consideriamo un suo elemento m. Diremo che m è massimo in B (rispetto alla relazione $\preceq$) se per ogni $x \in B$ si ha che $x \preceq m$. L'elemento m sarà invece detto minimo in B se per ogni $x \in B$ si ha che $m \preceq x$.*

Non sempre esiste un massimo (minimo risp.), ma quando ciò accade, tale elemento è unico. Se la relazione d'ordine è totale, un elemento massimale è anche massimo, ed un elemento minimale è anche minimo.

1.22 Esempio. *Consideriamo l'insieme $Y = \{a,b,c,d\}$, sia $A = \mathcal{P}(Y)$. Sia B il sottoinsieme di A costituito da tutte e sole le parti proprie di Y:*

$$B = A - \{Y\} = \left\{ Z \in \mathcal{P}(Y) \;\middle|\; Z \neq Y \right\} = \left\{ Z \in \mathcal{P}(Y) \;\middle|\; Z \subset Y \right\} .$$

Consideriamo in B la relazione $\subseteq$ di inclusione. Il sottoinsieme $S = \{a,b,c\}$ di A è un elemento di B, ed è massimale in B, ma non massimo. Osserviamo che ogni sottoinsieme di Y con tre elementi gode della stessa proprietà.

Affrontiamo ora lo studio delle applicazioni. Consideriamo quindi un'applicazione $f : A \to B$. Se $x \in A$ e $y = f(x) \in B$, scriveremo anche, talvolta, $x \mapsto y$. Il sottoinsieme

$$\operatorname{im} f = f(A) := \{ \, y \in B \mid \exists x \in A \mid y = f(x) \, \} \subseteq B$$

si dice *immagine* di f. Può accadere che $\operatorname{im} f = B$ oppure che $\operatorname{im} f \subset B$. Inoltre, se $y \in \operatorname{im} f$, può accadere che sia unico l'elemento $x \in A$ tale che $y = f(x)$, oppure possono esistere vari elementi $x, x', x'', \dots \in A$ tali che

$$y = f(x) = f(x') = f(x''), \dots \ ,$$

cioè y può essere immagine di vari elementi del dominio. Per ogni $y \in B$ è definito un sottoinsieme di A, detto controimmagine di y mediante f e denotato con il simbolo $f^{-1}(y)$, al modo seguente:

$$f^{-1}(y) = \{ \, x \in A \mid f(x) = y \, \} \ .$$

1.23 Osservazione. *È bene tener presente che*

- *Il sottoinsieme $f^{-1}(y)$ può essere vuoto;*
- $f^{-1}(y) \neq \emptyset \iff y \in \operatorname{im} f$;
- *Se $f^{-1}(y) \neq \emptyset$, esso può essere costituito da uno o più elementi del dominio A.*

Più in generale, se $Y \subseteq B$ è un sottoinsieme del codominio, si definisce controimmagine di Y il sottoinsieme

$$f^{-1}(Y) = \{ \, x \in A \mid f(x) \in Y \, \} \subseteq A \ .$$

1.24 Definizione. *Sia $f : A \to B$ un'applicazione. Essa si dice iniettiva se per ogni $x', x'' \in A$ tali che $x' \neq x''$ si abbia che $f(x') \neq f(x'')$.*

Ciò si può esprimere anche dicendo che per ogni $y \in \operatorname{im} f$ esiste un unico elemento $x \in A$ tale che $y = f(x)$, oppure che vale la seguente implicazione:

$$f(x') = f(x'') \quad \Longrightarrow \quad x' = x'' \ .$$

1.25 Definizione. *Sia $f : A \to B$ un'applicazione. Essa si dice suriettiva se $\operatorname{im} f = B$, ovvero se per ogni $y \in B$ esiste almeno un elemento $x \in A$ tale che $y = f(x)$.*

1.26 Definizione. *Sia $f : A \to B$ un'applicazione. Essa si dice biettiva se è sia iniettiva che suriettiva.*

1.27 Esempio. *Sia A l'insieme degli abitanti di una certa città, ad esempio Napoli, in un certo istante. Definiamo un'applicazione $f : A \to \mathbb{N}_0$ al modo seguente. Se a è un abitante* *di Napoli, sia $f(a)$ la sua età. È chiaro che f è un'applicazione, in quanto ogni abitante ha sicuramente una età ben definita. f non è iniettiva, in quanto molti abitanti di Napoli sono coetanei, e non è suriettiva. Se scegliamo infatti un numero naturale m abbastanza grande, ad esempio $m = 1000$, certamente non potremo trovare un abitante di Napoli con quell'età, ovvero $\nexists\, a \in A$ tale che $f(a) = 1000$. Quindi $1000 \notin \operatorname{im} f$.*

1.28 ESEMPIO. *Definiamo un'applicazione $f : \mathbb{R} \to \mathbb{R}$ ponendo $x \mapsto x^2$, per ogni $x \in \mathbb{R}$. f non è iniettiva (ad esempio abbiamo che $f(2) = f(-2) = 4$) e neppure suriettiva (ad esempio scelto -1 nel codominio $\mathbb{R}$, è ben noto che non esiste alcun elemento del dominio $\mathbb{R}$ che abbia -1 come quadrato). Osserviamo che*

$$f^{-1}(0) = \{0\} \quad ; \quad f^{-1}(1) = \{1, -1\} \quad ; \quad f^{-1}(-1) = \emptyset \ .$$

Possiamo ora definire due nuove applicazioni che agiscono come f, ma su insiemi diversi. Definiamo

$$g : \mathbb{R}_0^+ \longrightarrow \mathbb{R} \quad ; \quad h : \mathbb{R}_0^+ \longrightarrow \mathbb{R}_0^+$$

ponendo, come nel caso di f, $g(x) = x^2$ e $h(x) = x^2$, per ogni $x \in \mathbb{R}_0^+$. Si verifica agevolmente che g è un'applicazione iniettiva, ma non suriettiva, mentre h è sia iniettiva che suriettiva, e quindi è biettiva.

Questo esempio suggerisce la seguente

1.29 DEFINIZIONE. *Sia $f : A \to B$ un'applicazione. Scelto un qualunque sottoinsieme non vuoto $\bar{A}$ del dominio A e un sottoinsieme $\bar{B}$ del codominio B tale che si abbia $\operatorname{im} f \subseteq \bar{B}$, l'applicazione $\bar{f} : \bar{A} \to \bar{B}$ definita ponendo $\bar{f}(x) = f(x)$, per ogni $x \in \bar{A}$, si dice restrizione di f.*

Se $\bar{A} \subset A$ e $\bar{B} = B$, la restrizione $\bar{f}$ si indica anche con il simbolo $f_{|\bar{A}}$. Nell'esempio precedente, g ed h sono entrambe restrizioni di f, e in particolare $g = f_{|\mathbb{R}_0^+}$.

1.30 ESEMPIO. *Sia A un qualunque insieme non vuoto. L'applicazione $id_A : A \to A$ definita ponendo $id_A(x) = x$ per ogni $x \in A$ si dice applicazione identica di A e si indica spesso con il simbolo id. Se $X \subseteq A$, la restrizione dell'identità al sottoinsieme X sul dominio si dice applicazione d'inclusione, o più semplicemente inclusione, di X in A, e si scrive talvolta $id_{|X} : X \hookrightarrow A$.*

Concludiamo questa breve esposizione sulle applicazioni introducendo i concetti di applicazione composta ed inversa.

1.31 DEFINIZIONE. *Siano A, B, C degli insiemi non vuoti, e siano $f : A \to B$ e $g : B \to C$ due applicazioni. Definiamo una nuova applicazione $h : A \to C$ ponendo, per ogni $x \in A$, $h(x) = g\big(f(x)\big)$. L'applicazione h si dice composta di f e g e si scrive $h = g \circ f$, o anche $h = gf$.*

Per evidenziare che h è la composta di f e g, si scrive talvolta

$$h : A \xrightarrow{f} B \xrightarrow{g} C \ .$$

Consideriamo ora il caso particolare in cui $A = C$.

1.32 Definizione. *Siano $f : A \to B$ e $g : B \to A$ due applicazioni. Diremo che g è inversa di f se accade che $gf = id_A$ e $fg = id_B$.*

In altri termini deve accadere che, per ogni $x \in A$, si abbia che $gf(x) = x$ e che, per ogni $y \in B$, si abbia $fg(y) = y$. Osserviamo che non sempre esiste un'applicazione inversa di un'applicazione data f, ma se esiste è unica, e si indica con il simbolo f^{-1}. Va sottolineato che, nel caso in cui f sia dotata di inversa, per ogni elemento y del codominio di f il simbolo $f^{-1}(y)$ indica sia l'immagine dell'elemento y mediante f^{-1} (che è un elemento del dominio di f) che la controimmagine dell'elemento y mediante f (che è un sottoinsieme del dominio di f costituito da un solo elemento, ovvero un *singleton*), un chiaro abuso di notazione che è però di uso comune. Si dovrà capire dal contesto il significato del simbolo.

1.33 Esempio. *Consideriamo l'applicazione $h : \mathbb{R}_0^+ \longrightarrow \mathbb{R}_0^+$ definita, come sopra, ponendo $h(x) = x^2$, per ogni numero reale non negativo x, e l'applicazione $k : \mathbb{R}_0^+ \longrightarrow \mathbb{R}_0^+$ definita ponendo, per ogni numero reale non negativo x, $k(x) = \sqrt{x}$. Si verifica agevolmente che k è l'inversa di h, cioè $k = h^{-1}$. Si scriverà quindi, ad esempio, $h^{-1}(4) = 2$ oppure $h^{-1}(4) = \{2\}$ a seconda che si voglia indicare l'immagine tramite h^{-1} di 4, oppure la controimmmagine di 4 mediante h.*

Le applicazioni dotate di inversa si caratterizzano al modo seguente.

1.34 Proposizione. *Un'applicazione è dotata di inversa se e solo se essa è biettiva.*

Dimostrazione. Sia $f : A \to B$ un'applicazione. Supponiamo che f sia biettiva. Definiamo un'applicazione $g : B \to A$ al modo seguente. Per ogni $y \in B$, osserviamo che esiste un unico elemento $x \in A$ tale che $f(x) = y$. L'esistenza è dovuta al fatto che f è suriettiva, l'unicità al fatto che f è iniettiva. Poniamo allora $g(y) = x$. È agevole verificare che g è l'inversa di f. Supponiamo ora che f sia dotata di inversa, diciamo g. Per ogni $y \in B$ consideriamo l'elemento $x = g(y) \in A$. Abbiamo che

$$f(x) = fg(y) = id_B(y) = y \ ,$$

e quindi f è suriettiva. Siano ora $x', x'' \in A$ e sia $f(x') = f(x'') = y$. Allora

$$\begin{aligned} g(y) &= gf(x') = id_A(x') = x' \\ &= gf(x'') = id_A(x'') = x'' \end{aligned}$$

e quindi f è anche iniettiva. □

1.2 Generalità sulle strutture algebriche

In alcuni insiemi, come ad esempio quelli numerici, o dei vettori liberi o applicati, siamo abituati a considerare delle operazioni (addizione, moltiplicazione, ...). A seconda delle operazioni che consideriamo, e delle loro proprietà, possiamo sviluppare certi strumenti di calcolo. Precisiamo dunque tali concetti introducendo le nozioni di operazione (binaria) interna ed esterna, e di struttura algebrica.

1.35 Definizione. *Sia S un insieme non vuoto. Una operazione (binaria) interna di S è un'applicazione $f : S \times S \to S$.*

È d'uso comune associare un simbolo, ad esempio $*$, a tale operazione e scrivere $a * b$ invece di $f(a, b)$, per ogni $(a, b) \in S \times S$.

1.36 Definizione. *L'operazione interna $*$ di S è associativa se*

$$(a * b) * c = a * (b * c) \quad \forall\ a, b, c \in S\ .$$

1.37 Esempio. *L'addizione $+$ e la moltiplicazione $\cdot$ negli insiemi numerici $\mathbb{N}$, $\mathbb{N}_0$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ sono operazioni interne associative. La divisione $:$ non è una operazione (non si può dividere per 0 in $\mathbb{N}_0, \mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$ e si può effettuare la divisione tra i numeri naturali o interi relativi m, n se e solo se m è multiplo di n). In $\mathbb{Q} - \{0\}$ è sempre possibile effettuare la divisione, e pertanto in tale insieme la divisione è una operazione interna, ma è facile verificare che essa non è associativa. Infatti, ad esempio,*

$$(16 : 4) : 2 \neq 16 : (4 : 2)\ .$$

1.38 Definizione. *Sia S un insieme non vuoto. Una struttura algebrica (con una operazione interna) con sostegno S è una coppia $(S; *)$, dove $*$ è una operazione interna di S.*

Quasi sempre, con abuso di notazione, si scriverà S sia per indicare l'insieme di sostegno che per indicare la struttura.

1.39 DEFINIZIONE. *Sia S una struttura algebrica (con operazione interna $*$) e sia $T \subseteq S$. Diremo che T è una parte stabile di S se per ogni $a, b \in T$ si ha che $a * b \in T$.*

Si può facilmente verificare che l'intersezione di parti stabili è ancora una parte stabile.

1.40 DEFINIZIONE. *Sia S una struttura algebrica (con operazione interna $*$). Se l'operazione $*$ è associativa, la struttura si dice semigruppo.*

1.41 DEFINIZIONE. *Sia S una struttura algebrica (con operazione interna $*$) e sia $u \in S$. L'elemento u si dice neutro a destra se $x * u = x$ per ogni $x \in S$, neutro a sinistra se $u * y = y$ per ogni $y \in S$. Infine, u si dice neutro se lo è sia a destra che a sinistra.*

1.42 PROPOSIZIONE. *Sia S una struttura algebrica (con operazione interna $*$). Se $u' \in S$ è neutro a sinistra e $u'' \in S$ è neutro a destra, si ha che $u' = u''$. In particolare quindi esiste al più un elemento neutro.*

Dimostrazione. Poiché u' è neutro a sinistra, si ha che $u' * u'' = u''$. Ma poiché u'' è neutro a destra, si ha che $u' * u'' = u'$. Quindi $u' = u''$. $\square$

1.43 DEFINIZIONE. *Un semigruppo dotato di elemento neutro si dice monoide.*

1.44 DEFINIZIONE. *Sia S un monoide e sia u il suo elemento neutro. Un elemento $x \in S$ è simmetrizzabile a sinistra se esiste $x' \in S$ tale che $x' * x = u$, a destra se esiste $x'' \in S$ tale che $x * x'' = u$. Infine x è simmetrizzabile se esiste $y \in S$ tale che $y * x = u = x * y$. Gli elementi x', x'', y si dicono rispettivamente simmetrico a sinistra, simmetrico a destra, simmetrico di x.*

1.45 PROPOSIZIONE. *Sia S un monoide e sia $x \in S$. Se x' è un simmetrico sinistro di x e x'' è un simmetrico destro di x, allora $x' = x''$.*

Dimostrazione. Denotiamo con u l'elemento neutro del monoide. Abbiamo che

$$\begin{aligned} x' &= x' * u \\ &= x' * (x * x'') \\ &= (x' * x) * x'' \quad \text{(per l'associatività)} \\ &= u * x'' \\ &= x'' \,, \end{aligned}$$

cioè $x' = x''$. □

Osserviamo che in generale il simmetrico sinistro (o destro), se esiste, non è unico. La proposizione precedente ci assicura però che il simmetrico, se esiste, è unico.

1.46 Esempio. *Consideriamo gli insiemi numerici $\mathbb{N}$, $\mathbb{N}_0$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ con l'addizione. Tale operazione è associativa e quindi tutte le strutture considerate sono semigruppi. Lo 0 è l'elemento neutro, e quindi $\mathbb{N}_0$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ sono monoidi. In $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ ogni elemento è dotato di simmetrico (l'opposto). Consideriamo ora $\mathbb{N}$, $\mathbb{N}_0$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ con la moltiplicazione, anch'essa associativa. In questo caso 1 è l'elemento neutro, e tutte le strutture sono monoidi. In $\mathbb{N}$ ed $\mathbb{N}_0$ l'unico elemento simmetrizzabile è l'elemento neutro, che è simmetrico di se stesso, come avviene in generale. In $\mathbb{Z}$ è simmetrizzabile anche -1, che è simmetrico di se stesso. In $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ ogni elemento non nullo è simmetrizzabile, ed il simmetrico è l'inverso o reciproco. In $\mathbb{Q}^+$, $\mathbb{R}^+$ ogni elemento è simmetrizzabile. Come già osservato, la divisione non può considerarsi come operazione in $\mathbb{N}$, $\mathbb{N}_0$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ in quanto, dati due elementi a, b non sempre è possibile effettuare la divisione $a : b$. In particolare, in $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ è possibile effettuare tale operazione se e solo se $b \neq 0$. In $\mathbb{Q}^+$, $\mathbb{R}^+$, $\mathbb{Q} - \{0\}$, $\mathbb{R} - \{0\}$, $\mathbb{C} - \{0\}$ la divisione è una operazione, ma non è associativa.*

In analogia con il caso numerico, quando si usa il simbolo $+$, l'operazione è detta addizione (notazione additiva) ed il simmetrico di un elemento x è detto opposto e denotato $-x$. Quando invece si usa il simbolo $\cdot$ l'operazione è detta moltiplicazione (notazione moltiplicativa) ed il simmetrico di un elemento x è detto inverso o reciproco e denotato con x^{-1} o anche $1/x$. È frequente che in notazione moltiplicativa il simbolo $\cdot$ sia omesso (e in tal caso si dice che l'operazione è denotata per giustapposizione).

1.47 Definizione. *Un monoide G è un gruppo se ogni suo elemento è simmetrizzabile.*

1.48 Definizione. *Una operazione $*$ in un insieme non vuoto S si dice commutativa se si ha che $a * b = b * a$ per ogni $a, b \in S$.*

1.49 Definizione. *Un gruppo G la cui operazione è commutativa prende il nome di gruppo abeliano.*

1.50 Esempio. *Consideriamo gli insiemi numerici* $\mathbb{N}$, $\mathbb{N}_0$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ *con l'addizione. Per quanto già osservato, vediamo che* $\mathbb{N}$, $\mathbb{N}_0$ *non sono gruppi, mentre* $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ *sono gruppi abeliani. Se consideriamo invece le strutture numeriche moltiplicative, osserviamo che lo* 0 *non è mai invertibile. Esempi di gruppi (abeliani) moltiplicativi sono* $\mathbb{Q}^+$, $\mathbb{R}^+$, $\mathbb{Q}-\{0\}$, $\mathbb{R}-\{0\}$, $\mathbb{C}-\{0\}$.

1.51 Esempio. *Consideriamo l'insieme*

$$\mathcal{H} = \{\,1, -1, i, -i, j, -j, k, -k\,\}\ ,$$

ed in esso definiamo una moltiplicazione mediante la seguente tabella

$\cdot$	1	-1	i	$-i$	j	$-j$	k	$-k$
1	1	-1	i	$-i$	j	$-j$	k	$-k$
-1	-1	1	$-i$	i	$-j$	j	$-k$	k
i	i	$-i$	-1	1	k	$-k$	$-j$	j
$-i$	$-i$	i	1	-1	$-k$	k	j	$-j$
j	j	$-j$	$-k$	k	-1	1	i	$-i$
$-j$	$-j$	j	k	$-k$	1	-1	$-i$	i
k	k	$-k$	j	$-j$	$-i$	i	-1	1
$-k$	$-k$	k	$-j$	j	i	$-i$	1	-1

dove il prodotto tra due elementi x *e* y *si ottiene selezionando* x *sulla prima colonna e* y *sulla prima riga e determinando l'elemento della tabella sull'intersezione della riga di* x *e della colonna di* y. *La struttura* $(\mathcal{H};\cdot)$ *così ottenuta è un gruppo non abeliano: il gruppo dei quaternioni.*

Possiamo considerare anche strutture con più operazioni. In particolare siamo interessati alle strutture con due operazioni interne. È d'uso comune denotare tali operazioni con i simboli $+, \cdot$ (addizione e moltiplicazione).

1.52 Definizione. *Sia* A *un insieme non vuoto e siano* $+, \cdot$ *due sue operazioni interne. La terna* $(A; +, \cdot)$ *si dice anello se*

(i) *la struttura* $(A, +)$ *è un gruppo abeliano;*

(ii) *la struttura* $(A, \cdot)$ *è un semigruppo;*

(iii) $x \cdot (a' + a'') = x \cdot a' + x \cdot a''$, *per ogni* $x, a', a'' \in A$;

(iv) $(a' + a'') \cdot y = a' \cdot y + a'' \cdot y$, *per ogni* $a', a'', y \in A$.

Le ultime due proprietà sono note come *distributività* (a sinistra e a destra) della moltiplicazione rispetto all'addizione. Abbiamo qui implicitamente adottato la convenzione che la moltiplicazione ha la precedenza sull'addizione, evitando così di dover usare alcune ulteriori parentesi. Quando la moltiplicazione è commutativa, l'anello è detto commutativo. Se poi esiste l'elemento neutro rispetto alla moltiplicazione, l'anello è detto unitario.

1.53 Esempio. *Consideriamo gli insiemi numerici $\mathbb{N}$, $\mathbb{N}_0$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ con l'addizione e la moltiplicazione. Per quanto già osservato, vediamo che $\mathbb{N}$, $\mathbb{N}_0$ non sono anelli, mentre $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ sono anelli commutativi unitari. Il sottoinsieme $\mathbb{P} \subset \mathbb{Z}$ dei numeri pari è stabile rispetto ad entrambe le operazioni ed è anch'esso un anello commutativo, ma non è unitario.*

1.54 Esempio. *L'insieme $\mathbb{R}[x]$ dei polinomi nell'indeterminata x a coefficienti reali, con le usuali operazioni di addizione e moltiplicazione è un ulteriore esempio notevole di anello commutativo unitario. I coefficienti possono anche essere presi in un qualunque altro anello, ad esempio gli interi, i razionali o i complessi. Si può considerare anche l'insieme $\mathbb{R}[x_1, x_2, \ldots, x_n]$ dei polinomi a coefficienti reali in n indeterminate. Anche questo insieme ha una struttura di anello, con le usuali operazioni di addizione e moltiplicazione tra polinomi.*

1.55 Definizione. *Un anello unitario A di dice corpo se, privato dello 0, è un gruppo rispetto alla moltiplicazione. Se poi la moltiplicazione è anche commutativa, la struttura prende il nome di campo.*

Tradizionalmente, gli elementi di un campo si dicono scalari.

1.56 Esempio. *Gli anelli commutativi unitari $\mathbb{Z}$ e $\mathbb{R}[x]$ non sono campi. Gli insiemi numerici dei razionali, reali e complessi sono invece altrettanti esempi di struttura di campo.*

1.57 Esempio. *Consideriamo il sottoinsieme $A \subset \mathbb{R}$ così definito:*

$$A = \left\{ a + b\sqrt{2} \;\middle|\; a, b \in \mathbb{Q} \right\} .$$

È agevole verificare che A è stabile rispetto ad entrambe le operazioni ed è un campo, che contiene propriamente il campo dei numeri razionali ed è contenuto propriamente nel campo reale. Tale campo si dice estensione di $\mathbb{Q}$ mediante $\sqrt{2}$.

1.58 Esempio. *Consideriamo l'insieme $\mathbb{H}$ delle espressioni formali del tipo $\mathbf{w} = a + bi + cj + dk$ dove $a, b, c, d \in \mathbb{R}$. Definiamo due operazioni $+, \cdot$ come segue. Se $\mathbf{w}', \mathbf{w}'' \in \mathbb{H}$, ad esempio $\mathbf{w}' = a' + b'i + c'j + d'k$, $\mathbf{w}'' = a'' + b''i + c''j + d''k$, poniamo*

$$\mathbf{w}' + \mathbf{w}'' = a' + a'' + (b' + b'')i + (c' + c'')j + (d' + d'')k \ ; \quad \mathbf{w}' \cdot \mathbf{w}'' = \alpha + \beta i + \gamma j + \delta k$$

dove

$$\alpha = a'a'' - b'b'' - c'c'' - d'd'' \ , \ \beta = a'b'' + b'a'' + c'd'' - d'c'' \ ,$$
$$\gamma = a'c'' - b'd'' + c'a'' + d'b'' \ , \ \delta = a'd'' + bc'' - c'b'' + d'a'' \ .$$

Osserviamo che la moltiplicazione è definita formalmente rispettando la tabella dell'Esempio 1.51 sui simboli i, j, k e la consueta regola dei segni, e imponendo la distributività. La verifica del fatto che $(\mathbb{H}; +)$ è un gruppo abeliano è agevole. La verifica del fatto che $(\mathbb{H} - \{0\}; \cdot)$ è un gruppo richiede qualche calcolo in più. La costruzione della struttura $(\mathbb{H}; +, \cdot)$ ricalca quella dei numeri complessi: i, j, k sono tutte radici di -1. La moltiplicazione non è commutativa. Infatti, ad esempio, $ij = k$, $ji = -k$. Pertanto $(\mathbb{H}; +, \cdot)$ è un corpo, ma non un campo, ed è noto come il corpo dei quaternioni *introdotto dal matematico irlandese W. R. Hamilton.*

Sia ora A una struttura con una operazione di addizione associativa, e supponiamo di voler sommare non necessariamente due ma più elementi, ad esempio gli n elementi $x_1, x_2, \ldots, x_n$, dove n è un qualunque numero naturale. Useremo allora, talvolta, la notazione di *sommatoria* $\sum$, scegliendo un qualunque simbolo, ad esempio i, detto *indice di sommatoria*, che varia da 1 ad n, e scriveremo

$$\sum_{i=1}^{n} x_i$$

invece di

$$x_1 + x_2 + \cdots + x_n \ .$$

Naturalmente si può utilizzare una notazione analoga per altre operazioni. Ad esempio, nel caso della moltiplicazione si utilizza il simbolo $\prod$ e si parla di *produttoria*.

1.59 Esempio. *Se vogliamo considerare la somma dei primi* 1000 *numeri naturali, possiamo scrivere*

$$1 + 2 + 3 + 4 + 5 + 6 + 7 + 8 + 9 + \cdots + 999 + 1000 = \sum_{n=1}^{1000} n = 500500 \ .$$

Nel prossimo capitolo introdurremo una nuova struttura algebrica nella quale si utilizza un tipo di operazione diverso da quello considerato finora. Si tratta di una operazione *esterna*, nel senso della seguente definizione.

1.60 Definizione. *Siano S, K insiemi non vuoti. Una operazione esterna μ in S, con operatori in K, è un'applicazione*

$$\mu : K \times S \longrightarrow S \ .$$

Gli elementi di K si dicono scalari.

Scelto un simbolo, ad esempio $\cdot$ per rappresentare l'operazione μ, scriveremo, come per le operazioni interne, $a \cdot x$ invece di $\mu(a, x)$, per ogni $a \in S$, $x \in K$, ovvero, come si suol dire, useremo la notazione moltiplicativa.

Capitolo 2

Spazi vettoriali

In questo capitolo introdurremo e studieremo una struttura algebrica che giocherà un ruolo chiave nella nostra esposizione: la struttura di *spazio vettoriale* su un campo.

2.1 Generalità, dipendenza lineare

2.1 DEFINIZIONE. *Sia $\mathbb{K}$ un campo e V un insieme non vuoto. In V siano definite due operazioni: una interna, detta addizione e denotata con il simbolo $+$, ed una esterna, con operatori in $\mathbb{K}$, detta moltiplicazione e denotata con il simbolo $\cdot$. La struttura $(V;+,\cdot)$ è uno spazio vettoriale sul campo $\mathbb{K}$, o anche $\mathbb{K}$-spazio, se sono verificati i seguenti assiomi:*

(1) *la struttura $(V;+)$ è un gruppo abeliano;*

(2) $\alpha \cdot (\beta \cdot \mathbf{v}) = (\alpha\beta) \cdot \mathbf{v}, \quad \forall\, \alpha, \beta \in \mathbb{K},\ \forall\, \mathbf{v} \in V;$

(3) $1 \cdot \mathbf{v} = \mathbf{v}, \quad \forall\, \mathbf{v} \in V;$

(4) $(\alpha + \beta) \cdot \mathbf{v} = \alpha \cdot \mathbf{v} + \beta \cdot \mathbf{v}, \quad \forall\, \alpha, \beta \in \mathbb{K},\ \forall\, \mathbf{v} \in V;$

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

L'assioma (ii) è detto *associatività mista*. Gli assiomi (iv) e (v) esprimono proprietà di distributività. Il lettore noterà che si usa lo stesso simbolo $+$ per indicare l'addizione in $\mathbb{K}$ e in V. Come in ogni struttura algebrica, l'insieme V prende il nome di sostegno della struttura $(V;+,\cdot)$. Per semplicità scriveremo V anche per indicare la terna $(V;+,\cdot)$. Gli elementi di V sono detti *vettori*, quelli di $\mathbb{K}$ *scalari*. Il simbolo $\mathbf{0}$ indicherà sempre l'elemento neutro di V rispetto all'addizione, detto anche *vettore nullo*; inoltre, per ogni vettore $\mathbf{v} \in V$, scriveremo $-\mathbf{v}$ per indicare il vettore opposto. Dagli assiomi ora enunciati, discendono varie proprietà, che andiamo ad esaminare.

2.2 Proposizione. *Sia V uno spazio vettoriale. Valgono le seguenti proprietà:*

(i) *per ogni $\mathbf{v} \in V$ si ha che $0 \cdot \mathbf{v} = \mathbf{0}$;*

(ii) *per ogni $\mathbf{v} \in V$ si ha che $(-1) \cdot \mathbf{v} = -\mathbf{v}$;*

(iii) *per ogni $\lambda \in \mathbb{K}$ si ha che $\lambda \cdot \mathbf{0} = \mathbf{0}$;*

(iv) *se $\lambda \cdot \mathbf{v} = \mathbf{0}$ e $\lambda \neq 0$ si ha che $\mathbf{v} = \mathbf{0}$.*

Dimostrazione. (i) Per ogni $\mathbf{v} \in V$ si ha che

$$0 \cdot \mathbf{v} = (0+0) \cdot \mathbf{v} = 0 \cdot \mathbf{v} + 0 \cdot \mathbf{v} \,,$$

e quindi, sommando ad entrambi i membri il vettore $-(0 \cdot \mathbf{v})$, otteniamo

$$\begin{aligned} \mathbf{0} &= -(0 \cdot \mathbf{v}) + 0 \cdot \mathbf{v} = -(0 \cdot \mathbf{v}) + (0 \cdot \mathbf{v} + 0 \cdot \mathbf{v}) \\ &= (-(0 \cdot \mathbf{v}) + 0 \cdot \mathbf{v}) + 0 \cdot \mathbf{v} = \mathbf{0} + 0 \cdot \mathbf{v} \\ &= 0 \cdot \mathbf{v} \;. \end{aligned}$$

(ii) Per ogni $\mathbf{v} \in V$ si ha che

$$\mathbf{0} = 0 \cdot \mathbf{v} = ((-1)+1) \cdot \mathbf{v} = (-1) \cdot \mathbf{v} + 1 \cdot \mathbf{v} = (-1) \cdot \mathbf{v} + \mathbf{v} \;,$$

e quindi, sommando ad entrambi i membri il vettore $-\mathbf{v}$ otteniamo

$$\begin{aligned} (-\mathbf{v}) &= \mathbf{0} + (-\mathbf{v}) = ((-1) \cdot \mathbf{v} + \mathbf{v}) + (-\mathbf{v}) \\ &= (-1) \cdot \mathbf{v} + (\mathbf{v} + (-\mathbf{v})) = (-1) \cdot \mathbf{v} + \mathbf{0} \\ &= (-1) \cdot \mathbf{v} \;. \end{aligned}$$

(iii) Per ogni $\lambda \in \mathbb{K}$ si ha che

$$\lambda \cdot \mathbf{0} = \lambda \cdot (\mathbf{0} + \mathbf{0}) = \lambda \cdot \mathbf{0} + \lambda \cdot \mathbf{0} \,,$$

e quindi, sommando ad entrambi i membri il vettore $-(\lambda \cdot \mathbf{0})$, otteniamo l'asserto, imitando la dimostrazione del primo punto.

(iv) Sia λ uno scalare non nullo, sia $\mathbf{v}$ un vettore qualunque, e supponiamo che $\lambda \cdot \mathbf{v} = \mathbf{0}$. Allora

$$\mathbf{v} = 1 \cdot \mathbf{v} = (\lambda^{-1}\lambda) \cdot \mathbf{v} = \lambda^{-1} \cdot (\lambda \cdot \mathbf{v}) = \lambda^{-1} \cdot \mathbf{0} = \mathbf{0} \;.$$

□

Il lettore osservi che in ogni uguaglianza della dimostrazione appena svolta si utilizza un opportuno assioma della Definizione 2.1. Quando si opera come nella dimostrazione dei punti (i) e (ii) si dice che si procede per *cancellazione*. D'ora in avanti ometteremo quasi sempre il simbolo $\cdot$ nella notazione di moltiplicazione esterna, così come è d'uso comune fare per la moltipicazione interna di $\mathbb{K}$, ovvero, come si suol dire, l'operazione sarà indicata per giustapposizione.

2.3 Esempio. **Spazio banale.** *Sia V un singleton, ovvero un insieme con un solo elemento, e denotiamo tale elemento con il simbolo $\mathbf{0}$. Con le ovvie operazioni banali di addizione interna e di moltiplicazione esterna, la struttura $(V;+,\cdot)$ è uno spazio vettoriale su $\mathbb{K}$, detto* spazio vettoriale banale, o nullo, *e $\mathbf{0}$ è il suo vettore nullo.*

2.4 Esempio. **Spazio numerico.** *Per ogni $n \in \mathbb{N}$ consideriamo la n-ma potenza cartesiana $\mathbb{K}^n$ del campo $\mathbb{K}$, ovvero l'insieme*

$$\mathbb{K}^n = \{ (a_1, \dots, a_n) \mid a_i \in \mathbb{K},\ \forall\, i \}$$

delle n-ple di elementi di $\mathbb{K}$. In $\mathbb{K}^n$ si definiscono le operazioni di addizione (interna) e di moltiplicazione (esterna) ponendo

$$(a_1, \dots, a_n) + (b_1, \dots, b_n) = (a_1 + b_1, \dots, a_n + b_n) \ ;$$
$$\lambda \cdot (a_1, \dots, a_n) = (\lambda a_1, \dots, \lambda a_n)$$

per ogni $(a_1, \dots, a_n), (b_1, \dots, b_n) \in \mathbb{K}^n$ e per ogni $\lambda \in \mathbb{K}$. Si verifica agevolmente che $(\mathbb{K}^n;+,\cdot)$ è uno spazio vettoriale su $\mathbb{K}$. In particolare se $n = 1$ lo spazio vettoriale $\mathbb{K}^1$ può identificarsi con il campo $\mathbb{K}$, ovvero $\mathbb{K}$ può vedersi come spazio vettoriale su se stesso. $\mathbb{K}^n$ prende il nome di n-spazio vettoriale numerico *su $\mathbb{K}$. Un suo elemento sarà detto vettore numerico di ordine n. Il vettore nullo di $\mathbb{K}^n$ è manifestamente il vettore $\mathbf{0} = (0, \dots, 0)$. Definiamo un'applicazione*

$$\mu : \mathbb{K}^n \times \mathbb{K}^n \longrightarrow \mathbb{K}$$

al modo seguente. Se $\mathbf{a} = (a_1, \dots, a_n)$, $\mathbf{b} = (b_1, \dots, b_n)$, poniamo

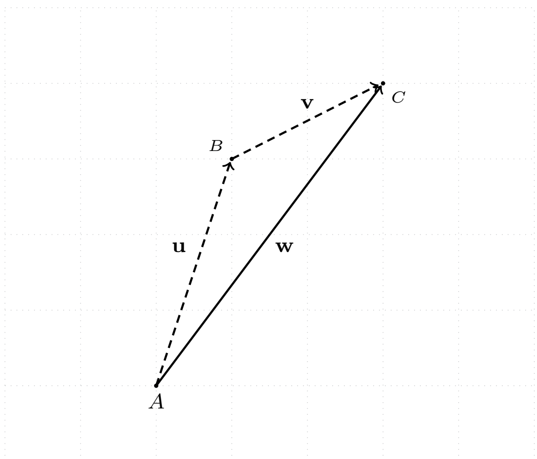
$$\mu(\mathbf{a}, \mathbf{b}) = \sum_{i=1}^{n} a_i b_i \ .$$

L'applicazione μ non è una operazione nel senso già acquisito, ma è comunque nota come prodotto scalare standard *in $\mathbb{K}^n$. Di solito si scrive $\mathbf{a} \cdot \mathbf{b}$, o anche $\langle \mathbf{a}, \mathbf{b} \rangle$, invece di $\mu(\mathbf{a}, \mathbf{b})$.*

2.5 Esempio. **Vettori liberi.** *Consideriamo il piano $\mathbb{E}^2$ della geometria elementare e l'insieme quoziente dei vettori liberi ordinari V^2 dell'Esempio 1.11. Ricordiamo nuovamente che per ogni vettore libero ordinario $\mathbf{u} \in V^2$ e per ogni $A \in \mathbb{E}^2$ esiste un unico $B \in \mathbb{E}^2$ tale che $\mathbf{u} = [AB]$. Consideriamo ora un altro vettore libero ordinario $\mathbf{v}$, e sia C l'unico punto di $\mathbb{E}^2$ tale che $\mathbf{v} = [BC]$. Poniamo*

$$\mathbf{u} + \mathbf{v} = [AC] \ .$$

Questa costruzione è detta talvolta regola del triangolo.



Abbiamo in tal modo definito una operazione di addizione in V^2; è agevole verificare che tale operazione è associativa e che $\forall\ P \in \mathbb{E}^2$ il vettore rappresentato dal segmento degenere $[PP]$ funge da elemento neutro rispetto a tale addizione. Tale vettore sarà detto vettore nullo *e sarà indicato con il simbolo* $\mathbf{0}$. *Infine ogni vettore libero ordinario $[PQ]$ è dotato di opposto $[QP]$. Osserviamo che tale addizione è commutativa e quindi $(V^2;+)$ è un gruppo abeliano. Vogliamo ora definire una moltiplicazione esterna $\cdot$ in V^2 con operatori in $\mathbb{R}$. Per ogni scalare $\lambda \in \mathbb{R}$ poniamo $\lambda \cdot \mathbf{0} = \mathbf{0}$. Per ogni vettore libero ordinario $[PQ]$ poniamo $0 \cdot [PQ] = \mathbf{0}$. Siano ora λ uno scalare non nullo e $[PQ]$ un vettore libero ordinario non nullo (ovvero sia $P \neq Q$). Sia r la retta passante per P e Q e denotiamo con ℓ, ℓ' le due semirette di r di origine P. Sia ad esempio ℓ quella contenente Q. Se $\lambda \in \mathbb{R}^+$, sia T l'unico punto di ℓ tale che il segmento PT abbia lunghezza λ, avendo usato la lunghezza di $[PQ]$ come unità di misura. Poniamo allora $\lambda \cdot [PQ] = [PT]$. Se invece $\lambda \in \mathbb{R}^-$, sia T' l'unico punto di ℓ' tale che il segmento PT' abbia lunghezza $-\lambda$. Poniamo allora $\lambda \cdot [PQ] = [PT']$. Si verifica, con semplici costruzioni geometriche elementari, che la struttura $(V^2;+,\cdot)$ è uno spazio vettoriale, che prende il nome di* spazio vettoriale dei vettori liberi ordinari del piano. *In modo analogo si costruisce, a partire dallo spazio euclideo tridimensionale $\mathbb{E}^3$, lo spazio vettoriale V^3 dei vettori liberi ordinari dello spazio.*

2.6 Esempio. **Vettori applicati.** *Consideriamo il piano $\mathbb{E}^2$ della geometria elementare. Se $A, B \in \mathbb{E}^2$ il segmento orientato AB è anche detto vettore applicato, A è il suo punto di applicazione o primo estremo, B il secondo estremo. Fissiamo un punto $O \in \mathbb{E}^2$. Consideriamo l'insieme $\mathbb{E}^2_O$ dei vettori applicati in O. Se $P, Q \in \mathbb{E}^2$ esiste un unico punto $R \in \mathbb{E}^2$ tale che il quadrilatero $OPRQ$ sia un parallelogramma (eventualmente degenere). Definiamo una operazione di addizione in $\mathbb{E}^2_O$, ponendo $OP + OQ = OR$. Questa costruzione è detta talvolta* regola del parallelogramma. *Si può procedere anche, equivalentemente, considerando il vettore libero α rappresentato da OP e il vettore libero β rappresentato da OQ ed osservare che il punto R di cui sopra è l'unico tale che $\alpha + \beta$ sia rappresentato da OR. Si definisce anche una moltiplicazione esterna con operatori in $\mathbb{R}$ in analogia con l'esempio precedente, e si ottiene una struttura di spazio vettoriale su $\mathbb{E}^2_O$, detta spazio dei vettori applicati nel piano (con origine, o punto di applicazione, O).*

2.7 Esempio. **Polinomi.** *Sia $V = \mathbb{K}[x]$, l'insieme dei polinomi sul campo $\mathbb{K}$ nell'indeterminata x. La struttura $(V;+)$ è un gruppo abeliano. Definiamo una operazione esterna su V con operatori in $\mathbb{K}$ al modo seguente. Siano $f \in V$, $\lambda \in \mathbb{K}$, ad esempio sia*

$$f = a_0 + a_1 x + \cdots + a_n x^n = \sum_{h=0}^{n} a_h x^h \ .$$

Poniamo allora $\lambda \cdot f = g$ dove

$$g = \lambda a_0 + \lambda a_1 x + \cdots + \lambda a_n x^n \ .$$

In altri termini il prodotto di uno scalare per un polinomio non è altro che il prodotto tra due polinomi, nel caso in cui il primo sia semplicemente uno scalare, ovvero il polinomio nullo o di grado 0. È agevole verificare che la struttura $(V;+,\cdot)$ così ottenuta è uno spazio vettoriale su $\mathbb{K}$. Nell'Esempio 1.54 avevamo utilizzato anche l'operazione di moltiplicazione interna, ottenendo così una struttura di anello. Vediamo quindi che uno stesso insieme, $\mathbb{K}[x]$ in questo caso, è sostegno di strutture diverse, a seconda delle operazioni considerate.

2.8 Esempio. **Campi e sottocampi.** *Sia $V = \mathbb{R}$ il campo dei numeri reali e $\mathbb{K} = \mathbb{Q}$ il campo dei numeri razionali. Consideriamo in V l'usuale addizione tra numeri reali. Inoltre, per ogni numero razionale k e per ogni numero reale λ, definiamo il loro prodotto (esterno) considerando di nuovo l'usuale prodotto tra numeri reali, considerando che un numero razionale è anche un numero reale. Con queste operazioni l'insieme dei numeri reali può vedersi come spazio vettoriale sul campo dei numeri razionali.*

2.9 Definizione. *Una n-pla, ovvero una sequenza ordinata, $\mathcal{S} = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$ di vettori sarà detta anche* sistema ordinato. *L'intero positivo n sarà detto ordine di $\mathcal{S}$.*

Quando l'ordine in cui compaiono i vettori non è essenziale, parleremo semplicemente di *sistema* e scriveremo $\mathcal{S} = [\mathbf{v}_1, \ldots, \mathbf{v}_n]$. Consideriamo dunque un sistema $\mathcal{S} = [\mathbf{v}_1, \ldots, \mathbf{v}_n]$. Utilizzeremo, con abuso di notazione, alcune notazioni di tipo insiemistico. Ad esempio, se $\mathbf{w}$ è uno degli elementi di $\mathcal{S}$ scriveremo $\mathbf{w} \in \mathcal{S}$. Diremo poi che t è la *molteplicità* di $\mathbf{w}$ in $\mathcal{S}$ se il vettore $\mathbf{w}$ compare t volte in $\mathcal{S}$. Sia ora $\mathcal{S}' = [\mathbf{w}_1, \ldots, \mathbf{w}_m]$ un altro sistema. Se accade che ogni vettore che appartiene a $\mathcal{S}$ con una certa molteplicità, appartiene anche ad $\mathcal{S}'$ con molteplicità non minore, diremo che $\mathcal{S}$ è incluso in $\mathcal{S}'$, o anche che $\mathcal{S}$ è un sottosistema di $\mathcal{S}'$, e scriveremo $\mathcal{S} \subseteq \mathcal{S}'$. Ad esempio, se

$$\mathcal{S} = [\mathbf{u}, \mathbf{v}, \mathbf{u}, \mathbf{w}] \ , \quad \mathcal{S}' = [\mathbf{u}, \mathbf{v}, \mathbf{w}] \ , \quad \mathcal{S}'' = [\mathbf{u}, \mathbf{v}, \mathbf{u}, \mathbf{w}, \mathbf{w}]$$

avremo che $\mathcal{S}' \subseteq \mathcal{S} \subseteq \mathcal{S}''$, ma $\mathcal{S} \not\subseteq \mathcal{S}'$. Osserviamo che il vettore $\mathbf{u}$ ha molteplicità 2 in $\mathcal{S}$, 1 in $\mathcal{S}'$, 2 in $\mathcal{S}''$.

2.10 Definizione. *Sia $\mathcal{S} = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$ un sistema ordinato di vettori e sia $(\alpha_1, \ldots, \alpha_n) \in \mathbb{K}^n$ una n-pla di scalari. Il vettore*

$$\mathbf{v} = \sum_{j=1}^{n} \alpha_j \mathbf{v}_j$$

si dice combinazione lineare di $\mathcal{S}$, ovvero dei vettori $\mathbf{v}_1, \ldots, \mathbf{v}_n$, con coefficienti $\alpha_1, \ldots, \alpha_n$.

Osserviamo esplicitamente che se $\alpha_1 = \cdots = \alpha_n = 0$, tale combinazione lineare risulterà nulla, qualunque siano i vettori $\mathbf{v}_1, \ldots, \mathbf{v}_n$ considerati. Ci si può chiedere invece se una combinazione lineare *non banale* di $\mathcal{S}$, ovvero con coefficienti non tutti nulli, può ancora essere nulla.

2.11 Definizione. *Il sistema ordinato $\mathcal{S}$ si dice (linearmente) dipendente se esiste una n-pla non banale di scalari $(\alpha_1, \ldots, \alpha_n) \in \mathbb{K}^n$ tale che risulti*

$$\sum_{j=1}^{n} \alpha_j \mathbf{v}_j = \mathbf{0} \ . \tag{2.1}$$

Ricordiamo che la n-pla di scalari $(\alpha_1, \ldots, \alpha_n)$ si dice non banale quando non tutti gli scalari $\alpha_1, \ldots, \alpha_n$ sono nulli. La (2.1) prende il nome di relazione di dipendenza del sistema ordinato $\mathcal{S}$. È chiaro che la definizione di dipendenza lineare non dipende dall'ordine con cui si considerano i vettori. Talvolta si dice anche che i vettori $\mathbf{v}_1, \ldots, \mathbf{v}_n$ sono dipendenti. In alcuni casi si riconosce immediatamente che un sistema è dipendente. Ad esempio, se il vettore nullo appartiene al sistema, oppure se nel sistema c'è una ripetizione, allora certamente il sistema è dipendente. Più in generale, vale la seguente proposizione di caratterizzazione.

2.12 DEFINIZIONE. *Sia* $\mathbf{v}$ *un vettore e* $\mathcal{S} = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$ *un sistema ordinato. Diremo che* $\mathbf{v}$ *dipende da* $\mathcal{S}$ *se esiste una* n*-pla di scalari* $(\alpha_1, \ldots, \alpha_n) \in \mathbb{K}^n$ *tale che risulti*

$$\mathbf{v} = \sum_{j=1}^{n} \alpha_j \mathbf{v}_j \ .$$

In particolare, osserviamo che il vettore nullo dipende da ogni sistema, e che ogni vettore appartenente ad un sistema dipende dal sistema stesso.

2.13 LEMMA. *Sia* $\mathbf{v}$ *un vettore e siano* $\mathcal{S} = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$, $\mathcal{S}' = (\mathbf{w}_1, \ldots, \mathbf{w}_m)$ *due sistemi ordinati. Se* $\mathbf{v}$ *dipende da* $\mathcal{S}$ *e ogni vettore di* $\mathcal{S}$ *dipende da* $\mathcal{S}'$*, allora* $\mathbf{v}$ *dipende da* $\mathcal{S}'$*.*

Dimostrazione. Omessa. □

Nella situazione descritta nel lemma precedente, si dice anche che il sistema $\mathcal{S}$ dipende dal sistema $\mathcal{S}'$. Anche in questo caso l'ordine non è rilevante. È chiaro che se $\mathcal{S}$ è un sottosistema di $\mathcal{S}'$, allora $\mathcal{S}$ dipende da $\mathcal{S}'$.

2.14 PROPOSIZIONE. *Il sistema* $\mathcal{S}$ *è dipendente se e solo se esiste un vettore in* $\mathcal{S}$ *che dipende dai rimanenti.*

Dimostrazione. Omessa. □

Osserviamo esplicitamente che se il sistema $\mathcal{S}$ è dipendente e si ha che $\mathcal{S} \subseteq \mathcal{S}'$, allora anche il sistema $\mathcal{S}'$ è dipendente.

2.15 DEFINIZIONE. *Il sistema ordinato* $\mathcal{S} = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$ *si dice indipendente se non è dipendente, ovvero se l'unica combinazione lineare nulla di* $\mathcal{S}$ *è quella banale, ovvero a coefficienti tutti nulli. In formule, deve essere vera la seguente implicazione:*

$$\sum_{j=1}^{n} \alpha_j \mathbf{v}_j = \mathbf{0} \quad \Longrightarrow \quad \alpha_j = 0 \ \forall j \ .$$

Anche in questo caso l'ordine non è rilevante. Talvolta diremo allora semplicemente che i vettori $\mathbf{v}_1, \ldots, \mathbf{v}_n$ sono indipendenti. Da quanto detto finora segue che se $\mathcal{S}$ è indipendente, il vettore nullo non appartiene al sistema, ed i vettori del sistema sono a due a due distinti. Nella teoria degli spazi vettoriali gioca un ruolo fondamentale lo studio dei sistemi di vettori e dei loro sottosistemi indipendenti massimali.

2.16 LEMMA. *Siano $\mathcal{S} = [\mathbf{v}_1, \ldots, \mathbf{v}_n]$, $\mathcal{S}' = [\mathbf{w}_1, \ldots, \mathbf{w}_m]$ due sistemi di vettori. Supponiamo che $\mathcal{S}$ sia indipendente e sia contenuto in $\mathcal{S}'$. $\mathcal{S}$ è indipendente massimale in $\mathcal{S}'$ se e solo se ogni vettore di $\mathcal{S}'$ dipende da $\mathcal{S}$.*

Dimostrazione. Omessa. □

2.2 Basi e dimensione

2.17 DEFINIZIONE. *Il sistema ordinato $\mathcal{S} = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$ si dice sistema di generatori di V se ogni vettore di V dipende da $\mathcal{S}$. In tal caso diremo anche che $\mathcal{S}$ genera V, ovvero che i vettori $\mathbf{v}_1, \ldots, \mathbf{v}_n$ generano V, o anche che i vettori $\mathbf{v}_1, \ldots, \mathbf{v}_n$ sono generatori di V, e scriveremo*

$$V = \mathcal{L}(\mathcal{S}) = \mathcal{L}(\mathbf{v}_1, \ldots, \mathbf{v}_n) \; . \tag{2.2}$$

Anche in questo caso l'ordine non è rilevante.

2.18 DEFINIZIONE. *Uno spazio vettoriale V si dice finitamente generato se ammette un sistema di generatori.*

Naturalmente una definizione relativa agli spazi vettoriali riveste un certo interesse se alcuni spazi vettoriali soddisfano tale definizione ed altri no.

2.19 ESEMPIO. *Consideriamo lo spazio vettoriale $V = \mathbb{K}[x]$ dei polinomi sul campo $\mathbb{K}$ dell'Esempio 2.7. Tale spazio non è finitamente generato. Sia infatti*

$$\mathcal{S} = [\, f_1, \ldots, f_m \,]$$

un sistema di elementi di V. Se t è il massimo tra 0 e massimo dei gradi dei polinomi non banali di $\mathcal{S}$, allora si verifica che i polinomi di grado maggiore di t non possono esprimersi come combinazione lineare dei polinomi $f_1, \ldots, f_m$. Quindi in ogni caso $\mathcal{S}$ non è un sistema di generatori di V.

È chiaro dalla definizione che, se $\mathcal{S}$ è un sistema di generatori di V e $\mathcal{S} \subseteq \mathcal{S}'$, anche $\mathcal{S}'$ è un sistema di generatori. Introduciamo ora una nozione di grande importanza.

2.20 Definizione. *Un sistema ordinato $\mathcal{B} = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$ si dice base ordinata, o anche riferimento, se è indipendente e genera V.*

Sottolineiamo ancora che le nozioni di dipendenza ed indipendenza lineare di un sistema, di sistema di generatori e di base sono tutte indipendenti dall'ordine con cui vengono considerati i vettori del sistema. D'ora in avanti non specificheremo se un sistema è ordinato. Inoltre indicheremo, con abuso di notazione, con lo stesso simbolo $\mathcal{S}$ un sistema ordinato di vettori $(\mathbf{v}_1, \ldots, \mathbf{v}_n)$ e anche lo stesso sistema (non ordinato) $[\mathbf{v}_1, \ldots, \mathbf{v}_n]$.

Osserviamo esplicitamente che lo spazio banale non possiede basi, in quanto in esso non esistono sistemi indipendenti. Anche uno spazio vettoriale non finitamente generato non ha basi, in quanto non possiede sistemi di generatori. È utile avere dei criteri per riconoscere se un sistema è una base.

2.21 Teorema di caratterizzazione delle basi. *Sia V uno spazio vettoriale non banale. Un sistema ordinato di vettori $\mathcal{B} = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$ di V è una base ordinata se e solo se vale una delle seguenti condizioni:*

(a) $\mathcal{B}$ è un sistema indipendente massimale di vettori di V;

(b) $\mathcal{B}$ è un sistema di generatori minimale di vettori di V;

(c) $\forall\, \mathbf{v} \in V,\ \exists!\ (\alpha_1, \ldots, \alpha_n) \in \mathbb{K}^n \mid \mathbf{v} = \sum_{j=1}^n \alpha_j \mathbf{e}_j$.

Dimostrazione. Sia $\mathcal{B} = [\mathbf{e}_1, \ldots, \mathbf{e}_n]$. Supponiamo che $\mathcal{B}$ sia una base e proviamo che valgono le condizioni (a), (b), (c).
Proviamo che vale la (a). Già sappiamo che $\mathcal{B}$ è indipendente. Proviamo che è anche massimale, verificando che ogni sistema che contiene $\mathcal{B}$ propriamente è dipendente. Sia $\mathcal{B}'$ un sistema contenente propriamente $\mathcal{B}$. Se in $\mathcal{B}'$ c'è un vettore che compare più di una volta, allora $\mathcal{B}'$ è dipendente. Se invece in $\mathcal{B}'$ non ci sono ripetizioni, deve esistere un vettore $\mathbf{v} \in \mathcal{B}'$ tale che $\mathbf{v} \notin \mathcal{B}$. Poiché $\mathcal{B}$ genera V, $\mathbf{v}$ dipende da $\mathcal{B}$ e quindi dipende anche dai rimanenti vettori di $\mathcal{B}'$. Pertanto, anche in questo caso, $\mathcal{B}'$ è dipendente.
Proviamo che vale la (b). Già sappiamo che $\mathcal{B}$ genera V. Proviamo che $\mathcal{B}$ è minimale come sistema di generatori. Per assurdo, sia $\mathcal{B}'$ un sistema di generatori contenuto propriamente in $\mathcal{B}$. Esiste allora un vettore $\mathbf{e}_j \in \mathcal{B}$ tale che $\mathbf{e}_j \notin \mathcal{B}'$. Poiché $\mathcal{B}'$ genera V, $\mathbf{e}_j$ dipende da $\mathcal{B}'$ e quindi anche dai rimanenti vettori di $\mathcal{B}$, che risulta quindi dipendente, in contraddizione con la sua ipotizzata indipendenza.
Proviamo che vale la (c). Già sappiamo che $\mathcal{B}$ è un sistema di generatori. Quindi per ogni $\mathbf{v} \in V$ esistono degli scalari $\alpha_1, \ldots, \alpha_n$ tali che

$$\mathbf{v} = \sum_{j=1}^n \alpha_j \mathbf{e}_j \ .$$

Vogliamo provare che tali scalari sono univocamente determinati. Se $\beta_1, \ldots, \beta_n$ sono degli scalari tali che

$$\mathbf{v} = \sum_{j=1}^{n} \beta_j \mathbf{e}_j$$

avremo che

$$\sum_{j=1}^{n} (\beta_j - \alpha_j)\mathbf{e}_j = \mathbf{0}$$

e poiché $\mathcal{B}$ è indipendente, si avrà che $\beta_j - \alpha_j = 0$, ovvero $\beta_j = \alpha_j$, per ogni j, e cioè gli scalari $\alpha_1, \ldots, \alpha_n$ sono univocamente determinati.
Proviamo ora che ognuna delle condizioni (a), (b), (c) implica che $\mathcal{B}$ è una base. Supponiamo che valga la (a). Allora già sappiamo che $\mathcal{B}$ è indipendente. Vogliamo provare che $\mathcal{B}$ genera V. Sia dunque $\mathbf{v}$ un qualunque vettore di V. Il sistema

$$\mathcal{B}' = [\mathbf{e}_1, \ldots, \mathbf{e}_n, \mathbf{v}]$$

contiene propriamente $\mathcal{B}$ e quindi, per la massimalità di $\mathcal{B}$ come sistema indipendente, $\mathcal{B}'$ deve risultare dipendente. Pertanto esistono degli scalari non tutti nulli $\beta_1, \ldots, \beta_n, \alpha$ tali che

$$\sum_{j=1}^{n} \beta_j \mathbf{e}_j + \alpha \mathbf{v} = \mathbf{0}\ . \tag{2.3}$$

Se $\alpha = 0$ la (3) si riduce a

$$\sum_{j=1}^{n} \beta_j \mathbf{e}_j = \mathbf{0} \tag{2.4}$$

con qualcuno degli scalari β_j non nullo. Ma allora la (2.4) è una relazione di dipendenza di $\mathcal{B}$ e questa è una contraddizione. Deve quindi essere $\alpha \neq 0$ e dalla (2.3) si deduce che

$$\mathbf{v} = -\alpha^{-1}\beta_1\mathbf{e}_1 + \cdots - \alpha^{-1}\beta_n\mathbf{e}_n$$

e cioè v dipende da $\mathcal{B}$ e $\mathcal{B}$ genera V.
Supponiamo che valga la (b). Sappiamo che $\mathcal{B}$ genera V. Vogliamo provare che $\mathcal{B}$ è indipendente. Per assurdo supponiamo che $\mathcal{B}$ sia dipendente. Esiste allora un vettore di $\mathcal{B}$ che dipende dai rimanenti. Sia esso, per fissare le idee, $\mathbf{e}_1$. Ciò vuol dire che esistono degli scalari $\beta_2, \ldots, \beta_n$ tali che

$$\mathbf{e}_1 = \sum_{j=2}^{n} \beta_j \mathbf{e}_j\ .$$

Poiché $\mathcal{B}$ genera V, per ogni $\mathbf{v} \in V$ esistono degli scalari $\alpha_1, \ldots, \alpha_n$ tali che

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

e quindi

$$\begin{aligned}\mathbf{v} &= \alpha_1(\sum_{j=2}^{n}\beta_j\mathbf{e}_j) + \alpha_2\mathbf{e}_2 + \cdots + \alpha_n\mathbf{e}_n \\ &= (\alpha_1\beta_2 + \alpha_2)\mathbf{e}_2 + \cdots + (\alpha_1\beta_n + \alpha_n)\mathbf{e}_n\end{aligned}$$

e $\mathbf{v}$ dipende da $[\mathbf{e}_2, \ldots, \mathbf{e}_n]$. Ma allora $[\mathbf{e}_2, \ldots, \mathbf{e}_n]$ genera V, e ciò è in contraddizione con l'ipotizzata minimalità di $\mathcal{B}$ come sistema di generatori.
Supponiamo infine che valga la (c). Da tale condizione si deduce banalmente che $\mathcal{B}$ genera V. Rimane da verificare l'indipendenza di $\mathcal{B}$. Poiché $\mathbf{0} \in V$, per la (c) $\mathbf{0}$ si esprime in unico modo come combinazione lineare di $\mathcal{B}$. D'altra parte è chiaro che

$$\mathbf{0} = 0\mathbf{e}_1 + \cdots + 0\mathbf{e}_n$$

e quindi l'unica combinazione lineare nulla di $\mathcal{B}$ è quella banale, ovvero a coefficienti tutti nulli, e $\mathcal{B}$ è indipendente. □

La condizione (c) si esprime anche dicendo che ogni vettore dello spazio si esprime in modo unico come combinazione lineare di $\mathcal{B}$.

2.22 Definizione. *Sia $\mathcal{B} = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$ una base di V e sia $\mathbf{v} \in V$. Gli scalari $\alpha_1, \ldots, \alpha_n$ di cui alla condizione* (c) *del teorema precedente si dicono componenti del vettore $\mathbf{v}$ in $\mathcal{B}$.*

Il seguente risultato, che potrebbe essere enunciato anche in modo più generale, è essenziale nello studio delle basi di uno spazio vettoriale.

2.23 Lemma di Steinitz. *Consideriamo una base $\mathcal{B} = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$ di V e sia $\mathcal{S}' = (\mathbf{w}_1, \ldots, \mathbf{w}_m)$ un sistema (qualunque) di vettori. Se $m > n$, il sistema $\mathcal{S}'$ è dipendente.*

Dimostrazione. Se qualche vettore di $\mathcal{S}$ è nullo, allora $\mathcal{S}$ è dipendente. Sia invece $\mathbf{v}_j \neq \mathbf{0}$ per ogni j. Poiché $\mathcal{B}$ è un sistema di generatori, $\mathbf{v}_1$ si può esprimere come combinazione lineare di $\mathcal{B}$; esistono cioè degli scalari $\alpha_1, \ldots, \alpha_n$ tali che

$$\mathbf{v}_1 = \alpha_1\mathbf{e}_1 + \cdots + \alpha_n\mathbf{e}_n \ . \tag{2.5}$$

Gli α_i non possono essere tutti nulli (altrimenti si avrebbe $\mathbf{v}_1 = \mathbf{0}$). Sia ad esempio $\alpha_1 \neq 0$. Tale assunzione non è restrittiva, in quanto in un sistema non conta l'ordine dei vettori. Dalla (2.5) si deduce allora che

$$\mathbf{e}_1 = \alpha_1^{-1}\mathbf{v}_1 - \alpha_1^{-1}\alpha_2\mathbf{e}_2 + \cdots - \alpha_1^{-1}\alpha_n\mathbf{e}_n \ .$$

Proviamo che il sistema $[\mathbf{v}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n]$ genera V. Sia $\mathbf{v} \in V$. Poiché $\mathcal{B}$ è un sistema di generatori, esistono degli scalari $\beta_1, \ldots, \beta_n$ tali che

$$\begin{aligned}\mathbf{v} &= \beta_1\mathbf{e}_1 + \beta_2\mathbf{e}_2 + \cdots + \beta_n\mathbf{e}_n \\ &= \beta_1\big(\alpha_1^{-1}\mathbf{v}_1 - \alpha_1^{-1}\alpha_2\mathbf{e}_2 + \cdots - \alpha_1^{-1}\alpha_n\mathbf{e}_n\big) + \beta_2\mathbf{e}_2 + \cdots + \beta_n\mathbf{e}_n \\ &= \beta_1\alpha_1^{-1}\mathbf{v}_1 + (\beta_2 - \beta_1\alpha_1^{-1}\alpha_2)\mathbf{e}_2 + \cdots + (\beta_n - \beta_1\alpha_1^{-1}\alpha_n)\mathbf{e}_n\end{aligned}$$

Pertanto ogni vettore di V si esprime come combinazione lineare dei vettori $\mathbf{v}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ e tali vettori generano V. In particolare $\mathbf{v}_2$ si potrà esprimere come combinazione lineare di tali vettori: esisteranno quindi degli scalari $\gamma_1, \ldots, \gamma_n$ tali che

$$\mathbf{v}_2 = \gamma_1\mathbf{v}_1 + \gamma_2\mathbf{e}_2 + \cdots + \gamma_n\mathbf{e}_n \ .$$

I γ_i non sono tutti nulli, altrimenti si avrebbe $\mathbf{v}_2 = \mathbf{0}$. Se accade che

$$\gamma_2 = \cdots = \gamma_n = 0 \ ,$$

si deduce che

$$\mathbf{v}_2 = \gamma_1\mathbf{v}_1$$

e quindi il sistema $\mathcal{S}$ è linearmente dipendente. Supponiamo invece che qualcuno degli scalari $\gamma_2, \ldots, \gamma_n$, ad esempio γ_2, sia diverso da 0. Avremo allora che

$$\mathbf{e}_2 = -\gamma_1\gamma_2^{-1}\mathbf{v}_1 + \gamma_2^{-1}\mathbf{v}_2 - \gamma_2^{-1}\gamma_3\mathbf{e}_3 + \cdots - \gamma_2^{-1}\gamma_n\mathbf{e}_n \ .$$

Come già fatto in precedenza, si prova che il sistema

$$[\mathbf{v}_1, \mathbf{v}_2, \mathbf{e}_3, \ldots, \mathbf{e}_n]$$

genera V. Si procede in modo analogo sostituendo ai vettori $\mathbf{e}_i$ i vettori $\mathbf{v}_i$ e si trova, dopo altri $n-2$ passaggi, che il sistema $[\mathbf{v}_1, \ldots, \mathbf{v}_n]$ genera V. In particolare $\mathbf{v}_m$ dipende da tale sistema, e quindi $\mathcal{S}$ è dipendente. □

Tale Lemma è del tutto equivalente al seguente enunciato.

2.24 Lemma. *Sia $\mathcal{B} = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$ una base di V e sia $\mathcal{S}' = (\mathbf{w}_1, \ldots, \mathbf{w}_m)$ un sistema (qualunque) di vettori. Se il sistema $\mathcal{S}'$ è indipendente allora $m \leq n$.*

Una importante conseguenza del Lemma di Steinitz è il seguente

2.25 Teorema di equipotenza delle basi. *Sia V uno spazio vettoriale non banale e finitamente generato. Le basi di V sono tutte equipotenti.*

Dimostrazione. Siano $\mathcal{B}', \mathcal{B}''$ due basi, di cardinalità s, t rispettivamente, di uno spazio vettoriale V finitamente generato. Applicando il Lemma 2.24 con $\mathcal{B}'$ nel ruolo di $\mathcal{B}$ e $\mathcal{B}''$ nel ruolo di $\mathcal{S}$ si trova che $t \leq s$. Scambiando i ruoli di $\mathcal{B}', \mathcal{B}''$ si verifica invece che $t \geq s$. Quindi $s = t$. □

Abbiamo osservato che affinché esistano basi è necessario che lo spazio sia non banale e finitamente generato. La seguente proposizione ci assicura che tale condizione è anche sufficiente.

2.26 Teorema di estrazione di una base. *Sia V uno spazio vettoriale non banale e sia $\mathcal{S} = [\mathbf{v}_1, \dots, \mathbf{v}_m]$ un sistema di generatori di V. Da $\mathcal{S}$ si può estrarre una base di V.*

Dimostrazione. Sia $\mathcal{S} = [\mathbf{v}_1, \dots, \mathbf{v}_m]$. Se $\mathcal{S}$ è un sistema minimale di generatori, allora $\mathcal{S}$ è una base di V. Se $\mathcal{S}$ non è minimale, esiste un indice i tale che il sistema $[\mathbf{v}_1, \dots, \mathbf{v}_{i-1}, \mathbf{v}_{i+1}, \dots, \mathbf{v}_m]$ genera ancora V. Se tale sistema di generatori è minimale, abbiamo trovato una base estratta da $\mathcal{S}$. Altrimenti possiamo rimuovere un altro vettore e procedere in modo analogo fino a che non si perviene ad un sistema minimale di generatori (che sarà necessariamente di ordine almeno 1, in quanto lo spazio non è banale). $\square$

Pertanto ogni spazio vettoriale non banale ammette basi, ed esse sono tutte equipotenti. Ha pertanto senso la seguente definizione.

2.27 Definizione. *Sia V uno spazio vettoriale finitamente generato. La dimensione di V, che si indica con il simbolo* $\dim V$, *è* 0 *se V è lo spazio banale; è invece la cardinalità di una sua base se V è non banale.*

2.28 Esempio. *Consideriamo il 2-spazio numerico $V = \mathbb{K}^2$ sul campo $\mathbb{K}$ e il sistema*

$$\mathcal{B}_2 = \left[\,(1,0),(0,1)\,\right] .$$

È agevole verificare che $\mathcal{B}_2$ è indipendente e genera $\mathbb{K}^2$. Pertanto $\mathcal{B}_2$ è una base di $\mathbb{K}^2$, ogni altra base ha ordine 2 e $\dim \mathbb{K}^2 = 2$. In modo analogo, per ogni $n \in \mathbb{N}$, consideriamo lo spazio numerico $V = \mathbb{K}^n$ su $\mathbb{K}$ e il sistema

$$\mathcal{B}_n = \left[\,(1,0,\dots,0),(0,1,\dots,0),\ \dots\ (0,0,\dots,1)\,\right] .$$

È agevole verificare che $\mathcal{B}_n$ è una base di $\mathbb{K}^n$, ogni altra base ha ordine n e $\dim \mathbb{K}^n = n$. Il sistema $\mathcal{B}_n$ prende il nome di base canonica, *oppure* standard, *oppure* naturale *o ancora* usuale *di $\mathbb{K}^n$.*

2.29 Esempio. *Consideriamo lo spazio vettoriale V^2 dei vettori liberi ordinari del piano. Un qualunque sistema formato da due vettori non nulli e non paralleli è una base di V^2, ogni altra base ha ordine 2 e $\dim V^2 = 2$. Consideriamo ora lo spazio vettoriale V^3 dei vettori liberi ordinari dello spazio. Un qualunque sistema formato da tre vettori non nulli e non complanari è una base di V^3, ogni altra base ha ordine 3 e $\dim V^3 = 3$. Ricordiamo che tre vettori liberi dello spazio $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ sono* non complanari *se, fissato un punto O e determinati i punti P_1, P_2, P_3 tali che $\mathbf{u}_1 = [OP_1]$, $\mathbf{u}_2 = [OP_2]$, $\mathbf{u}_3 = [OP_3]$, si ha che i punti O, P_1, P_2, P_3 sono non complanari (ovvero non esiste un piano che li contiene tutti).*

Abbiamo già osservato come da un sistema di generatori di uno spazio non banale si possa estrarre una base eliminando eventualmente alcuni vettori. Vediamo ora come si può ottenere una base aggiungendo opportunamente vettori ad un sistema indipendente.

2.30 Teorema di completamento di una base. *Sia V uno spazio vettoriale finitamente generato e sia $\mathcal{S} = [\mathbf{v}_1, \dots, \mathbf{v}_r]$ un sistema indipendente di vettori di V. Esistono allora dei vettori $\mathbf{v}_{r+1}, \dots, \mathbf{v}_n$ tali che il sistema $\mathcal{S}' = [\mathbf{v}_1, \dots, \mathbf{v}_r, \mathbf{v}_{r+1}, \dots, \mathbf{v}_n]$ sia una base di V.*

Dimostrazione. Sia $\dim V = n$. Se $\mathcal{S}$ è un sistema di generatori di V, allora $\mathcal{S}$ è una base di V (e $n = r$). In caso contrario, si può scegliere un vettore $\mathbf{v}_{r+1} \in V$ che non dipende da $\mathcal{S}$. Proviamo che il sistema $\mathcal{S}' = [\mathbf{v}_1, \dots, \mathbf{v}_r, \mathbf{v}_{r+1}]$ è indipendente. Siano $\alpha_1, \dots, \alpha_r, \beta \in \mathbb{K}$ tali che

$$\alpha_1\mathbf{v}_1 + \cdots + \alpha_r\mathbf{v}_r + \beta\mathbf{v}_{r+1} = \mathbf{0} \ . \tag{2.6}$$

Se fosse $\beta \neq 0$, avremmo che

$$\mathbf{v}_{r+1} = (-\beta^{-1}\alpha_1)\mathbf{v}_1 + \cdots + (-\beta^{-1}\alpha_r)\mathbf{v}_r$$

e quindi $\mathbf{v}_{r+1}$ dipenderebbe da $\mathcal{S}$ ed $\mathcal{S}'$ risulterebbe dipendente. Dunque $\beta = 0$ e la (2.6) si riduce a

$$\alpha_1\mathbf{v}_1 + \cdots + \alpha_r\mathbf{v}_r = \mathbf{0} \ .$$

Pertanto, poiché $\mathcal{S}$ è indipendente, $\alpha_1 = \cdots = \alpha_r = 0$. Quindi $\mathcal{S}'$ è indipendente. Se $\mathcal{S}'$ è un sistema di generatori, allora $\mathcal{S}'$ è una base di V. Altrimenti si procede in modo analogo per trovare altri vettori $\mathbf{v}_{r+2}, \dots, \mathbf{v}_n$, ottenendo così una base di V, in quanto sarà sicuramente impossibile determinare un sistema indipendente di ordine maggiore di n. □

Il teorema ora enunciato è utile sotto il profilo teorico, ma non è costruttivo, ovvero non spiega come trovare i vettori $\mathbf{v}_{r+1}, \dots, \mathbf{v}_n$ utilizzati per completare a base il sistema $\mathcal{S}$. Talvolta è utile allora la seguente riformulazione (più debole) del teorema precedente.

2.31 Teorema di completamento bis. *Sia V uno spazio vettoriale di dimensione n, sia $\mathcal{B} = [\mathbf{e}_1, \dots, \mathbf{e}_n]$ una base di V, e sia $\mathcal{S} = [\mathbf{v}_1, \dots, \mathbf{v}_r]$ un sistema indipendente di vettori di V. Allora i vettori $\mathbf{v}_{r+1}, \dots, \mathbf{v}_n$ tali che il sistema $\mathcal{S}' = [\mathbf{v}_1, \dots, \mathbf{v}_r, \mathbf{v}_{r+1}, \dots, \mathbf{v}_n]$ sia una base di V, esistenti in base al teorema precedente, possono essere scelti tra i vettori di $\mathcal{B}$.*

Dimostrazione. È chiaro che $r \leq n$. Se $r = n$ in base a risultati già acquisiti $\mathcal{S}$ è una base. Pertanto se $r < n$ vuol dire che $\mathcal{S}$ non è una base, e poiché è indipendente, non è un sistema di generatori di V. Esiste allora un vettore $\mathbf{e}_{i_1}$ in $\mathcal{B}$ che non dipende da $\mathcal{S}$. Infatti, se ogni vettore di $\mathcal{B}$ dipendesse da $\mathcal{S}$, poiché

$\mathcal{B}$ è una base, ogni vettore di V dipenderebbe da $\mathcal{S}$. È agevole provare che il sistema
$$\mathcal{S}' = [\mathbf{v}_1, \ldots, \mathbf{v}_r, \mathbf{e}_{i_1}]$$
è indipendente. Se $r+1=n$, allora $\mathcal{S}'$ è una base. Altrimenti si procede in modo analogo e si scelgono dei vettori $\mathbf{e}_{i_2}, \ldots, \mathbf{e}_{i_{n-r}}$ in $\mathcal{B}$ tali che il sistema
$$\mathcal{S}'' = [\mathbf{v}_1, \ldots, \mathbf{v}_r, \mathbf{e}_{i_1}, \ldots, \mathbf{e}_{i_{n-r}}]$$
sia indipendente. Poiché $\mathcal{S}''$ consta di n vettori, esso sarà una base di V, del tipo richiesto nell'enunciato. □

2.32 Esempio. *Consideriamo il 2-spazio numerico $V = \mathbb{K}^2$ sul campo $\mathbb{K}$ e il sistema*
$$\mathcal{B}' = \left[\,(1,1), (1,-1)\,\right] .$$
È agevole verificare che $\mathcal{B}'$ è indipendente. Poiché $\mathcal{B}_2$ è una base, in base al Lemma di Steinitz, non possono esistere sistemi indipendenti di ordine maggiore di 2. Quindi $\mathcal{B}'$ è indipendente massimale, cioè è una base. Osserviamo che, per ogni coppia $(a,b) \in \mathbb{K}^2$, è verificata l'uguaglianza
$$(a,b) = \frac{a+b}{2}(1,1) + \frac{a-b}{2}(1,-1)) .$$
Pertanto, le componenti del vettore (a,b) in $\mathcal{B}_2$ sono proprio gli scalari a, b, ma le componenti dello stesso vettore (a,b) *nella base $\mathcal{B}'$ sono gli scalari $\frac{a+b}{2}, \frac{a-b}{2}$.*

2.33 Esempio. *Consideriamo il 2-spazio numerico $V = \mathbb{R}^2$ sul campo $\mathbb{R}$ e il sistema*
$$\mathcal{S} = \left[\,(1,1), (1,2), (2,1)\,\right] .$$
È agevole verificare che $\mathcal{S}$ genera $\mathbb{R}^2$. Infatti se consideriamo un generico elemento $(a,b) \in \mathbb{R}^2$, osserviamo che, per ogni $h \in \mathbb{R}$, è verificata l'uguaglianza
$$(a,b) = (2a-b-3h)(1,1) + (-a+b+h)(1,2) + h(2,1)$$
e quindi la coppia (a,b) si esprime (in infiniti modi differenti) come combinazione lineare dei vettori di $\mathcal{S}$. Ad esempio, se consideriamo la coppia $(3,5)$, scelto $h=2$, abbiamo
$$(3,5) = (-5)(1,1) + 4(1,2) + 2(2,1) .$$
Poiché $\mathcal{B}_2$ è una base, in base al Lemma di Steinitz, non possono esistere sistemi indipendenti di ordine maggiore di 2. Quindi $\mathcal{S}$ è dipendente e non è una base.

2.34 Esempio. *Consideriamo il 3-spazio numerico $V = \mathbb{R}^3$ sul campo $\mathbb{R}$ e il sistema*
$$\mathcal{S} = \left[\,(1,0,1), (1,1,1)\,\right] .$$
È agevole verificare che $\mathcal{S}$ è un sistema indipendente. Consideriamo, infatti, una generica combinazione lineare di $\mathcal{S}$, del tipo $x(1,0,1) + y(1,1,1)$ e imponiamo che essa sia uguale al vettore nullo $(0,0,0)$:
$$x(1,0,1) + y(1,1,1) = (0,0,0) .$$
Effettuando i calcoli troviamo
$$(x+y, y, x+y) = (0,0,0) ,$$
e quindi, uguagliando componente per componente, deduciamo che deve essere $x=0$ e $y=0$. Poiché $\mathcal{B}_3$ è una base, $\mathcal{S}$ non può essere una base (ha solo 2 elementi), e quindi non può essere un sistema di generatori. Possiamo ottenere una base aggiungendo ad $\mathcal{S}$ un opportuno vettore che, in base al Teorema di completamento bis, può essere scelto nella base standard $\mathcal{B}_3$. Osserviamo che, in questo caso, possiamo scegliere sia il vettore $(1,0,0)$ che il vettore $(0,0,1)$, ma non $(0,1,0)$.

2.3 Sottospazi

Introdurremo ora la nozione di *sottospazio* di uno spazio vettoriale. Sia V uno spazio vettoriale su $\mathbb{K}$ e sia $W \subseteq V$.

2.35 DEFINIZIONE. *Il sottoinsieme W si dice sottospazio (vettoriale) di V se sono soddisfatte le seguenti condizioni:*

(i) $\mathbf{0} \in W$;

(ii) $\forall\ \mathbf{u}, \mathbf{v} \in W$ *si ha che* $\mathbf{u} + \mathbf{v} \in W$;

(iii) $\forall\ \alpha \in \mathbb{K},\ \forall\ \mathbf{v} \in W$ *si ha che* $\alpha\mathbf{v} \in W$.

La *(i)* ci assicura, in particolare, che $W \neq \emptyset$. La *(ii)* e la *(iii)* rappresentano la stabilità del sottoinsieme W rispetto alle operazioni di addizione (interna) e moltiplicazione (esterna), ovvero ci assicurano che W è, come si suol dire, una *parte stabile* di V, ed equivalgono alla condizione seguente:

$$\forall\ \alpha, \beta \in \mathbb{K},\ \forall\ \mathbf{u}, \mathbf{v} \in W \quad \text{si ha che} \quad \alpha\mathbf{u} + \beta\mathbf{v} \in W\ .$$

Tale condizione coinvolge simultaneamente le operazioni di addizione e moltiplicazione ed è nota come condizione di stabilità rispetto alla formazione di combinazioni lineari con due vettori e due scalari. In realtà il fatto che si utilizzino solo due vettori e due scalari è del tutto irrilevante, e la condizione può scriversi, in forma più generale, al modo seguente:

$$\forall\ k \in \mathbb{N},\ \forall\ \alpha_1, \ldots, \alpha_k \in \mathbb{K},\ \forall\ \mathbf{u}_1, \ldots, \mathbf{u}_k \in W \quad \text{si ha che} \quad \sum_{\ell=1}^{k} \alpha_\ell \mathbf{u}_\ell \in W\ .$$

2.36 PROPOSIZIONE. *Sia V uno spazio vettoriale e sia W un suo sottospazio. Con le operazioni che W eredita da V, W è a sua volta uno spazio vettoriale.*

Dimostrazione. È lasciata al lettore per esercizio. □

Per indicare che W è un sottospazio di V scriveremo $W \leq V$. Useremo la notazione $W < V$ nel caso di inclusione stretta. Osserviamo che se $W \leq V$ e V è finitamente generato, anche W è finitamente generato e $\dim W \leq \dim V$. Il segno di uguaglianza tra le dimensioni vale inoltre se e solo se $W = V$. Il lettore potrà verificare facilmente, per esercizio, che l'intersezione di due o più sottospazi di V è ancora un sottospazio di V. Un modo molto utile per costruire e descrivere i sottospazi è il seguente. Sia $\mathcal{S}$ un sistema di vettori, oppure un semplice sottoinsieme di V. Indichiamo con il simbolo $\mathcal{L}(\mathcal{S})$ l'insieme di tutte le possibili combinazioni lineari di vettori di $\mathcal{S}$. Per convenzione, se $\mathcal{S} = \emptyset$ poniamo $\mathcal{L}(\mathcal{S}) = \{\mathbf{0}\}$, lo spazio banale, o nullo.

2.37 Proposizione. *Per ogni sistema, o anche sottoinsieme, $\mathcal{S}$ dello spazio vettoriale V, $\mathcal{L}(\mathcal{S})$ è un sottospazio di V.*

Dimostrazione. È lasciata al lettore per esercizio. □

Il sottospazio $\mathcal{L}(\mathcal{S})$ viene comunemente detto *sottospazio generato* da $\mathcal{S}$. In altri termini, $\mathcal{L}$ può vedersi come applicazione dell'insieme delle parti di V nell'insieme dei sottospazi di V, ed in quanto tale gode di interessanti proprietà, che saranno evidenziate negli esercizi, dette subadditività ed idempotenza. Osserviamo che se $\mathcal{S}$ è un sistema di generatori di V si ha che $\mathcal{L}(\mathcal{S}) = V$, e ciò è consistente con le notazioni già utilizzate nella (2.2). È poi chiaro che $\mathcal{L}(V) = V$, e infine $\mathcal{L}(W) = W$ per ogni sottospazio W di V. Si potrebbe dimostrare facilmente che per ogni sistema, o anche sottoinsieme, $\mathcal{S}$ dello spazio vettoriale V, il sottospazio $\mathcal{L}(\mathcal{S})$ è il più piccolo sottospazio di V contenente $\mathcal{S}$, o anche che $\mathcal{L}(\mathcal{S})$ è l'intersezione di tutti i sottospazi di V che contengono $\mathcal{S}$.

2.38 Esempio. *Consideriamo lo spazio vettoriale numerico $V = \mathbb{R}^2$ di dimensione 2 su $\mathbb{R}$ ed i suoi sottoinsiemi*

$$A = \left\{ (a,b) \in \mathbb{R}^2 \;\middle|\; a,b \geq 0 \right\} ;$$
$$B = \left\{ (a,b) \in \mathbb{R}^2 \;\middle|\; a \cdot b \geq 0 \right\} ;$$
$$C = \left\{ (a,b) \in \mathbb{R}^2 \;\middle|\; a - b = 0 \right\} .$$

Il lettore potrà verificare che C è un sottospazio, ma A e B non lo sono.

2.39 Esempio. *Consideriamo lo spazio vettoriale numerico $V = \mathbb{R}^3$ di dimensione 3 su $\mathbb{R}$ ed i suoi sottoinsiemi*

$$A = \left\{ (a,b,c) \in \mathbb{R}^3 \;\middle|\; a + b + c = 0 \right\} ;$$
$$B = \left\{ (a,b,c) \in \mathbb{R}^3 \;\middle|\; a \geq b \geq c \right\} ;$$
$$C = \left\{ (a,b,c) \in \mathbb{R}^3 \;\middle|\; a - b = a + c = 0 \right\} .$$

Il lettore potrà verificare che A e C sono sottospazi, ma B non lo è.

2.40 Esempio. *Sia V uno spazio vettoriale non banale sul campo $\mathbb{K}$ e sia $\mathbf{u} \in V$ un vettore non nullo. L'insieme $A = \{ k\mathbf{u} \,|\, k \in \mathbb{K} \}$, costituito dai* multipli scalari *di $\mathbf{u}$, è un sottospazio di V, detto* retta vettoriale. *Si osservi che, in base alle notazioni introdotte, si ha che $A = \mathcal{L}(\mathbf{u})$.*

2.41 Esempio. *Consideriamo lo spazio vettoriale dei polinomi $V = \mathbb{R}[x]$ su $\mathbb{R}$ ed i suoi sottoinsiemi*

$$A = \left\{ p \in \mathbb{R}[x] \;\middle|\; \deg p \leq 3 \right\} ;$$
$$B = \left\{ p \in \mathbb{R}[x] \;\middle|\; \deg p = 3 \right\} ;$$
$$C = \left\{ p \in \mathbb{R}[x] \;\middle|\; \deg p \text{ è pari} \right\} .$$

Il lettore potrà verificare che A è un sottospazio, ma B e C non lo sono.

2.42 Proposizione. *Consideriamo un sistema $\mathcal{S}$ di ordine m e due suoi sottosistemi $\mathcal{S}'$ e $\mathcal{S}''$ di ordine h, k rispettivamente. Se $\mathcal{S}'$ è un sottosistema indipendente massimale in $\mathcal{S}$, allora $\mathcal{L}(\mathcal{S}) = \mathcal{L}(\mathcal{S}')$. Se poi anche $\mathcal{S}''$ è un sottosiste-*

ma indipendente massimale in $\mathcal{S}$, i sistemi $\mathcal{S}'$ e $\mathcal{S}''$ hanno lo stesso numero di elementi, cioè $h = k$.

Dimostrazione. In base a risultati già acquisiti, poiché $\mathcal{S}', \mathcal{S}''$ sono indipendenti massimali in $\mathcal{S}$, ogni vettore di $\mathcal{S}$ dipende da $\mathcal{S}'$ e da $\mathcal{S}''$. Consideriamo allora il sottospazio $W = L(\mathcal{S})$. Si verifica agevolmente che $\mathcal{S}, \mathcal{S}'$ sono entrambe basi di W, e quindi $h = k$. □

Nella situazione sopra descritta, si dice talvolta che $\mathcal{S}$ ha rango h. Vogliamo ora occuparci del problema della costruzione di nuovi sottospazi di uno spazio vettoriale. A tal proposito, consideriamo due sottospazi U, W di V. Abbiamo già osservato che la loro intersezione $U \cap W$ è ancora un sottospazio di V. La loro unione $U \cup W$ non è, in generale, un sottospazio. Un caso particolare è quello in cui $U \subseteq W$. È chiaro allora che $U \cup W = W$, che è un sottospazio. Altrimenti, possiamo considerare il sottospazio generato da $U \cup W$, ovvero $\mathcal{L}(U \cup W)$.

2.43 Definizione. *Sia V uno spazio vettoriale e siano U, W due sottospazi di V. Il sottospazio $\mathcal{L}(U \cup W)$ sarà detto sottospazio somma, oppure congiungente, di U e W, e sarà indicato con il simbolo $U + W$.*

Alcune proprietà dell'operazione ora introdotta di somma di sottospazi sono riassunte nella seguente proposizione.

2.44 Proposizione. *Sia V uno spazio vettoriale e siano U, W due sottospazi di V. Valgono le seguenti proprietà.*

(i) Se $U \subseteq W$, si ha che $U + W = U \cup W = \mathcal{L}(U \cup W) = W$;

(ii) In generale

$$U + W = \left\{ \mathbf{v} \in V \;\middle|\; \exists\, \mathbf{u} \in U,\ \exists\, \mathbf{w} \in W \;\middle|\; \mathbf{v} = \mathbf{u} + \mathbf{w} \right\} ;$$

(iii) Se i vettori $\mathbf{u}_1, \ldots, \mathbf{u}_h$ generano U e i vettori $\mathbf{w}_1, \ldots, \mathbf{w}_k$ generano W, si ha che

$$U + W = \mathcal{L}(\mathbf{u}_1, \ldots, \mathbf{u}_h, \mathbf{w}_1, \ldots, \mathbf{w}_k) .$$

Dimostrazione. È lasciata al lettore per esercizio. □

Il seguente risultato riveste un notevole interesse e mette in relazione le dimensioni di vari sottospazi vettoriali.

2.45 Teorema di Grassmann. *Sia V uno spazio vettoriale e siano U, W due suoi sottospazi finitamente generati. Allora anche il sottospazio $U + W$*

è finitamente generato e si ha che

$$\dim(U+W) = \dim U + \dim W - \dim(U \cap W) \ .$$

Dimostrazione. Se almeno uno dei due sottospazi è banale, il risultato è immediato. Supponiamo quindi che entrambi i sottospazi siano non banali. Sia $\dim U = s$, $\dim W = t$, $\dim U \cap W = h$. Ovviamente si ha che $h \leq s, t$. Sia $[\mathbf{e}_1, \ldots, \mathbf{e}_h]$ una base di $U \cap W$. I vettori di tale base sono indipendenti ed è quindi possibile trovare $s-h$ vettori $\mathbf{u}_{h+1}, \ldots, \mathbf{u}_s \in U$ e $t-h$ vettori $\mathbf{w}_{h+1}, \ldots, \mathbf{w}_t \in W$ tali che i sistemi

$$\mathcal{B}' = [\mathbf{e}_1, \ldots, \mathbf{e}_h, \mathbf{u}_{h+1}, \ldots, \mathbf{u}_s] \, , \ \mathcal{B}'' = [\mathbf{e}_1, \ldots, \mathbf{e}_h, \mathbf{w}_{h+1}, \ldots, \mathbf{w}_t]$$

risultino delle basi di U e W. Proviamo che il sistema

$$\mathcal{B} = [\mathbf{e}_1, \ldots, \mathbf{e}_h, \mathbf{u}_{h+1}, \ldots, \mathbf{u}_s, \mathbf{w}_{h+1}, \ldots, \mathbf{w}_t]$$

è una base di $U+W$. Verifichiamo che $\mathcal{B}$ genera $U+W$. Sia $\mathbf{v} \in U+W$. Esistono dei vettori $\mathbf{u} \in U$ e $\mathbf{w} \in W$ tali che $\mathbf{v} = \mathbf{u} + \mathbf{w}$. Poiché $\mathcal{B}', \mathcal{B}''$ sono basi di U e W, possiamo determinare (in modo peraltro univoco), degli scalari $\alpha_1, \ldots, \alpha_s, \beta_1, \ldots, \beta_t$ tali che

$$\mathbf{u} = \alpha_1 \mathbf{e}_1 + \cdots + \alpha_h \mathbf{e}_h + \alpha_{h+1} \mathbf{u}_{h+1} + \cdots + \alpha_s \mathbf{u}_s$$

$$\mathbf{w} = \beta_1 \mathbf{e}_1 + \cdots + \beta_h \mathbf{e}_h + \beta_{h+1} \mathbf{w}_{h+1} + \cdots + \beta_t \mathbf{w}_t \ .$$

Pertanto

$$\mathbf{v} = \mathbf{u} + \mathbf{w} = \sum_{i=1}^{h} (\alpha_i + \beta_i) \mathbf{e}_i + \sum_{j=h+1}^{s} \alpha_j \mathbf{u}_j + \sum_{k=h+1}^{t} \beta_k \mathbf{w}_k$$

e cioè $\mathbf{v}$ dipende da $\mathcal{B}$. Resta ora da provare che il sistema di generatori $\mathcal{B}$ è indipendente. Consideriamo degli scalari

$$\gamma_1, \ldots, \gamma_h, \delta_{h+1}, \ldots, \delta_s, \epsilon_{h+1}, \ldots, \epsilon_t$$

tali che

$$\sum_{i=1}^{h} \gamma_i \mathbf{e}_i + \sum_{j=h+1}^{s} \delta_j \mathbf{u}_j + \sum_{k=h+1}^{t} \epsilon_k \mathbf{w}_k = \mathbf{0} \ . \tag{2.7}$$

Si ha che

$$\underbrace{-\sum_{j=h+1}^{s} \delta_j \mathbf{u}_j}_{\in U} = \underbrace{\sum_{i=1}^{h} \gamma_i \mathbf{e}_i + \sum_{k=h+1}^{t} \epsilon_k \mathbf{w}_k}_{\in W} \in U \cap W \ .$$

Esisteranno dunque degli scalari $\lambda_1, \ldots, \lambda_h$ tali che

$$-\sum_{j=h+1}^{s} \delta_j \mathbf{u}_j = \sum_{i=1}^{h} \lambda_i \mathbf{e}_i$$

in quanto $[\mathbf{e}_1, \ldots, \mathbf{e}_h]$ è una base di $U \cap W$. Si ha quindi che

$$\sum_{j=h+1}^{s} \delta_j \mathbf{u}_j + \sum_{i=1}^{h} \lambda_i \mathbf{e}_i = \mathbf{0} .$$

Poiché $\mathcal{B}'$ è un sistema indipendente, deduciamo che $\delta_{h+1}, \ldots, \delta_s = 0$. Analogamente si prova che $\epsilon_{h+1}, \ldots, \epsilon_t = 0$ e dalla (2.7) si deduce che

$$\sum_{i=1}^{h} \gamma_i \mathbf{e}_i = \mathbf{0}$$

e quindi $\gamma_1, \ldots, \gamma_h = 0$, poiché il sistema $[\mathbf{e}_1, \ldots, \mathbf{e}_h]$ è indipendente. Abbiamo quindi provato che $\mathcal{B}$ è una base di $U + W$. D'altra parte la sua cardinalità è $h + (s - h) + (t - h) = s + t - h$ e ciò conclude la dimostrazione. □

La relazione che compare nell'enunciato è nota come *formula di Grassmann.* Vogliamo ora descrivere un caso particolare nella costruzione della somma di due sottospazi.

2.46 Definizione. *Sia V uno spazio vettoriale e siano U, W due sottospazi di V. Il sottospazio congiungente $U + W$ di U e W si dice somma diretta, e sarà indicato con il simbolo $U \oplus W$, se per ogni vettore $\mathbf{v} \in U + W$ i vettori $\mathbf{u}$ e $\mathbf{w}$ tali che $\mathbf{v} = \mathbf{u} + \mathbf{w}$ di cui alla definizione precedente sono univocamente determinati, ovvero, in formule,*

$$\forall \mathbf{v} \in U + W \ \exists !\, \mathbf{u} \in U,\ \exists !\, \mathbf{w} \in W \ \Big|\ \mathbf{v} = \mathbf{u} + \mathbf{w} .$$

.

La nozione di somma diretta di due sottospazi si caratterizza come segue.

2.47 Proposizione. *Sia V uno spazio vettoriale e siano U, W due sottospazi di V. Il sottospazio congiungente $U + W$ di U e W è una somma diretta se e solo se $U \cap W = \{\mathbf{0}\}$.*

Dimostrazione. Supponiamo che il sottospazio congiungente $U+W$ di U e W sia una somma diretta. Sia $\mathbf{v} \in U \cap W$. Se fosse $\mathbf{v} \neq \mathbf{0}$ avremmo che $\mathbf{v} \in U$, $-\mathbf{v} \in W$ e

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

e quindi il vettore nullo si decomporrebbe in due modi distinti in somma di vettori di U e di W. Supponiamo, viceversa, che $U \cap W = \{\mathbf{0}\}$, poniamo $V' = U + W$ e consideriamo un vettore $\mathbf{v} \in V'$. Esisteranno allora dei vettori $\mathbf{u} \in U$, $\mathbf{w} \in W$ tali che $\mathbf{v} = \mathbf{u} + \mathbf{w}$. Bisogna provare l'unicità di tale decomposizione.

Siano dunque $\mathbf{u}' \in U$, $\mathbf{w}' \in W$ tali che $\mathbf{v} = \mathbf{u} + \mathbf{w} = \mathbf{u}' + \mathbf{w}'$. Si ha che $\mathbf{u}' - \mathbf{u} = \mathbf{w} - \mathbf{w}' \in U \cap W = \{\mathbf{0}\}$ e quindi $\mathbf{u}' - \mathbf{u} = \mathbf{0} = \mathbf{w} - \mathbf{w}'$, ovvero $\mathbf{u} = \mathbf{u}'$, $\mathbf{w} = \mathbf{w}'$. □

Osserviamo esplicitamente che se i sottospazi U, W sono finitamente generati, alla luce di quanto appena enunciato e della formula di Grassmann, si deduce che il sottospazio congiungente $U+W$ è una somma diretta se e solo se $\dim(U+W) = \dim U + \dim W$.

Generalizziamo ora la costruzione di sottospazio congiungente e di somma diretta nel caso in cui si considerino non necessariamente solo due, ma un qualunque numero (finito) di sottospazi.

2.48 Definizione. *Sia V uno spazio vettoriale e siano $U_1, \ldots, U_k$ dei sottospazi di V. Il sottospazio $\mathcal{L}(U_1 \cup \cdots \cup U_k)$ sarà detto sottospazio somma, oppure congiungente, di $U_1, \ldots, U_k$, e sarà indicato con il simbolo $U_1 + \cdots + U_k$.*

In perfetta analogia con il caso di due sottospazi, abbiamo la seguente proposizione e la successiva definizione.

2.49 Proposizione. *Sia V uno spazio vettoriale e siano $U_1, \ldots, U_k$ dei sottospazi di V. Si ha che*

$$U_1 + \cdots + U_k = \left\{ \mathbf{v} \in V \;\middle|\; \forall j = 1 \ldots, k \; \exists \, \mathbf{u}_j \in U_j \;\middle|\; \mathbf{v} = \mathbf{u}_1 + \cdots + \mathbf{u}_k \right\} .$$

Dimostrazione. È lasciata al lettore per esercizio. □

2.50 Definizione. *Sia V uno spazio vettoriale e siano $U_1, \ldots, U_k$ dei sottospazi di V. Il sottospazio congiungente $U_1 + \cdots + U_k$ sarà detto somma diretta, e sarà indicato con il simbolo $U_1 \oplus \cdots \oplus U_k$, se per ogni vettore $\mathbf{v} \in U_1 + \cdots + U_k$ i vettori $\mathbf{u}_1, \ldots, \mathbf{u}_k$ di cui alla proposizione precedente sono univocamente determinati, ovvero, in formule,*

$$\forall \, \mathbf{v} \in U_1 + \cdots + U_k, \; \forall \, j = 1, \ldots, k \; \exists! \, \mathbf{u}_j \in U_j \;\middle|\; \mathbf{v} = \mathbf{u}_1 + \cdots + \mathbf{u}_k \; .$$

Anche in questo caso osserviamo che, se i sottospazi in questione sono finitamente generati, anche il sottospazio congiungente è finitamente generato, e risulta che esso è una somma diretta se e solo se

$$\dim(U_1 + \cdots + U_k) = \dim U_1 + \cdots + \dim U_k \; .$$

Consideriamo ora il caso particolare in cui lo spazio V sia somma diretta di alcuni suoi sottospazi, ovvero esistano $U_1, \ldots, U_k \leq V$ tali che

$$V = U_1 \oplus \cdots \oplus U_k \; .$$

Esiste una relazione tra le basi dei sottospazi $U_1, \ldots, U_k$ e le basi di V, espressa nella seguente proposizione.

2.51 Proposizione. *Sia V uno spazio vettoriale e siano $U_1, \ldots, U_k$ dei sottospazi di V tali che*

$$V = U_1 \oplus \cdots \oplus U_k \ .$$

Consideriamo inoltre, per ogni $j = 1, \ldots, k$, una base $\mathcal{B}_j$ di U_j. Allora il sistema unione $\mathcal{B}$ è una base di V.

Dimostrazione. Omessa. □

2.4 Applicazioni lineari

Siano V, V' due spazi vettoriali su uno stesso campo $\mathbb{K}$, e sia $f : V \to V'$ un'applicazione tra (i sostegni di) tali spazi vettoriali.

2.52 Definizione. *Diremo che f è un'applicazione lineare, ovvero un omomorfismo, se sono verificate le seguenti condizioni:*

(i) $f(\mathbf{u}+\mathbf{v}) = f(\mathbf{u}) + f(\mathbf{v})$, $\forall\, \mathbf{u}, \mathbf{v} \in V$;

(ii) $f(\alpha\mathbf{v}) = \alpha f(\mathbf{v})$, $\forall\, \alpha \in \mathbb{K}$, $\forall\, \mathbf{v} \in V$.

Le condizioni (i) ed (ii) equivalgono alla condizione

- $f(\alpha\mathbf{u} + \beta\mathbf{v}) = \alpha f(\mathbf{u}) + \beta f(\mathbf{v})$, $\forall\, \alpha, \beta \in \mathbb{K}$, $\forall\, \mathbf{u}, \mathbf{v} \in V$

o anche alla condizione

- $f\Big(\sum_{i=1}^{m} \alpha_i \mathbf{v}_i\Big) = \sum_{i=1}^{m} \alpha_i f(\mathbf{v}_i)$, $\forall\, \alpha_1, \ldots, \alpha_m \in \mathbb{K}$, $\forall\, \mathbf{u}_1, \ldots, \mathbf{u}_m \in V$,

dove $m \in \mathbb{N}$.

Lo spazio V è detto *dominio* dell'applicazione lineare f, lo spazio V' è detto invece *codominio*. Il sottoinsieme

$$f(V) = \big\{\, \mathbf{v}' \in V' \;\big|\; \exists\, \mathbf{v} \in V \;\big|\; f(\mathbf{v}) = \mathbf{v}' \,\big\} \subseteq V'$$

è detto *immagine* di f ed è di solito denotato $\operatorname{im} f$, e non va confuso con il codominio V'. Infatti, in generale $\operatorname{im} f \subseteq V'$ e vale l'uguaglianza se e solo se f è suriettiva.

Dalla definizione di applicazione lineare discendono le seguenti facili proprietà:

(i) $f(\mathbf{0}) = \mathbf{0}$;

(ii) $f(-\mathbf{v}) = -f(\mathbf{v})$, $\forall\, \mathbf{v} \in V$.

In seguito ci capiterà di considerare delle applicazioni lineari il cui codominio V' coincide con il campo $\mathbb{K}$. Esse prendono il nome di *forme lineari*.

Focalizziamo ora la nostra attenzione sul caso in cui il dominio $V = \mathbb{K}^n$ ed il codominio $V' = \mathbb{K}^m$ dell'applicazione lineare f sono entrambi spazi vettoriali numerici, con $m, n \in \mathbb{N}$. Per ogni $i = 1, \ldots, m$ definiamo un'applicazione $\pi_i : \mathbb{K}^m \to \mathbb{K}$, detta *proiezione* i-ma, ponendo

$$\pi_i(a_1, \ldots, a_m) = a_i \ .$$

È immediato verificare che tale applicazione è lineare. Sia ora $f : \mathbb{K}^n \to \mathbb{K}^m$ un'applicazione, e poniamo $f_i = \pi_i \circ f$. Ad esempio, se $n = 3$ e $m = 4$ e si ha che $f(a_1, a_2, a_3) = (b_1, b_2, b_3, b_4)$, allora $f_3(a_1, a_2, a_3) = b_3$. Le applicazioni composte f_i si dicono *componenti* di f. Abbiamo la seguente caratterizzazione delle applicazioni lineari nel caso numerico.

2.53 PROPOSIZIONE. *Consideriamo un'applicazione* $f : \mathbb{K}^n \longrightarrow \mathbb{K}^m$ *tra spazi numerici. Con le notazioni appena introdotte, abbiamo che l'applicazione* f *è lineare se e solo se tali sono le sue componenti* $f_1, \ldots, f_m$

Dimostrazione. Omessa. □

In altri termini un'applicazione tra spazi numerici è lineare se e solo se le sue componenti sono forme lineari. È pertanto importante poter caratterizzare le forme lineari su uno spazio numerico.

2.54 TEOREMA DI CARATTERIZZAZIONE DELLE FORME LINEARI. *Consideriamo un'applicazione* $g : \mathbb{K}^n \to \mathbb{K}$. *Tale applicazione* g *è lineare (e quindi è una forma) se e solo se esistono degli scalari* $\gamma_1, \ldots, \gamma_n$ *tali che si abbia*

$$g(x_1, \ldots, x_n) = \gamma_1 x_1 + \cdots + \gamma_n x_n \ , \ \forall\ (x_1, \ldots, x_n) \in \mathbb{K}^n \ .$$

Dimostrazione. Omessa. □

Gli scalari $\gamma_1, \ldots, \gamma_n$ si dicono talvolta *coefficienti* della forma lineare g.

2.55 ESEMPIO. *Siano* V, V' *due spazi vettoriali su un campo* $\mathbb{K}$ *e sia* $f : V \to V'$ *l'applicazione nulla, definita ponendo* $f(\mathbf{u}) = \mathbf{0}$ *per ogni* $\mathbf{u} \in V$. *Tale applicazione è lineare, e si dice* omomorfismo nullo.

2.56 ESEMPIO. *Sia* V *uno spazio vettoriale su un campo* $\mathbb{K}$. *L'applicazione identica* $id_V : V \to V$, *definita ponendo* $id_V(\mathbf{u}) = \mathbf{u}$ *per ogni* $\mathbf{u} \in V$, *è lineare.*

2.57 ESEMPIO. *Sia V uno spazio vettoriale su un campo $\mathbb{K}$ e sia $k \in \mathbb{K}$. L'applicazione $\psi_k : V \to V$ definita ponendo $\psi_k(\mathbf{u}) = k\mathbf{u}$, per ogni $\mathbf{u} \in V$ è lineare, ed è nota come* omotetìa *di ragione k. Si osservi che per $k = 1$ si ottiene l'identità, mentre per $k = 0$ si ottiene l'omomorfismo nullo.*

2.58 ESEMPIO. *Consideriamo lo spazio vettoriale numerico $V = \mathbb{R}^2$ e lo spazio vettoriale dei polinomi $V' = \mathbb{R}[x]$ sul campo reale. Definiamo un'applicazione $f : \mathbb{R}^2 \to \mathbb{R}[x]$ ponendo $f(a, b) = ax + b$. Tale applicazione è lineare.*

2.59 ESEMPIO. *Consideriamo lo spazio vettoriale dei polinomi $V = \mathbb{R}[x]$ sul campo reale. Consideriamo poi l'applicazione $D : \mathbb{R}[x] \to \mathbb{R}[x]$ che associa ad ogni polinomio la sua derivata. In altri termini, per ogni polinomio $p = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$, poniamo $D(p) = a_1 + 2a_2x + \cdots + na_nx^{n-1}$. Tale applicazione è lineare.*

2.60 DEFINIZIONE. *Sia V un $\mathbb{K}$-spazio vettoriale. Un'applicazione lineare $f : V \to V$ si dice endomorfismo.*

Le applicazioni lineari di cui negli Esempi 2.56, 2.57, 2.59 sono quindi endomorfismi. Torniamo ora al caso generale.

2.61 DEFINIZIONE. *Consideriamo un'applicazione lineare $f : V \to V'$. Il sottoinsieme*

$$\ker f = f^{-1}(\mathbf{0}) = \left\{ \mathbf{v} \in V \;\middle|\; f(\mathbf{v}) = \mathbf{0} \right\} \subseteq V$$

si dice nucleo di f.

2.62 PROPOSIZIONE. *Il sottoinsieme* $\ker f$ *è un sottospazio di V, ed il sottoinsieme* $\operatorname{im} f$ *è un sottospazio di V'.*

Dimostrazione. È lasciata al lettore per esercizio. □

2.63 DEFINIZIONE. *Consideriamo un'applicazione lineare $f : V \to V'$. Diremo che f è un monomorfismo se è iniettiva, un epimorfismo se è suriettiva, un isomorfismo se è biettiva.*

Da tale definizione si deduce che l'applicazione lineare f è un epimorfismo se e solo se $\operatorname{im} f = V'$. Nel caso in cui V' sia finitamente generato, ciò equivale a dire che $\dim \operatorname{im} f = \dim V'$. Esiste una caratterizzazione corrispondente per i monomorfismi.

2.64 Teorema di caratterizzazione dei monomorfismi. *Sia $f : V \to V'$ un'applicazione lineare. Tale applicazione f è un monomorfismo se e solo se* $\ker f = \{\mathbf{0}\}$.

Dimostrazione. Se f è iniettiva, per ogni vettore non nullo $\mathbf{v} \in V_*$ si ha che $f(\mathbf{v}) \neq f(\mathbf{0}) = \mathbf{0}$, e quindi $\mathbf{v} \notin \ker f$. Pertanto $\ker f$ si riduce al solo vettore nullo. Viceversa, supponiamo che $\ker f = \{\mathbf{0}\}$. Siano $\mathbf{u}, \mathbf{v} \in V$ e sia $\mathbf{u} \neq \mathbf{v}$. Allora non può accadere che $f(\mathbf{u}) = f(\mathbf{v})$, altrimenti risulterebbe $f(\mathbf{u} - \mathbf{v}) = \mathbf{0}$ pur essendo $\mathbf{u} - \mathbf{v} \neq \mathbf{0}$. Deve quindi essere $f(\mathbf{u}) \neq f(\mathbf{v})$ ed f è iniettiva. □

Osserviamo che, poiché un isomorfismo è un'applicazione lineare biettiva, esso è sicuramente dotato di inversa. La proposizione seguente ci assicura che tale inversa è ancora un'applicazione lineare.

2.65 Proposizione. *Sia $f : V \to V'$ un isomorfismo. Allora anche l'applicazione inversa $f^{-1} : V' \to V$ è lineare, ed anche biettiva, e cioè un isomorfismo.*

Dimostrazione. Poiché f è biettiva, ha senso considerare la sua inversa f^{-1} che risulta a sua volta biettiva. Resta da verificare che f^{-1} è una applicazione lineare. Siano dunque $\alpha, \beta \in \mathbb{K}$, $\mathbf{u}', \mathbf{v}' \in V'$. Essendo f biettiva, esiste un unico $\mathbf{u} \in V$ tale che $\mathbf{u}' = f(\mathbf{u})$ ed esiste un unico $\mathbf{v} \in V$ tale che $\mathbf{v}' = f(\mathbf{v})$. Pertanto $\mathbf{u} = f^{-1}(\mathbf{u}')$, $\mathbf{v} = f^{-1}(\mathbf{v}')$ e si ha che

$$\begin{aligned} f^{-1}(\alpha\mathbf{u}' + \beta\mathbf{v}') &= f^{-1}\big(\alpha f(\mathbf{u}) + \beta f(\mathbf{v})\big) \\ &= f^{-1}\big(f(\alpha\mathbf{u} + \beta\mathbf{v})\big) \\ &= \alpha\mathbf{u} + \beta\mathbf{v} \\ &= \alpha f^{-1}(\mathbf{u}') + \beta f^{-1}(\mathbf{v}') \end{aligned}$$

e quindi f^{-1} è lineare. □

Consideriamo ora un'applicazione lineare $f : V \to V'$, un sistema di vettori $\mathcal{S} = (\mathbf{v}_1, \ldots, \mathbf{v}_m)$ in V, ed infine il sistema $\mathcal{S}' = (f(\mathbf{v}_1), \ldots, f(\mathbf{v}_m))$ in V', costituito dai vettori immagine dei vettori di $\mathcal{S}$. Il seguente enunciato evidenzia alcune relazioni tra proprietà dei due sistemi.

2.66 Proposizione. *Nella situazione appena descritta, abbiamo che*

(i) se $\mathcal{S}$ è dipendente, anche $\mathcal{S}'$ è dipendente;

(ii) se $\mathcal{S}$ è indipendente ed f è un monomorfismo, anche $\mathcal{S}'$ è indipendente;

(iii) se $\mathcal{S}$ genera V, il sistema $\mathcal{S}'$ genera il sottospazio $\operatorname{im} f$ *di V';*

(iv) se $\mathcal{S}$ genera V ed f è un epimorfismo, il sistema $\mathcal{S}'$ genera V';

(v) se $\mathcal{S}$ è una base di V ed f è un isomorfismo, allora $\mathcal{S}'$ è una base di V'.

Dimostrazione. Omessa. □

Osserviamo che le parti (ii), (iv) e (v) della proposizione precedente si invertono, nel senso che se f è un'applicazione lineare e conserva l'indipendenza lineare, ovvero trasforma vettori indipendenti in vettori indipendenti, allora f è un monomorfismo, e, analogamente, se f trasforma sistemi di generatori di V in sistemi di generatori di V', allora f è un epimorfismo. Infine, se f trasforma basi di V in basi di V', allora f è un isomorfismo.

2.67 Definizione. *Due spazi vettoriali V e V' si dicono isomorfi, e si scrive $V \cong V'$, se esiste un isomorfismo $f : V \to V'$.*

In tal caso scriveremo talvolta $f : V \xrightarrow{\cong} V'$. Osserviamo che quando esiste un isomorfismo tra due spazi, esso non è, in generale, unico. Dalla proposizione precedente si deduce il seguente corollario.

2.68 Corollario. *Se gli spazi vettoriali V e V' sono isomorfi e finitamente generati, si ha che* $\dim V = \dim V'$.

Dimostrazione. È lasciata al lettore. □

Vedremo in seguito che vale anche il viceversa. A tale scopo studiamo ora un metodo per costruire un'applicazione lineare a partire da alcuni input iniziali. Consideriamo una base $\mathcal{B} = (\mathbf{e}_1, \dots, \mathbf{e}_n)$ dello spazio V ed un sistema $\mathcal{S}' = (\mathbf{v}'_1, \dots, \mathbf{v}'_n)$ di vettori di V', dello stesso ordine. Sottolineiamo che $\mathcal{S}'$ è del tutto arbitrario. Ad esempio può essere dipendente o indipendente, può contenere il vettore nullo, in esso possono comparire delle ripetizioni.

2.69 Proposizione. *Nella situazione sopra descritta, esiste una ed una sola applicazione lineare $f : V \to V'$ tale che $f(\mathbf{e}_i) = \mathbf{v}'_i$, per ogni $i = 1, \dots, n$.*

Dimostrazione. Omessa. □

Quando si utilizza questa proposizione, si dice che l'applicazione f è costruita definendo i suoi valori sulla base $\mathcal{B}$ e poi *estendendo per linearità*. Al di là della dimostrazione di tale enunciato, desideriamo dare un cenno sulla costruzione di f. Se $\mathbf{v} \in V$ è un qualunque vettore del dominio, il Teorema di caratterizzazione delle basi ci assicura che esiste un'unica n-pla di scalari $(\alpha_1, \dots, \alpha_n)$ tale che

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

Basta allora porre

$$f(\mathbf{v}) = \alpha_1 \mathbf{v}'_1 + \cdots + \alpha_n \mathbf{v}'_n \ .$$

Si verifica facilmente che l'applicazione così definita è lineare, si comporta come richiesto sui vettori della base, ed è l'unica con tali proprietà. Il lettore potrà inoltre dimostrare, per esercizio, che f è un monomorfismo se e solo se il sistema $\mathcal{S}'$ è indipendente, è un epimorfismo se e solo se il sistema $\mathcal{S}'$ genera V', un isomorfismo se e solo se il sistema $\mathcal{S}'$ è una base di V'. In particolare, se due spazi hanno la stessa dimensione, essi risultano isomorfi. Vale inoltre il seguente utile risultato.

2.70 Teorema dell'equazione dimensionale. *Siano V, V' due spazi vettoriali su un campo $\mathbb{K}$, e sia V finitamente generato. Sia inoltre $f : V \to V'$ un'applicazione lineare. Allora il sottospazio* $\operatorname{im} f$ *di V' è finitamente generato e si ha che*

$$\dim V = \dim \ker f + \dim \operatorname{im} f \ .$$

Dimostrazione. Omessa. □

2.71 Esempio. *Sia V uno spazio vettoriale finitamente generato su un campo $\mathbb{K}$, e sia $\mathcal{B} = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$ una sua base ordinata. Il Teorema di caratterizzazione delle basi, parte (c), ci assicura che per ogni vettore $\mathbf{u} \in V$ esiste un'unica n-pla $(\alpha_1, \ldots, \alpha_n)$, detta n-pla delle componenti di $\mathbf{u}$ rispetto alla base $\mathcal{B}$, tale che $\mathbf{u} = \alpha_1 \mathbf{e}_1 + \cdots + \alpha_n \mathbf{e}_n$. Definiamo un'applicazione $\Phi_{\mathcal{B}} : V \to \mathbb{K}^n$ associando ad ogni vettore $\mathbf{u}$ la sua n-pla delle componenti rispetto alla base $\mathcal{B}$. Tale applicazione è lineare, e si verifica agevolmente che essa è un isomorfismo, noto come* isomorfismo coordinato. *Osserviamo che possiamo definire l'isomorfismo coordinato anche ponendo, per ogni $i = 1, \ldots, n$, $\Phi_{\mathcal{B}}(\mathbf{e}_i) = (0, \ldots, 1, \ldots, 0)$ (l'i-mo elemento della base standard di $\mathbb{K}^n$, ovvero la n-pla in cui l'unico elemento non nullo, che è 1, compare nella posizione i-ma), ed estendendo poi per linearità.*

Capitolo 3

Matrici

3.1 Generalità sulle matrici

Sia S un insieme non vuoto.

3.1 Definizione. *Una matrice di tipo $m \times n$ su S è un'applicazione*

$$A : \{1, \dots, m\} \times \{1, \dots, n\} \longrightarrow S \ .$$

Pertanto ad ogni coppia $(i, j) \in \{1, \dots, m\} \times \{1, \dots, n\}$ risulta associato un elemento $A(i, j) \in S$ che si denota di solito con il simbolo $a_{i,j}$.

In questo capitolo assumeremo quasi sempre che l'insieme S sia un campo $\mathbb{K}$. Molti risultati valgono però in situazioni più generali (ad esempio quando S è un anello). È importante capire comunque che una matrice può essere definita su un insieme di elementi di qualunque natura.

Una matrice A è completamente determinata dagli scalari $a_{i,j}$ che possono essere disposti in una tabella per righe e per colonne, e quindi per convenzione scriveremo

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} \end{pmatrix}$$

o anche, in forma abbreviata

$$A = (a_{i,j}) \ .$$

I vettori numerici

$$A_1 = (a_{1,1}, a_{1,2}, \dots, a_{1,n}) \,, \dots, \ A_m = (a_{m,1}, a_{m,2}, \dots, a_{m,n})$$

si dicono righe di A. I vettori numerici

$$A^1 = \begin{pmatrix} a_{1,1} \\ \vdots \\ a_{m,1} \end{pmatrix}, \dots, A^n = \begin{pmatrix} a_{1,n} \\ \vdots \\ a_{m,n} \end{pmatrix}$$

si dicono colonne di A. Un vettore numerico di ordine n

$$(a_1, \dots, a_n)$$

può vedersi come matrice di tipo $1 \times n$ e prende il nome di *vettore (numerico) riga*. Analogamente, un vettore numerico di ordine m

$$\begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$$

scritto in forma verticale, può vedersi come matrice di tipo $m \times 1$ e prende il nome di *vettore (numerico) colonna*. Indicheremo con $\mathcal{M}_{m,n}(\mathbb{K})$, o più semplicemente con $\mathcal{M}_{m,n}$, l'insieme delle matrici di tipo $m \times n$ su $\mathbb{K}$.

3.2 Definizione. *Sia $A \in \mathcal{M}_{m,n}$. Si dice trasposta di A, e si indica con uno dei simboli ${}^tA$, ${}_tA$, A_t, la matrice $B = (b_{i,j}) \in \mathcal{M}_{n,m}$ definita ponendo $b_{i,j} = a_{j,i}$.*

La matrice ${}^tA$ si ottiene da A scambiando le righe con le colonne. Diremo che una matrice A è *quadrata* di ordine n se $A \in \mathcal{M}_{n,n}$, ovvero se A possiede n righe e n colonne. La matrice quadrata A sarà poi detta *diagonale* se per ogni coppia (i,j) tale che $i \neq j$ si ha che $a_{i,j} = 0$. Il vettore numerico

$$(a_{1,1}, \dots, a_{n,n})$$

prende il nome di *diagonale principale* di A. Diremo inoltre che A è una matrice *triangolare bassa*, ovvero triangolare inferiore, se per ogni coppia (i,j) tale che $i < j$ si ha che $a_{i,j} = 0$. A è invece una matrice *triangolare alta*, ovvero triangolare superiore, se per ogni coppia (i,j) tale che $i > j$ si ha che $a_{i,j} = 0$. Una matrice quadrata A si dirà poi *simmetrica* se $a_{i,j} = a_{j,i}$, *antisimmetrica* se $a_{i,j} = -a_{j,i}$, per ogni i,j. Osserviamo che gli elementi della diagonale principale di una matrice antisimmetrica sono tutti nulli, che una matrice è diagonale se e solo se essa è simultaneamente triangolare alta e bassa e che una matrice è simmetrica se e solo se coincide con la sua trasposta.

3.3 Esempio. *Se poniamo*

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

la trasposta di A è la matrice

$$ {}^tA = \begin{pmatrix} 1 & 0 & 1 \\ 2 & 2 & 3 \end{pmatrix} . $$

La matrice

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

è triangolare alta, mentre la sua trasposta

$$ {}^tB = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \\ 2 & 3 & 1 \end{pmatrix} $$

è triangolare bassa. Le matrici

$$ C = \begin{pmatrix} 1 & 1 & 3 \\ 1 & 2 & 2 \\ 3 & 2 & 0 \end{pmatrix} \quad ; \quad C' = \begin{pmatrix} 0 & 1 & 2 \\ -1 & 0 & -2 \\ -2 & 2 & 0 \end{pmatrix} $$

sono rispettivamente simmetrica ed antisimmetrica. Infine le matrici

$$ D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad ; \quad D' = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} $$

sono diagonali. La matrice in cui tutti gli scalari sono nulli si dice matrice nulla e si indica con il simbolo O, e si dovrà capire dal contesto il suo tipo.

Per ogni $n \in \mathbb{N}$ denotiamo con I_n la matrice

$$ I_n = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix} $$

che prende il nome di *matrice identità*, o anche *matrice identica*, di ordine n. Abbiamo che $I_n = (\delta_{i,j})$, dove $\delta_{i,j}$ è il simbolo di Kronecker (associato alla coppia (i,j)), definito ponendo

$$ \delta_{i,j} = \begin{cases} 1 & \text{se } i = j \\ 0 & \text{se } i \neq j \end{cases} \tag{3.1} $$

In $\mathcal{M}_{m,n}$ possiamo definire una addizione (interna) e una moltiplicazione esterna con operatori in $\mathbb{K}$ come segue. Sia λ uno scalare e siano $A = (a_{i,j}), B = (b_{i,j})$ due matrici di tipo $m \times n$. Poniamo

$$ A + B = C \ ; \quad \lambda \cdot A = D $$

dove $C = (c_{i,j}), D = (d_{i,j})$ e $c_{i,j} = a_{i,j} + b_{i,j},\ d_{i,j} = \lambda a_{i,j}$.

3.4 Proposizione. $(\mathcal{M}_{m,n}; +, \cdot)$ *è uno spazio vettoriale isomorfo a* $\mathbb{K}^{mn}$.

Dimostrazione. La verifica del fatto che $(\mathcal{M}_{m,n};+,\cdot)$ è uno spazio vettoriale è lasciata per esercizio. Si osserva poi che le mn matrici aventi un elemento uguale a 1 e tutti gli altri elementi nulli costituiscono una base per $(\mathcal{M}_{m,n};+,\cdot)$. □

L'elemento neutro dello spazio vettoriale $(\mathcal{M}_{m,n};+,\cdot)$ (rispetto all'addizione) è la matrice nulla O. Vogliamo ora definire una moltiplicazione tra matrici. A tale scopo consideriamo due matrici $A=(a_{i,j})\in\mathcal{M}_{m,n},\ B=(b_{j,h})\in\mathcal{M}_{n,s}$. Definiamo una matrice $C=(c_{i,h})\in\mathcal{M}_{m,s}$ ponendo

$$c_{i,h}=\sum_{j=1}^{n}a_{i,j}b_{j,h}\ .$$

In altri termini, si considerano la i-ma riga A_i di A e la h-ma colonna B^h di B e si effettua la somma dei prodotti delle componenti di ugual posto di tali vettori numerici. L'elemento $c_{i,h}$ è quindi il prodotto scalare standard dei vettori numerici A_i e B^h e possiamo scrivere

$$c_{i,h}=A_i\cdot B^h\ .$$

La matrice C così ottenuta si dice *prodotto righe per colonne* di A e B e si indica con $C=AB$.

Il lettore potrà provare per esercizio che per ogni $A\in\mathcal{M}_{m,n}$ si ha che

$$I_mA=A\ ;\quad AI_n=A\ .$$

3.5 PROPOSIZIONE. *Il prodotto righe per colonne tra matrici è distributivo rispetto alla somma, nel senso che se $A\in\mathcal{M}_{m,n}$ e $B,C\in\mathcal{M}_{n,s}$ si ha che*

$$A(B+C)=AB+AC\ .$$

Dimostrazione. Siano i,h tali che $1\le i\le m,\ 1\le h\le s$. Si ha che

$$A_i(B+C)^h=A_i(B^h+C^h)=A_iB^h+A_iC^h\ .$$

Pertanto l'elemento di posto (i,h) di $A(B+C)$ coincide con l'elemento di posto (i,h) di $AB+AC$. □

Analogamente si verifica che se $A,B\in\mathcal{M}_{m,n}$ e $C\in\mathcal{M}_{n,s}$ allora $(A+B)C=AC+BC$.

3.6 PROPOSIZIONE. *Il prodotto righe per colonne tra matrici è associativo, nel senso che se $A\in\mathcal{M}_{m,n},\ B\in\mathcal{M}_{n,s},\ C\in\mathcal{M}_{s,t}$ allora*

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

Dimostrazione. Osserviamo intanto che tutti i vari prodotti che compaiono nell'enunciato hanno senso e che $(AB)C,\ A(BC) \in \mathcal{M}_{m,t}$. Poniamo

$$D = (AB)C \ ; \quad E = A(BC)$$

e siano, come al solito,

$$A = (a_{i,j}),\ B = (b_{j,h}),\ C = (c_{h,k}),\ D = (d_{i,k}),\ E = (e_{i,k}) \ .$$

Vogliamo provare che $D = E$, ovvero che $d_{i,k} = e_{i,k}$ per ogni i, k. Per definizione

$$d_{i,k} = (AB)_i \cdot C^k \ .$$

Si ha che

$$(AB)_i = (A_i \cdot B^1, \dots, A_i \cdot B^s)$$

e d'altra parte

$$C^k = \begin{pmatrix} c_{1,k} \\ \vdots \\ c_{s,k} \end{pmatrix}$$

quindi

$$\begin{aligned} d_{i,k} &= (A_i \cdot B^1)c_{1,k} + \cdots + (A_i \cdot B^s)c_{s,k} \\ &= \sum_{h=1}^{s} (A_i \cdot B^h)c_{h,k} \\ &= \sum_{h=1}^{s} \Big(\sum_{j=1}^{n} a_{i,j} b_{j,h} \Big) c_{h,k} \\ &= \sum_{j,h} a_{i,j} b_{j,h} c_{h,k} \ . \end{aligned}$$

In modo analogo si verifica che

$$e_{i,k} = \sum_{j,h} a_{i,j} b_{j,h} c_{h,k}$$

e l'asserto è provato. □

3.7 Definizione. *Sia $A \in \mathcal{M}_{n,n}$ una matrice quadrata. Diremo che A è invertibile se esiste una matrice A' tale che $AA' = I_n = A'A$.*

Se la matrice A' esiste, essa è unica, è necessariamente quadrata dello stesso ordine di A e si denota con il simbolo A^{-1}.

È possibile, ed anche molto utile, introdurre una nozione di invertibilità più debole.

3.8 Definizione. *Sia $A \in \mathcal{M}_{m,n}$. Diremo che A è invertibile a destra se esiste una matrice $B \in \mathcal{M}_{n,m}$ (che prende il nome di inversa destra di A) tale che $AB = I_m$. Diremo invece che A è invertibile a sinistra se esiste una matrice $C \in \mathcal{M}_{n,m}$ (che prende il nome di inversa sinistra di A) tale che $CA = I_n$.*

Se $A \in \mathcal{M}_{n,n}$ è invertibile e A^{-1} è inversa di A, è chiaro che essa è anche inversa destra e sinistra di A. Viceversa, se $A \in \mathcal{M}_{n,n}$ e B, C sono rispettivamente una inversa destra ed una inversa sinistra di A, allora $B = C = A^{-1}$. Infatti

$$C = CI_n = C(AB) = (CA)B = I_nB = B \ .$$

Pertanto la matrice inversa, se esiste, è unica. È possibile dimostrare che se A è quadrata di ordine n e B è una sua inversa sinistra, essa è anche inversa destra, e cioè è proprio l'inversa di A.

Osserviamo esplicitamente che nel caso di matrici non quadrate le nozioni di invertibilità a destra e a sinistra sono indipendenti tra loro. In particolare, se esiste una inversa sinistra, essa non è unica e non esistono inverse destre. Anologamente, se esiste una inversa destra, essa non è unica e non esistono inverse sinistre.

3.9 Esempio. *Come osservato, il prodotto righe per colonne AB può effettuarsi solo quando il numero di colonne di A coincide con il numero di righe di B. In particolare, per ogni $n \in \mathbb{N}$, tale prodotto è una operazione interna in $\mathcal{M}_{n,n}$ ed è facile verificare che $\mathcal{M}_{n,n}$ è un monoide rispetto a tale operazione, il cui elemento neutro è I_n. Il sottoinsieme $GL_n(\mathbb{K})$ di $\mathcal{M}_{n,n}$ costituito dalle matrici quadrate invertibili è un gruppo e prende il nome di* gruppo generale lineare *(su $\mathbb{K}$). Il sottoinsieme O_n costituito dalle matrici invertibili tali che $A^{-1} = {}^tA$ è anch'esso un gruppo e prende il nome di* gruppo ortogonale.

Sia $A \in \mathcal{M}_{m,n}$ e siano

$$\mathcal{R}_A = [A_1, \dots, A_m] \ ; \quad \mathcal{C}_A = [A^1, \dots, A^n] \ .$$

È chiaro che $\mathcal{R}_A$ è un sistema di vettori di $\mathbb{K}^n$ e $\mathcal{C}_A$ è un sistema di vettori di $\mathbb{K}^m$ (anche se scritti in forma di colonna).

3.10 Proposizione. *Sia $\mathcal{S} = [A_{i_1}, \dots, A_{i_p}]$ un sistema indipendente massimale di vettori di $\mathcal{R}_A$ e $\mathcal{S}' = [A^{j_1}, \dots, A^{j_q}]$ un sistema indipendente massimale di vettori di $\mathcal{C}_A$. Si ha che $p = q$.*

Dimostrazione. Omessa. □

In base a risultati già acquisiti, sappiamo che i sistemi indipendenti massimali di righe o di colonne di una fissata matrice A sono tutti equipotenti. Pertanto ha senso la seguente

3.11 Definizione. *Sia $A \in \mathcal{M}_{m,n}$. Il rango di A è il numero intero non negativo ρ_A (indicato anche con il simbolo $\rho(A)$) definito come segue. Se A è la matrice nulla, si pone $\rho_A = 0$. Altrimenti, ρ_A è l'ordine di un (qualunque) sistema indipendente massimale di righe (o anche di colonne) di A.*

In alcuni casi è conveniente evidenziare alcuni blocchi di elementi di una matrice. A tal proposito, se $A = (a_{i,j}) \in \mathcal{M}_{m,n}$ e $B = (b_{h,k}) \in \mathcal{M}_{s,t}$ con $s \leq m$ e $t \leq n$, diremo che B è un *blocco* di A se esistono due indici non negativi i_0, j_0 tali che

$$b_{h,k} = a_{i_0+h, j_0+k} \qquad (\, h = 1, \ldots, s\, ;\; k = 1, \ldots, t\,)\ ,$$

ovvero se B è costituita dagli elementi di A individuati dalle righe di posto $i_0 + 1, \ldots, i_0 + s$ e dalle colonne di posto $j_0 + 1, \ldots, j_0 + t$. Possiamo descrivere tale situazione con la notazione

$$A = \left(\begin{array}{c|c|c} \quad & \quad & \quad \\ \hline & B & \\ \hline & & \end{array}\right) .$$

Ad esempio, se

$$A = \begin{pmatrix} 11 & 12 & 13 & 14 \\ 15 & 16 & 17 & 18 \\ 19 & 20 & 21 & 22 \\ 23 & 24 & 25 & 26 \end{pmatrix}$$

e

$$B = \begin{pmatrix} 11 & 12 \\ 15 & 16 \end{pmatrix} \ ; \quad C = \begin{pmatrix} 13 & 14 \\ 17 & 18 \end{pmatrix}$$

$$D = \begin{pmatrix} 19 & 20 \\ 23 & 24 \end{pmatrix} \ ; \quad E = \begin{pmatrix} 21 & 22 \\ 25 & 26 \end{pmatrix}$$

si scrive

$$A = \left(\begin{array}{c|c} B & C \\ \hline D & E \end{array}\right)$$

o anche

$$A = \left(\begin{array}{c|c} {\scriptstyle \begin{matrix} 11 & 12 \\ 15 & 16 \end{matrix}} & C \\ \hline D & E \end{array}\right)$$

e così via. Il lettore potrà verificare, per esercizio, che se $A = (A' | A'')$ è una matrice a blocchi con

$$A \in \mathcal{M}_{n,s+t} \ ; \quad A' \in \mathcal{M}_{n,s} \ ; \quad A'' \in \mathcal{M}_{n,t}$$

e $B \in \mathcal{M}_{m,n}$ allora

$$B \cdot A = \big(B \cdot A' \,|\, B \cdot A'' \big) \ .$$

Vogliamo ora introdurre una relazione di equivalenza nell'insieme $\mathcal{M}_{m,n}$. Indichiamo con $R_1, \ldots, R_m$ le righe e con $C^1, \ldots, C^n$ le colonne di una generica matrice.

3.12 Definizione. *Sia $A \in \mathcal{M}_{m,n}$. Una operazione elementare sulle righe di A è una trasformazione di uno dei seguenti tipi:*

(i) moltiplicare una riga A_i per uno scalare non nullo λ;

(ii) scambiare di posto due righe A_i, A_j;

(iii) sostituire la riga A_i con $A_i + \lambda A_j$, $(\lambda \in \mathbb{K},\ i \neq j)$.

Analoghe operazioni sono definite sulle colonne. L'operazione elementare (i) si denota con $R_i \rightsquigarrow \lambda R_i$, la (ii) con $R_i \leftrightsquigarrow R_j$, la (iii) con $R_i \rightsquigarrow R_i + \lambda R_j$. Analoghe notazioni si adottano anche per le operazioni elementari sulle colonne. Osserviamo che iterando l'operazione elementare (iii) si ottiene una operazione del tipo $R_i \rightsquigarrow R_i + \lambda_1 R_{j_1} + \cdots + \lambda_h R_{j_h}$.

Per meglio comprendere il significato delle operazioni introdotte, il lettore potrà verificare, a titolo di esercizio, che se $A, B \in \mathcal{M}_{m,n}$ e supponiamo che la matrice B sia ottenuta da A mediante una operazione elementare (sulle righe), i sottospazi di $\mathbb{K}^n$ generati rispettivamente dai sistemi delle righe di A e di B coincidono.

3.13 Definizione. *Siano $A, B \in \mathcal{M}_{m,n}$. Diremo che A e B sono equivalenti (e scriveremo $A \equiv B$) se B si ottiene da A mediante un numero finito di operazioni elementari.*

3.14 Lemma. *La relazione $\equiv$ è di equivalenza in $\mathcal{M}_{m,n}$.*

Dimostrazione. L'unica proprietà che merita un commento è quella simmetrica. Osserviamo a tal proposito che le operazioni elementari di tipo (i) e (ii) sono chiaramente reversibili. Supponiamo ora che la matrice B sia ottenuta dalla matrice A mediante l'operazione (di tipo (iii)) $R_h \rightsquigarrow R_h + \lambda R_k$, con $\lambda \neq 0$. Allora anche A si ottiene da B con un'analoga operazione, ovvero $R_h \rightsquigarrow R_h - \lambda R_k$. $\square$

3.15 Definizione. *Una matrice A si dice a scala (secondo le righe) se sono verificate le seguenti condizioni:*

(i) se $A_i = \mathbf{0}$ allora $A_{i+1} = \mathbf{0}$;

(ii) per ogni i, se $a_{i,j} \neq 0$ e $a_{i,h} = 0\ \forall\ h < j$ allora $a_{i+1,h} = 0\ \forall\ h \leq j$.

Non è difficile verificare che una matrice a scala è caratterizzata dalla seguente proprietà:

Il primo elemento non nullo di ogni riga (non nulla) è più a destra rispetto al primo elemento non nullo della riga precedente (e le eventuali righe nulle sono tutte in basso).

In ogni riga non banale il primo elemento non nullo prende il nome di *pivot*.

3.16 ESEMPIO. *Poniamo*

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

A e C sono a scala. B non lo è poiché $B_2 = \mathbf{0}$ ma $B_3 \neq \mathbf{0}$. Osserviamo che C si ottiene da B mediante la sequenza di operazioni $R_2 \leftrightsquigarrow R_3$, $R_3 \leftrightsquigarrow R_4$.

3.17 TEOREMA. *Ogni matrice è equivalente ad una matrice a scala.*

Dimostrazione. Dimostriamo l'asserto in $\mathcal{M}_{m,n}$ per induzione su m. Se $m = 1$, per ogni n le matrici di tipo $1 \times n$ si riducono a vettori riga, e quindi sono già a scala. Sia dunque $m > 1$ e supponiamo che l'asserto sia vero per matrici di $\mathcal{M}_{s,t}$, per ogni $s < m$ e per ogni t. Sia $A \in \mathcal{M}_{m,n}$, sia A^j la prima colonna non banale e sia $i = \min\{\, h \mid a_{h,j} \neq 0 \,\}$. In altri termini $a_{i,j}$ è il primo elemento non banale della prima colonna non banale:

$$A = \left(\begin{array}{c|c|c} O & O & \ddots \\ \hline O & a_{i,j} & \cdots \\ \hline O & \vdots & \ddots \end{array}\right)$$

Si deve intendere che gli elementi dei blocchi contenenti puntini possono annullarsi o non annullarsi. Con l'operazione elementare $R_i \leftrightsquigarrow R_1$ si ottiene la matrice

$$D = \left(\begin{array}{c|c|c|c|c} O & \begin{matrix} a_{i,j} \\ \vdots \end{matrix} & D^{j+1} & \cdots & D^n \end{array}\right) .$$

Tale operazione è ridondante se $i = 1$. Applichiamo ora, per ogni $h = 2, \ldots, m$, le operazioni elementari $R_h \rightsquigarrow R_h - \frac{d_{h,j}}{a_{i,j}} R_1$ ed otteniamo una matrice del tipo

$$E = \left(\begin{array}{c|c|c} & a_{i,j} & \cdots \\ \cline{2-3} O & \begin{matrix} 0 \\ \vdots \\ 0 \end{matrix} & F \end{array}\right)$$

dove $F \in \mathcal{M}_{m-1,t}$ con $t < n$. Per ipotesi induttiva quindi F si trasforma in una matrice a scala, diciamo H, e si ottiene la matrice a scala

$$L = \left(\begin{array}{c|c|c} & a_{i,j} & \cdots \\ \cline{2-3} O & \begin{matrix} 0 \\ \vdots \\ 0 \end{matrix} & H \end{array}\right)$$

Formalmente, osserviamo che se ad esempio per trasformare F in H si usa l'operazione elementare $R_p \leftrightsquigarrow R_q$, in relazione alla matrice E si dovrà usare l'operazione elementare $R_{p+1} \leftrightsquigarrow R_{q+1}$, e in generale si dovranno incrementare di 1 tutti gli indici di riga che compaiono nelle operazioni elementari usate per trasformare F in H, poiché le righe che occupano le posizioni $1, \ldots, m-1$ di F compaiono in posizione $2, \ldots, m$ rispettivamente in E. Poiché L si ottiene da A con un numero finito di operazioni elementari, si ha che $A \equiv L$, e l'asserto è provato. $\square$

Il metodo usato nella dimostrazione del precedente teorema prende il nome di *algoritmo di Gauss* per la riduzione a scala di una matrice. La riduzione a scala di una matrice, ovvero la ricerca di una matrice a scala equivalente ad una matrice data, è molto utile per vari motivi. Abbiamo ad esempio il seguente

3.18 Teorema. *Due matrici equivalenti hanno lo stesso rango.*

Dimostrazione. Basta osservare che le operazioni elementari non cambiano la dipendenza ed indipendenza dei sistemi di righe o di colonne. Pertanto un metodo per determinare il rango di una matrice è quello di trasformare tale matrice in una matrice a scala e poi determinare il rango di quest'ultima matrice. $\square$

3.19 Teorema. *Il rango di una matrice a scala è il numero di righe non banali.*

Dimostrazione. Consideriamo una generica matrice a scala di tipo $m \times n$

$$H = \begin{pmatrix} 0 & \ldots & 0 & a_{1,j_1} & \ldots & & a_{1,j_2} & \ldots & & & \\ 0 & \ldots & 0 & 0 & \ldots & 0 & a_{2,j_2} & \ldots & & & \\ & \vdots & & & \vdots & & & \vdots & & & \\ 0 & \ldots & 0 & 0 & \ldots & 0 & 0 & \ldots & 0 & a_{k,j_k} & \ldots \\ \hline & & & & & O & & & & & \end{pmatrix}$$

dove $a_{1,j_1}, \ldots, a_{k,j_k} \neq 0$, $k \leq m, n$ e $j_1 < \cdots < j_k$. È chiaro che le (eventuali) righe nulle, ovvero quelle di posto $k+1, \ldots, m$ non contribuiscono alla determinazione del rango e che $\rho_H \leq k$. Basta quindi provare che il sistema $[H_1, \ldots, H_k]$ di vettori numerici è indipendente. Siano $\alpha_1, \ldots, \alpha_k$ degli scalari tali che

$$\alpha_1 H_1 + \cdots + \alpha_k H_k = \mathbf{0} \ .$$

Da tale relazione vettoriale discendono le seguenti relazioni scalari:

$$
\begin{aligned}
0 &= \alpha_1 a_{1,j_1} \ ; \\
0 &= \alpha_1 a_{1,j_2} + \alpha_2 a_{2,j_2} \ ; \\
\vdots \quad & \quad \vdots \\
0 &= \alpha_1 a_{1,j_k} + \alpha_2 a_{2,j_k} + \cdots + \alpha_k a_{k,j_k} \ .
\end{aligned}
$$

Dalla prima di tali relazioni si ricava che $\alpha_1 = 0$. Sostituendo tale valore nelle relazioni successive si deduce, dalla seconda, che $\alpha_2 = 0$, ecc. Si prova pertanto che tutti gli scalari α_i sono nulli. $\square$

3.2 Determinanti

Indichiamo, per ogni $n \in \mathbb{N}$, con J_n l'insieme $\{1, 2, \ldots, n\}$ dei primi n numeri naturali. Prima di introdurre la nozione di *determinante*, fissiamo alcune notazioni. Sia $A \in \mathcal{M}_{m,n}$. Se $i_1, \ldots, i_s \in J_m$, $j_1, \ldots, j_t \in J_n$, indichiamo con $A^{j_1,\ldots,j_t}_{i_1,\ldots,i_s}$ la matrice

$$
A^{j_1,\ldots,j_t}_{i_1,\ldots,i_s} = \begin{pmatrix} a_{i_1,j_1} & a_{i_1,j_2} & \cdots & a_{i_1,j_t} \\ a_{i_2,j_1} & a_{i_2,j_2} & \cdots & a_{i_2,j_t} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i_s,j_1} & a_{i_s,j_2} & \cdots & a_{i_s,j_t} \end{pmatrix}
$$

che si ottiene considerando solo gli elementi di A che si trovano simultaneamente sulle righe di posto $i_1, \ldots, i_s$ e sulle colonne di posto $j_1, \ldots, j_t$. In generale, gli indici $i_1, \ldots, i_s$ non sono a due a due distinti, e così pure $j_1, \ldots, j_t$. Ad esempio, se

$$
A = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 6 & 7 & 8 & 9 & 10 \\ 11 & 12 & 13 & 14 & 15 \end{pmatrix} \in \mathcal{M}_{3,5}
$$

si avrà

$$
A^{2,4}_{1,2} = \begin{array}{c} \\ \rightarrow \\ \rightarrow \\ \\ \end{array} \!\!\! \begin{array}{c} \begin{array}{ccccc} & \downarrow & & \downarrow & \end{array} \\ \begin{pmatrix} 1 & \mathbf{2} & 3 & \mathbf{4} & 5 \\ 6 & \mathbf{7} & 8 & \mathbf{9} & 10 \\ 11 & 12 & 13 & 14 & 15 \end{pmatrix} \end{array} = \begin{pmatrix} 2 & 4 \\ 7 & 9 \end{pmatrix} \ ; \quad A^{2,4,5}_{3,1,1} = \begin{pmatrix} 12 & 14 & 15 \\ 2 & 4 & 5 \\ 2 & 4 & 5 \end{pmatrix} \ .
$$

Nell'esempio appena fatto, le frecce indicano le righe e le colonne selezionate. Se gli indici di riga $i_1, \ldots, i_s$ sono a due a due distinti e compaiono nell'ordine naturale, e così pure gli indici di colonna $j_1, \ldots, j_t$, la matrice $A^{j_1,\ldots,j_t}_{i_1,\ldots,i_s}$ prende il nome di *matrice subordinata* (o anche *sottomatrice*) di A individuata dalle righe di posto $i_1, \ldots, i_s$ e dalle colonne di posto $j_1, \ldots, j_t$. In particolare, indicheremo con $A_{(i,j)}$ la matrice subordinata

$$
A_{(i,j)} = A^{1,2,\ldots,\hat{j},\ldots,n}_{1,2,\ldots,\hat{i},\ldots,m}
$$

individuata da tutte le righe (nel loro ordine naturale) tranne la i-ma e da tutte le colonne (nel loro ordine naturale) tranne la j-ma. In altri termini, $A_{(i,j)}$ si ottiene da A cancellando la i-ma riga e la j-ma colonna:

$$A_{(i,j)} = \begin{pmatrix} a_{1,1} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & \cdots & a_{m,j-1} & a_{m,j+1} & \cdots & a_{m,n} \end{pmatrix},$$

ovvero

$$A_{(i,j)} = \begin{array}{cc} & \begin{array}{ccccccc} \phantom{a_{1,1}} & & & \downarrow & & & \end{array} \\ \begin{array}{c} \\ \\ \\ \rightarrow \\ \\ \\ \\ \end{array} & \begin{pmatrix} a_{1,1} & \cdots & a_{1,j-1} & a_{1,j} & a_{1,j+1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ a_{i,1} & \cdots & a_{i,j-1} & \boldsymbol{a_{i,j}} & a_{i,j+1} & \cdots & a_{i,n} \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & \cdots & a_{m,j-1} & a_{m,j} & a_{m,j+1} & \cdots & a_{m,n} \end{pmatrix} \end{array},$$

dove le frecce indicano la riga e la colonna da cancellare, i cui elementi sono stati indicati in rosso. La matrice $A_{(i,j)}$ si dice *matrice complementare* dell'elemento $a_{i,j}$ in A. Osserviamo esplicitamente che se A è quadrata, anche $A_{(i,j)}$ è quadrata. In generale una matrice subordinata può essere quadrata oppure no, indipendentemente dal fatto che la matrice di partenza lo sia.

Sia ora $\mathbb{K}$ un campo, e sia $A \in \mathcal{M}_{m,n}$. Useremo le seguenti notazioni

$$A = (A^1, \ldots, A^n) \ ; \quad A = \begin{pmatrix} A_1 \\ \vdots \\ A_m \end{pmatrix}$$

per evidenziare le colonne o le righe che compongono A. In effetti le espressioni appena introdotte possono vedersi come particolari decomposizioni in blocchi di A. Vogliamo definire, per ogni $n \in \mathbb{N}$, un'applicazione

$$\det : \mathcal{M}_{n,n} \longrightarrow \mathbb{K}$$

che soddisfi le seguenti *proprietà* $\mathcal{D}$. Sia $A \in \mathcal{M}_{n,n}$ e siano B^k, C^k due vettori numerici colonna di ordine n e poniamo

$$B = (A^1, \ldots, A^{k-1}, B^k, A^{k+1}, \ldots, A^n) \ ,$$

$$C = (A^1, \ldots, A^{k-1}, C^k, A^{k+1}, \ldots, A^n) \ .$$

Proprietà $\mathcal{D}$.

$(i)_{\mathcal{D}}$ $\det I_n = 1$;

$(ii)_{\mathcal{D}}$ Se $\exists\ j \mid A^j = A^{j+1}$ allora $\det A = 0$;

$(iii)_{\mathcal{D}}$ Se $A^k = \beta B^k + \gamma C^k$, si ha che $\det A = \beta \det B + \gamma \det C \ \forall\ \beta, \gamma \in \mathbb{K}$.

La $(iii)_{\mathcal{D}}$ si esprime dicendo che l'applicazione det è lineare sulle colonne.

3.20 Esempio. *Una matrice 1×1 è del tipo (λ) dove λ è uno scalare. Se si pone*

$$\det(\lambda) = \lambda$$

è immediato verificare che le $\mathcal{D}$ sono soddisfatte. Sia $n = 2$; una matrice 2×2 è del tipo

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

Se si pone $\det A = ad - bc$ è agevole verificare che le $\mathcal{D}$ sono soddisfatte.

Il principale risultato di questo capitolo è espresso dal seguente enunciato.

3.21 Teorema di esistenza e unicità dei determinanti. *Per ogni intero positivo n esiste un'unica applicazione*

$$\det : \mathcal{M}_{n,n}(\mathbb{K}) \longrightarrow \mathbb{K}$$

che soddisfi le proprietà $\mathcal{D}$.

La dimostrazione di tale teorema è abbastanza complessa ed è omessa. Tale risultato rende lecita la seguente

3.22 Definizione. *L'applicazione* det *dell'enunciato precedente si dice applicazione determinante. Se A è una matrice quadrata di ordine n, lo scalare* $\det A$ *prende il nome di determinante di A.*

Talvolta il determinante di una matrice quadrata A viene indicato, oltre che con il simbolo $\det A$, anche con i simboli $|A|$, $D(A)$.

Dalle proprietà $\mathcal{D}$ si deducono anche le seguenti ulteriori proprietà.

(iv) Se in una matrice si scambiano tra loro due colonne, il determinante cambia di segno.

(v) Se la matrice A ha due colonne uguali, allora $\det A = 0$.

Il seguente teorema fornisce un metodo operativo per il calcolo dei determinanti, in quanto riduce il calcolo del determinante di una matrice quadrata di ordine n a calcoli legati a determinanti di matrici più piccole.

3.23 Teorema. *Sia $A \in \mathcal{M}_{n,n}$ e sia h un indice di riga. Si ha che*

$$\det A = \sum_{j=1}^{n} (-1)^{h+j} a_{h,j} \det A_{(h,j)} \ . \tag{3.2}$$

Questo teorema è un caso particolare di un risultato più generale che enunceremo tra breve, noto come il 1^o Teorema di Laplace. Quando si calcola il determinante utilizzando la (3.2), detta anche *Regola di Laplace*, si dice che si sviluppa $\det A$ secondo la h-ma riga. Dal Teorema 3.23 si deducono ulteriori proprietà dei determinanti.

(vi) Se la matrice A possiede una riga nulla, $\det A = 0$.

(vii) Se si aggiunge ad una colonna della matrice A una combinazione lineare delle altre colonne, il determinante della matrice così ottenuta coincide con $\det A$.

Il seguente teorema, di dimostrazione non facile, consente di utilizzare tutte le proprietà relative alle colonne anche sulle righe, e viceversa.

3.24 Teorema. *Sia $A \in \mathcal{M}_{n,n}$. Si ha che* $\det A = \det {}^tA$.

Poiché ${}^tA$ si ottiene da A scambiando le righe con le colonne, il risultato ora acquisito ci assicura che in tutti gli enunciati riguardanti i determinanti le righe e le colonne di una matrice possono essere scambiate. Ad esempio l'applicazione det è lineare sulle righe, se una matrice A possiede una colonna nulla allora $\det A = 0$, se una matrice A possiede due righe uguali allora $\det A = 0$, e così via. In particolare si verifica che si può sviluppare il determinante di una matrice quadrata anche secondo una colonna. Se si sceglie ad esempio la j-ma, si ottiene l'espressione

$$\det A = \sum_{i=1}^{n} (-1)^{i+j} a_{i,j} \det(A_{(i,j)})$$

che prende il nome di sviluppo di Laplace secondo la colonna j-ma. Talvolta lo scalare

$$(-1)^{i+j} \det A_{(i,j)}$$

si dice *complemento algebrico* dell'elemento di posto (i,j) in A, e quindi la Regola di Laplace può anche esprimersi con il seguente enunciato, valido sia per le righe che per le colonne

3.25 1^o Teorema di Laplace. *Il determinante di una matrice quadrata A è la somma dei prodotti degli elementi di una linea (riga o colonna) di A moltiplicati per i loro complementi algebrici.*

3.26 Definizione. *Sia $A \in \mathcal{M}_{n,n}$. Se $\det A \neq 0$ diremo che A è non singolare, o anche non degenere.*

Vediamo ora l'effetto delle operazioni elementari sul determinante di una matrice. Sia $A \in \mathcal{M}_{n,n}$. Se B si ottiene da A mediante $R_i \rightsquigarrow \lambda R_i$, dalle $\mathcal{D}$ discende che $\det B = \lambda \det A$. Se invece B si ottiene da A mediante $R_i \leftrightsquigarrow R_j$, allora $\det B = -\det A$. Se infine B si ottiene da A mediante $R_i \rightsquigarrow R_i + \lambda R_j$, allora $\det B = \det A$. Un discorso analogo vale se le operazioni vengono effettuate sulle colonne. In generale quindi se $A \equiv B$ allora

$$\det A = 0 \iff \det B = 0$$

ovvero A è singolare se e solo se tale è B. Le operazioni elementari sulle matrici consentono anche di dimostrare, con procedimenti sostanzialmente elementari, un importante risultato conosciuto come il Teorema di Binet, che consente di calcolare il determinante del prodotto di due matrici quadrate. Occorre premettere alcune considerazioni.

3.27 Lemma. *Siano $A \in \mathcal{M}_{n,n}$, $C \in \mathcal{M}_{m,n}$, $D \in \mathcal{M}_{m,m}$ e sia inoltre $O \in \mathcal{M}_{n,m}$ la matrice nulla. Posto*

$$E = \left(\begin{array}{c|c} A & O \\ \hline C & D \end{array}\right) \in \mathcal{M}_{m+n,m+n}$$

si ha che

$$\det E = \det A \cdot \det D \ .$$

3.28 Lemma. *Siano $A \in \mathcal{M}_{n,n}$, $B \in \mathcal{M}_{n,m}$, $D \in \mathcal{M}_{m,m}$ e sia inoltre $O \in \mathcal{M}_{m,n}$ la matrice nulla. Posto*

$$E = \left(\begin{array}{c|c} A & B \\ \hline O & D \end{array}\right) \in \mathcal{M}_{m+n,m+n}$$

si ha che

$$\det E = \det A \cdot \det D \ .$$

3.29 Teorema di Binet. *Siano $A, B \in \mathcal{M}_{n,n}$. Si ha che*

$$\det AB = \det A \cdot \det B \ .$$

Il Teorema di Binet ci consente, tra l'altro, di affrontare alcune questioni di invertibilità di matrici.

3.30 Corollario. *Sia $A \in \mathcal{M}_{n,n}$. Se A è invertibile si ha che* $\det A \neq 0$ *ed inoltre*

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

3.31 Osservazione. *Se $\mathbb{B}$ è un anello commutativo unitario, possiamo ancora definire, formalmente allo stesso modo, un'applicazione*

$$\det : \mathcal{M}_{n,n}(\mathbb{B}) \longrightarrow \mathbb{B} \ .$$

Naturalmente non tutte le proprietà dei determinanti di matrici su un campo saranno valide, ma resta valido, ad esempio, il Teorema di Binet. In seguito useremo tale generalizzazione nel caso $\mathbb{B} = \mathbb{K}[x]$, l'anello dei polinomi in una indeterminata a coefficienti in un campo $\mathbb{K}$. Ad esempio, sia

$$A = \begin{pmatrix} 3 & x^2+2 \\ x & x-1 \end{pmatrix} \ .$$

Tale matrice ha coefficienti in $\mathbb{B} = \mathbb{R}[x]$ e quindi il suo determinante sarà a sua volta un polinomio a coefficienti reali, ovvero

$$\det A = 3(x-1) - x(x^2+2) = -x^3 + x - 3 \in \mathbb{R}[x] \ .$$

Il non annullarsi del determinante è in realtà una condizione necessaria e sufficiente per l'invertibilità di una matrice quadrata. Per provare ciò, procediamo come segue.

3.32 2^o Teorema di Laplace. *Sia A una matrice quadrata di ordine n. Per ogni coppia (h,k) di indici distinti si ha che*

$$\sum_{j=1}^{n} (-1)^{k+j} a_{h,j} \det A_{(k,j)} = 0 \ ; \ \sum_{i=1}^{n} (-1)^{i+k} a_{i,h} \det A_{(i,k)} = 0 \ . \tag{3.3}$$

Il 2^o Teorema di Laplace si enuncia anche dicendo che è nulla la somma dei prodotti degli elementi di una riga (o colonna) di A per i complementi algebrici degli elementi di un'altra riga (colonna rispettivamente). Il 1^o e il 2^o Teorema di Laplace si sintetizzano nelle seguenti formule:

$$\sum_{j=1}^{n} (-1)^{k+j} a_{h,j} \det A_{(k,j)} = \delta_{h,k} \cdot \det A \tag{3.4}$$

dove $\delta_{h,k}$ è il simbolo di Kronecker definito nella (3.1).

$$\sum_{i=1}^{n} (-1)^{i+k} a_{i,h} \det A_{(i,k)} = \delta_{h,k} \cdot \det A \ . \tag{3.5}$$

Proviamo ora l'inverso del Corollario 3.30.

3.33 Proposizione. *Sia A una matrice quadrata e sia* $\det A \neq 0$. *Allora A è invertibile e la sua inversa $B = (b_{s,t})$ si ottiene ponendo*

$$b_{s,t} = (-1)^{s+t} \frac{\det A_{(t,s)}}{\det A} \ .$$

Da questa osservazione, dall'unicità dell'inversa e dal Corollario 3.30 si deduce il seguente

3.34 Teorema. *Sia $A \in \mathcal{M}_{n,n}$. A è invertibile se e solo se* $\det A \neq 0$. *In tal caso $A^{-1} = (b_{i,j})$, dove*

$$b_{i,j} = (-1)^{i+j} \frac{\det A_{(j,i)}}{\det A} \ .$$

Si ha inoltre che

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

Dal Teorema di Binet e dal Teorema 3.34 si deduce che se A, B sono matrici quadrate di ordine n e si ha che $A \cdot B = I_n$ allora A è invertibile e B è la sua inversa. Infatti per il Teorema di Binet si ha che

$$\det A \cdot \det B = \det(A \cdot B) = \det I_n = 1$$

e quindi $\det A \neq 0$ e in base al Teorema 3.34 la matrice A risulta invertibile. Pertanto B, che è l'inversa destra di A, deve coincidere con A^{-1}. A posteriori osserviamo quindi che se una matrice quadrata ha inversa destra (oppure sinistra) essa è l'inversa, ed è unica.

Il lettore potrà provare, a titolo di esercizio, che il determinante di una matrice triangolare è dato dal prodotto degli elementi della diagonale principale.

3.3 Matrici e dipendenza lineare

La nozione di determinante può essere usata per studiare la dipendenza o l'indipendenza di sistemi di vettori numerici. I risultati che otterremo per i sistemi di vettori numerici colonna sono naturalmente validi anche per i vettori numerici riga.

3.35 Teorema. *Siano $A^1, \ldots, A^n \in \mathbb{K}^n$ n vettori numerici colonna di ordine n e consideriamo il sistema $\mathcal{S} = [A^1, \ldots, A^n]$. Se $\mathcal{S}$ è dipendente, posto $A = (A^1, \ldots, A^n)$, si ha che* $\det A = 0$.

3.36 Corollario. *Sia $A \in \mathcal{M}_{n,n}$. Se* $\det A \neq 0$ *allora il sistema $\mathcal{S} = [A^1, \ldots, A^n]$ è indipendente.*

3.37 Corollario. *Sia* $A \in \mathcal{M}_{n,n}$ *e sia* $\det A \neq 0$. *Per ogni vettore numerico colonna* $\mathbf{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$ *esiste un'unica* n*-pla* $(\lambda_1, \dots, \lambda_n)$ *di scalari tale che*

$$\mathbf{b} = \sum_{j=1}^{n} \lambda_j A^j \ .$$

Useremo ora la teoria dei determinanti per fornire un metodo per il calcolo del rango di una matrice. Sia ora $A \in \mathcal{M}_{m,n}$ e sia $A_{i_1,\dots,i_p}^{j_1,\dots,j_p}$ una sottomatrice quadrata di A.

3.38 Definizione. *Il determinante* $\det A_{i_1,\dots,i_p}^{j_1,\dots,j_p}$ *prende il nome di minore di* A *(associato alla sottomatrice* $A_{i_1,\dots,i_p}^{j_1,\dots,j_p}$*). Se* $i_h = j_h$ *per ogni* h*, tale minore si dice principale.*

Consideriamo ora il sistema $\mathcal{C}_A = [A^1, \dots, A^n]$ delle colonne di A, che è un sistema di vettori numerici colonna di $\mathbb{K}^m$.

3.39 Proposizione. *Se* $n \leq m$ *ed esistono degli indici di riga* $i_1, \dots, i_n$ *tali che*

$$\det A_{i_1,\dots,i_n}^{1,\dots,n} \neq 0$$

allora $\mathcal{C}_A$ *è indipendente.*

Un ragionamento analogo vale se scambiamo le righe con le colonne.

3.40 Proposizione. *Sia* $m \leq n$ *e supponiamo che esistano degli indici di colonna* $j_1, \dots, j_m$ *tali che*

$$\det A_{1,\dots,m}^{j_1,\dots,j_m} \neq 0 \ .$$

Allora $[A_1, \dots, A_m]$ *è indipendente.*

È possibile usare i determinanti anche per studiare la dipendenza o indipendenza lineare di sistemi di vettori non numerici. A tale scopo, se V è uno spazio vettoriale di dimensione n, consideriamo l'isomorfismo coordinato

$$\Phi : V \longrightarrow \mathbb{K}^n$$

di V rispetto ad una fissata base ordinata. Poiché Φ è un isomorfismo, un sistema $\mathcal{S} = [\mathbf{v}_1, \dots, \mathbf{v}_m]$ di vettori di V è indipendente se e solo se tale è il sistema $\mathcal{S}' = [\Phi(\mathbf{v}_1), \ldots, \Phi(\mathbf{v}_m)]$ di vettori numerici di $\mathbb{K}^n$ e si possono quindi applicare i risultati fin qui esposti a proposito dei sistemi di vettori numerici.

Il metodo per il calcolo del rango di una matrice A che ora esporremo, è basato sullo studio dei minori di A. Se $A^{j_1,\ldots,j_q}_{i_1,\ldots,i_p}$ è una sottomatrice di A, i è un indice di riga e j è un indice di colonna, la sottomatrice $A^{j_1,\ldots,j_q,j}_{i_1,\ldots,i_p,i}$ prende il nome di *orlato* di $A^{j_1,\ldots,j_q}_{i_1,\ldots,i_p}$ mediante la i-ma riga e la j-ma colonna.

3.41 Teorema degli orlati. *Consideriamo una matrice $A \in \mathcal{M}_{m,n}$ e indichiamo con $\mathcal{R}_A$, $\mathcal{C}_A$ i sistemi di righe e colonne di A. Sia B una sua sottomatrice quadrata di ordine p tale che $\det B \neq 0$. Se ogni orlato B' della sottomatrice B risulta degenere, allora il sottosistema di $\mathcal{R}_A$ costituito dalle righe di A che passano per B è un sottosistema indipendente massimale di $\mathcal{R}_A$ e, analogamente, il sottosistema di $\mathcal{C}_A$ costituito dalle colonne di A che passano per B è un sottosistema indipendente massimale di $\mathcal{C}_A$. In particolare, $\rho(A) = p$.*

Il Teorema degli orlati fornisce un metodo per determinare il rango di una matrice A, che supponiamo non banale. Chiamiamo *sottomatrice fondamentale* di A una sottomatrice quadrata di un certo ordine, diciamo p, avente determinante non nullo e tale che ogni minore di ordine maggiore di p è nullo, ovvero una sottomatice con le proprietà della matrice B del Teorema degli orlati. Il minore associato ad una sottomatrice fondamentale di dice *minore fondamentale*. Se $A^{j_1,\ldots,j_p}_{i_1,\ldots,i_p}$ è una sottomatrice fondamentale, è chiaro che ogni suo orlato ha determinante nullo e quindi, in base al Teorema degli Orlati,

$$\mathcal{S} = [A_{i_1}, \ldots, A_{i_p}]$$

è un sistema massimale di righe indipendenti di A,

$$\mathcal{T} = [A^{j_1}, \ldots, A^{j_p}]$$

è un sistema massimale di colonne indipendenti di A ed inoltre $\rho(A) = p$. Di conseguenza ogni altro minore fondamentale di A avrà ordine p. Viceversa, se $\rho(A) = p$ e $\mathcal{S}$ è un sistema massimale di righe indipendenti di A, la sottomatrice $A^{1,\ldots,n}_{i_1,\ldots,i_p}$ ha rango p, poiché le sue p righe sono indipendenti. Allora, in base al Teorema degli orlati, deve esistere una sottomatrice quadrata di ordine p di $A^{1,\ldots,n}_{i_1,\ldots,i_p}$ con determinante non nullo. Essa sarà individuata da tutte le righe e da p colonne, ad esempio quelle di posto $j_1, \ldots, j_p$, che risultano indipendenti nella sottomatrice $A^{1,\ldots,n}_{i_1,\ldots,i_p}$. Pertanto la matrice $A^{j_1,\ldots,j_p}_{i_1,\ldots,i_p}$ è una sottomatrice fondamentale di A. Infatti ogni suo orlato ha determinante nullo, perché altrimenti troveremmo $p+1$ righe indipendenti. Quindi le colonne $A^{j_1}, \ldots, A^{j_p}$ sono indipendenti in A. A posteriori possiamo quindi dedurre che una sottomatrice fondamentale è una sottomatrice avente determinante non nullo e ordine pari al rango di A.

3.42 Corollario. *Ogni sottomatrice quadrata di A avente determinante non nullo è contenuta in una sottomatrice fondamentale, ovvero può essere orlata in modo opportuno (ripetutamente) fino ad ottenere una sottomatrice fondamentale.*

Pertanto, per determinare il rango di una matrice non banale $A = (a_{i,j})$ si può procedere al modo seguente. Si considera un elemento non nullo a_{i_1,j_1} di A e si orla la sottomatrice $A_{i_1}^{j_1}$ alla ricerca di una sottomatrice di ordine 2 con determinante non nullo. Se non esiste alcun orlato con questa proprietà, ovvero se per ogni scelta degli indici i_2, j_2 si ha che $\det A_{i_1,i_2}^{j_1,j_2} = 0$, allora $\rho(A) = 1$. Altrimenti $\rho(A) \geq 2$ e si continua ad orlare finché non si costruisce un minore fondamentale, il cui ordine sarà il rango di A. Una facile conseguenza del Teorema degli orlati è il seguente enunciato, che inverte il Corollario 3.36:

3.43 Proposizione. *Sia $\mathcal{S} = [A^1, \dots, A^n]$ un sistema indipendente di vettori numerici colonna di ordine n. Posto $A = (A^1, \dots, A^n)$ si ha che* $\det A \neq 0$.

Pertanto il determinante di una matrice A è non nullo se e solo se il sistema dei vettori colonna, o anche dei vettori riga, di A è indipendente, ovvero se e solo se $\rho(A) = n$ (e anche se e solo se A è invertibile).

Capitolo 4

Sistemi di equazioni lineari

4.1 Generalità sui sistemi lineari

Sia $f \in \mathbb{K}[x_1, \ldots, x_n]$ un polinomio di grado m in n indeterminate sul campo $\mathbb{K}$. L'espressione

$$f(x_1, \ldots, x_n) = 0$$

prende il nome di *equazione algebrica* di grado m e rappresenta il problema della ricerca delle radici di f ovvero delle n-ple $\xi = (\xi_1, \ldots, \xi_n) \in \mathbb{K}^n$ tali che

$$f(\xi_1, \ldots, \xi_n) = 0 \ .$$

Le indeterminate $x_1, \ldots, x_n$ si dicono *incognite* dell'equazione. Una radice ξ di f si dice anche *soluzione* dell'equazione. Se $m = 1$ tale equazione si dice *lineare*. In tal caso, essendo f un polinomio di primo grado, esistono degli scalari $a_1, \ldots, a_n$, detti coefficienti, e b, detto termine noto, tali che

$$f(x_1, \ldots, x_n) = a_1 x_1 + \cdots + a_n x_n - b$$

e quindi l'equazione può scriversi nella forma

$$a_1 x_1 + \cdots + a_n x_n - b = 0$$

o anche

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

4.1 Definizione. *Un sistema lineare di m equazioni in n incognite (ovvero un sistema lineare $m \times n$) è una espressione del tipo*

$$\left\{ \begin{array}{rcl} a_{1,1}x_1 + \cdots + a_{1,n}x_n &=& b_1 \\ a_{2,1}x_1 + \cdots + a_{2,n}x_n &=& b_2 \\ \vdots \qquad\qquad\quad \vdots && \vdots \\ a_{m,1}x_1 + \cdots + a_{m,n}x_n &=& b_m \end{array} \right. \tag{4.1}$$

e rappresenta il problema della ricerca delle eventuali soluzioni comuni alle equazioni lineari

$$\begin{aligned} a_{1,1}x_1 + \cdots + a_{1,n}x_n &= b_1 \\ a_{2,1}x_1 + \cdots + a_{2,n}x_n &= b_2 \\ \vdots \qquad\qquad \vdots \quad &\quad \vdots \\ a_{m,1}x_1 + \cdots + a_{m,n}x_n &= b_m \end{aligned}$$

Osserviamo che non ha alcuna importanza l'ordine in cui si considerano le equazioni del sistema lineare (4.1). Inoltre, in (4.1) le equazioni non sono necessariamente a due a due distinte. Un sistema lineare si dirà *compatibile* se ammette almeno una soluzione, *incompatibile* altrimenti. Diremo inoltre che due sistemi lineari nello stesso numero di incognite sono equivalenti se ammettono le stesse soluzioni. In particolare i sistemi incompatibili (nello stesso numero di incognite) sono tutti equivalenti e due sistemi equivalenti hanno necessariamente lo stesso numero di incognite. Se un sistema lineare (ad n incognite, è denotato, ad esempio, con il simbolo $(*)$, indicheremo con il simbolo $\mathrm{Sol}(*)$ l'insieme delle sue soluzioni, che è un sottoinsieme di $\mathbb{K}^n$.

4.2 Definizione. *Il sistema lineare* (4.1) *si dice omogeneo se*

$$b_1 = \cdots = b_m = 0 \ .$$

4.3 Definizione. *Dato un sistema lineare* $m \times n$ *della forma* (4.1) *il sistema lineare omogeneo*

$$\left\{ \begin{aligned} a_{1,1}x_1 + \cdots + a_{1,n}x_n &= 0 \\ a_{2,1}x_1 + \cdots + a_{2,n}x_n &= 0 \\ \vdots \qquad\qquad \vdots \quad &\quad \vdots \\ a_{m,1}x_1 + \cdots + a_{m,n}x_n &= 0 \end{aligned} \right. \tag{4.2}$$

prende il nome di sistema lineare omogeneo associato a (4.1).

Vogliamo ora scrivere una espressione del tipo (4.1) in forma compatta, usando notazioni matriciali. A tale scopo estendiamo formalmente la nozione di prodotto righe per colonne tra matrici anche al caso in cui gli elementi delle matrici siano di natura arbitraria. Con questa convenzione, posto

$$A = (a_{i,j}) \ ; \quad B = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \ ; \quad X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

i sistemi lineari (4.1),(4.2) possono scriversi nella forma

$$AX = B \ ; \quad AX = \mathbf{0}$$

rispettivamente, o anche nella forma

$$\begin{cases} A_1X = b_1 \\ A_2X = b_2 \\ \vdots \quad \vdots \\ A_mX = b_m \end{cases} ; \qquad \begin{cases} A_1X = 0 \\ A_2X = 0 \\ \vdots \quad \vdots \\ A_mX = 0 \end{cases}$$

dove $A_1, \ldots, A_m$ sono le righe di A e le espressioni AX, A_iX sono prodotti righe per colonne (formali). Sottolineiamo il fatto che i simboli $a_{i,j}, b_h$ rappresentano degli scalari, mentre gli x_k rappresentano le incognite del sistema lineare. Una ulteriore rappresentazione dei sistemi lineari (4.1),(4.2) è la seguente

$$A^1x_1 + \cdots + A^nx_n = B \ ; \quad A^1x_1 + \cdots + A^nx_n = \mathbf{0} \tag{4.3}$$

dove $A^1, \ldots, A^n$ sono le colonne di A.

4.4 Definizione. *La matrice A si dice matrice incompleta, o matrice dei coefficienti, o anche prima matrice, del sistema lineare* (4.1). *La matrice $A' = (A|B)$ si dice invece matrice completa, o anche seconda matrice, del sistema lineare* (4.1).

Cercheremo ora di stabilire dei criteri per capire se un sistema lineare è compatibile e per determinarne le eventuali soluzioni.

4.5 Teorema. *Il sistema lineare* (4.1) *è compatibile se e solo se il vettore numerico colonna dei termini noti B dipende linearmente dalle colonne di A.*

Dimostrazione. Sia $\xi = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}$ una soluzione del sistema lineare in questione. Ciò vuol dire, usando l'espressione (4.3), I parte, del sistema lineare, che

$$\xi_1A^1 + \cdots + \xi_nA^n = B \tag{4.4}$$

e quindi B dipende dal sistema di vettori colonna $[A^1, \ldots, A^n]$.
Viceversa, se B dipende dal sistema di vettori colonna $[A^1, \ldots, A^n]$, esistono degli scalari $\xi_1, \ldots, \xi_n$ tali che valga la (4.4) e dunque la n-pla $\xi = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}$ è una soluzione del sistema lineare (4.1). □

Osserviamo che, poiché il vettore nullo dipende da ogni sistema di vettori, ogni sistema lineare omogeneo è compatibile, ammettendo almeno la soluzione banale $\mathbf{0} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$. Si verifica immediatamente, d'altra parte, che se $\mathbf{0}$ è soluzione di un sistema lineare, tale sistema lineare è necessariamente omogeneo. Nel seguito indicheremo con

$$\mathcal{C} = [A^1, \dots, A^n] \ ; \quad \mathcal{R} = [A_1, \dots, A_m]$$

i sistemi dei vettori colonna e dei vettori riga della matrice A e con $\mathcal{C}'$, $\mathcal{R}'$ i sistemi dei vettori colonna e dei vettori riga della matrice A'. Osserviamo che poiché $\mathcal{C} \subseteq \mathcal{C}'$ si ha che $\rho(A) \leq \rho(A')$.

4.6 Teorema di Rouché-Capelli. *Il sistema lineare* (4.1) *è compatibile se e solo se*

$$\rho(A) = \rho(A') \ . \tag{4.5}$$

Dimostrazione. Sia (4.1) compatibile. In base al Teorema 4.5, B dipende dal sistema $\mathcal{C}$. Pertanto se $\rho(A) = p$ e

$$\mathcal{D} = [A^{j_1}, \dots, A^{j_p}]$$

è un sistema indipendente massimale di colonne di A, ogni vettore di $\mathcal{C}$ dipende da $\mathcal{D}$ e quindi anche B dipende da $\mathcal{D}$. Ma allora $\mathcal{D}$ è anche un sistema indipendente massimale di colonne di A'. Pertanto $\rho(A') = p$ e la prima parte dell'enunciato è dimostrata.
Viceversa, supponiamo che valga la (4.5), e sia $\mathcal{D}$ un sistema indipendente massimale di colonne di A. Poiché $\rho(A') = \rho(A) = p$, $\mathcal{D}$ è anche un sistema indipendente massimale in $\mathcal{C}'$. Pertanto ogni colonna di A' dipende da $\mathcal{D}$. In particolare B dipende da $\mathcal{D}$ e quindi anche da $\mathcal{C}$ e, sempre in base al Teorema 4.5, il sistema lineare (4.1) risulta compatibile. □

Se il sistema lineare (4.1) è comparibile e si ha che $\rho(A) = \rho(A') = k$, si dice anche che il sistema lineare è *compatibile di rango k*.

4.2 Metodi di risoluzione

Analizzeremo ora due procedimenti per lo studio della compatibilità e la ricerca delle eventuali soluzioni di un sistema lineare. Il primo procedimento proposto è noto come il *metodo dei determinanti.*

In base al Teorema di Rouché-Capelli, per stabilire se il sistema lineare (4.1) è compatibile basta calcolare i ranghi di A ed A' e confrontarli. La ricerca delle eventuali soluzioni di (4.1) è invece più laboriosa. Diamo quindi una caratterizzazione per tali soluzioni.

4.7 TEOREMA. *Sia ξ una soluzione del sistema lineare* (4.1). *Allora le soluzioni di* (4.1) *sono tutti e soli i vettori numerici del tipo $\xi + \lambda$ dove $\lambda = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}$ è una soluzione del sistema lineare omogeneo associato* (4.2).

Dimostrazione. Sia $\mu = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix}$ una soluzione di (4.1). Si ha che $A\mu = B$. D'altra parte si ha anche che $A\xi = B$. Pertanto

$$A(\mu - \xi) = A\mu - A\xi = B - B = \mathbf{0}$$

e quindi, posto $\lambda = \mu - \xi$, si ha che λ è una soluzione del sistema lineare omogeneo associato e che $\mu = \xi + \lambda$.
Viceversa, se λ è una soluzione del sistema lineare omogeneo associato, si ha che $A\lambda = \mathbf{0}$ e quindi

$$A(\xi + \lambda) = A\xi + A\lambda = B + \mathbf{0} = B$$

e $\xi + \lambda$ è una soluzione di (4.1). □

In pratica quindi, la ricerca delle eventuali soluzioni di un sistema lineare si riduce alla determinazione di una soluzione di tale sistema e alla ricerca delle soluzioni del sistema lineare omogeneo ad esso associato. Ci occuperemo ora della risoluzione di un particolare tipo di sistema lineare.

4.8 DEFINIZIONE. *Un sistema lineare del tipo*

$$\begin{cases} a_{1,1}x_1 + \cdots + a_{1,n}x_n = b_1 \\ a_{2,1}x_1 + \cdots + a_{2,n}x_n = b_2 \\ \quad \vdots \qquad\qquad \vdots \quad\;\; \vdots \\ a_{n,1}x_1 + \cdots + a_{n,n}x_n = b_n \end{cases} \tag{4.6}$$

si dice quadrato di ordine n. Se la matrice incompleta A di tale sistema è non singolare, il sistema si dice di Cramer.

4.9 TEOREMA DI CRAMER. *Consideriamo un sistema lineare di Cramer* (4.6). *Tale sistema è compatibile e ammette un'unica soluzione*

$$\xi = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}$$

che è descritta come segue. Posto

$$F_j = (A^1, \dots, A^{j-1}, B, A^{j+1}, \dots, A^n)$$

si ha che

$$\xi_j = \frac{\det F_j}{\det A} \tag{4.7}$$

per ogni $j = 1, \dots, n$.

Dimostrazione. Poiché $\det A \neq 0$, si ha che $\rho(A) = n$ e quindi anche $\rho(A') = n$ e il sistema (4.6) risulta compatibile, per il Teorema di Rouché-Capelli. Proviamo ora che se ξ è una soluzione di (4.6) essa deve esprimersi attraverso le (4.7). Sia dunque ξ una soluzione di (4.6). Poiché $\det A \neq 0$, per il Teorema 3.34 la matrice A risulta invertibile. Pertanto

$$\xi = I_n \cdot \xi = A^{-1} \cdot A \cdot \xi = A^{-1} \cdot B\ . \tag{4.8}$$

Posto $A^{-1} = (c_{i,j})$ la (4.8) equivale alle relazioni scalari

$$\begin{aligned}
\xi_j &= (A^{-1})_j \cdot B \\
&= \sum_{k=1}^{n} c_{j,k} b_k \\
&= \sum_{k=1}^{n} (-1)^{j+k} \frac{\det A_{(k,j)}}{\det A} b_k \\
&= \frac{1}{\det A} \cdot \sum_{k=1}^{n} (-1)^{j+k} b_k \det A_{(k,j)} \\
&= \frac{\det F_j}{\det A} \qquad\qquad (\, j = 1, \dots, n\,)
\end{aligned}$$

□

La (4.7) è talvolta detta *Regola di Cramer.*

4.10 Corollario. *Un sistema lineare omogeneo quadrato*

$$AX = \mathbf{0}$$

ammette soluzioni non banali se e solo se $\det A = 0$.

Dimostrazione. Già sappiamo che un sistema lineare omogeneo è sempre compatibile ed ammette la soluzione banale. Se $\det A \neq 0$, tale soluzione è l'unica possibile, in base al Teorema di Cramer. Viceversa, se il sistema omogeneo in questione ammette solo la soluzione banale, ciò vuol dire che il vettore colonna nullo si esprime in modo unico (quello banale) come combinazione lineare delle colonne di A. Pertanto tali colonne sono indipendenti e si ha che $\det A \neq 0$. □

Torniamo ora al caso generale e consideriamo un sistema lineare del tipo (4.1).

4.11 Definizione. *Consideriamo un sistema lineare del tipo* (4.1). *Se*

$$\rho(A) = m \leq n$$

diremo che il sistema lineare è ridotto in forma normale.

4.12 Proposizione. *Un sistema lineare* (4.1) *ridotto in forma normale è sempre compatibile.*

Dimostrazione. Già sappiamo che $m = \rho(A) \leq \rho(A')$. D'altra parte il rango di A' non può superare il numero m di righe di A'. Quindi $\rho(A) = \rho(A') = m$ e il sistema lineare risulta compatibile per il Teorema di Rouché-Capelli. $\square$

4.13 Lemma. *Sia* $\mathcal{T}' = [A'_{i_1}, \dots, A'_{i_p}]$ *un sistema di righe della matrice* A' *e sia* A'_k *una riga di* A' *che dipende da* $\mathcal{T}'$. *Allora ogni soluzione del sistema lineare*

$$\begin{cases} A_{i_1} X = b_{i_1} \\ A_{i_2} X = b_{i_2} \\ \quad \vdots \qquad \vdots \\ A_{i_p} X = b_{i_p} \end{cases} \tag{4.9}$$

è anche soluzione dell'equazione

$$A_k X = b_k \ .$$

Dimostrazione. Sia ξ una soluzione del sistema lineare (4.9). Abbiamo che

$$A_{i_1}\xi = b_{i_1} \ ; \dots \ ; \ A_{i_p}\xi = b_{i_p} \ .$$

Poiché A'_k dipende da $\mathcal{T}'$, esistono degli scalari $\lambda_1, \dots, \lambda_p$ tali che

$$A'_k = \lambda_1 A'_{i_1} + \cdots + \lambda_p A'_{i_p}$$

ovvero

$$A_k = \lambda_1 A_{i_1} + \cdots + \lambda_p A_{i_p} \ ; \quad b_k = \lambda_1 b_{i_1} + \cdots + \lambda_p b_{i_p} \ .$$

Ma allora

$$\begin{aligned} A_k\xi &= (\lambda_1 A_{i_1} + \cdots + \lambda_p A_{i_p})\xi \\ &= \lambda_1 A_{i_1}\xi + \cdots + \lambda_p A_{i_p}\xi \\ &= \lambda_1 b_{i_1} + \cdots + \lambda_p b_{i_p} \\ &= b_k \end{aligned}$$

e questo è quanto si voleva provare. $\square$

4.14 Teorema. *Supponiamo che il sistema lineare* (4.1) *sia compatibile e che si abbia*

$$\rho(A) = \rho(A') = p \ .$$

Se $\mathcal{T} = [A_{i_1}, \ldots, A_{i_p}]$ *è un sistema indipendente massimale di righe di* A, *il sistema lineare* (4.1) *è equivalente al sistema lineare* (4.9), *che è quindi ancora compatibile ed è ridotto in forma normale.*

Dimostrazione. Il sistema di vettori $\mathcal{T}' = [A'_{i_1}, \ldots, A'_{i_p}]$ è un sistema indipendente massimale di righe di A'. Infatti $\mathcal{T}'$ è indipendente, poiché se esistesse una relazione di dipendenza tra i vettori di $\mathcal{T}'$, da essa si dedurrebbe una relazione di dipendenza tra i vettori di $\mathcal{T}$. $\mathcal{T}'$ è poi massimale rispetto a tale proprietà, in quanto $\rho(A') = p$ e quindi non possono esistere sistemi di righe indipendenti di A' di ordine maggiore di p. Si deduce allora che ogni riga di A' dipende da $\mathcal{T}'$ e, in base al Lemma 4.13, ogni soluzione di (4.9) è anche soluzione di ogni equazione del sistema lineare (4.1), e cioè è soluzione del sistema lineare (4.1) stesso. È poi evidente che ogni soluzione del sistema lineare (4.1) è anche soluzione del sistema lineare (4.9). Infine, è immediato verificare che il sistema lineare (4.9) è ridotto in forma normale, e ciò conclude la dimostrazione. □

Il Teorema 4.14 ci consente di concentrare la nostra attenzione sui sistemi lineari compatibili ridotti in forma normale. Infatti se il sistema lineare (4.1) è compatibile e $p = \rho(A) = \rho(A') < m$, da esso possiamo estrarre un sistema del tipo (4.9), dove $\mathcal{T} = [A_{i_1}, \ldots, A_{i_p}]$ è un sistema massimale indipendente di righe di A, che risulta compatibile, ridotto in forma normale ed equivalente a quello di partenza. Supponiamo quindi d'ora in avanti che il sistema lineare (4.1) sia compatibile e ridotto in forma normale, ovvero che si abbia

$$\rho(A) = \rho(A') = m \leq n \ .$$

Se $m = n$, tale sistema lineare è di Cramer. Infatti esso è quadrato ed inoltre $\det A \neq 0$ poiché $\rho(A) = m$. Possiamo quindi determinare la sua unica soluzione con la regola di Cramer. Se invece $m < n$, procediamo come segue. Sia $\mathcal{D} = [A^{j_1}, \ldots, A^{j_m}]$ un sistema massimale indipendente di colonne di A (e quindi anche di A'). Gli indici $j_1, \ldots, j_m$ sono allora a due a due distinti e $\{ j_1, \ldots, j_m \} \subset \{ 1, \ldots, n \}$. Sia $q = n - m$ e sia

$$\{ k_1, \ldots, k_q \} = \{ 1, \ldots, n \} - \{ j_1, \ldots, j_m \} \ .$$

Scriviamo il sistema (4.1) come segue

$$\begin{cases} a_{1,j_1} x_{j_1} + \cdots + a_{1,j_m} x_{j_m} = b_1 - a_{1,k_1} x_{k_1} - \cdots - a_{1,k_q} x_{k_q} \\ \quad \vdots \qquad\qquad \vdots \qquad\qquad \vdots \\ a_{m,j_1} x_{j_1} + \cdots + a_{m,j_m} x_{j_m} = b_m - a_{m,k_1} x_{k_1} - \cdots - a_{m,k_q} x_{k_q} \end{cases} \tag{4.10}$$

Se fissiamo arbitrariamente degli scalari $\xi_{k_1}, \dots, \xi_{k_q}$ e li sostituiamo al 2^o membro della (4.10) al posto delle incognite $x_{k_1}, \dots, x_{k_q}$ otteniamo il sistema lineare

$$\begin{cases} a_{1,j_1}x_{j_1}+\dots+a_{1,j_m}x_{j_m}=b_1-a_{1,k_1}\xi_{k_1}-\dots-a_{1,k_q}\xi_{k_q} \\ \quad\vdots \qquad\qquad\quad \vdots \qquad\qquad\quad \vdots \\ a_{m,j_1}x_{j_1}+\dots+a_{m,j_m}x_{j_m}=b_m-a_{m,k_1}\xi_{k_1}-\dots-a_{m,k_q}\xi_{k_q} \end{cases} \tag{4.11}$$

La matrice dei coefficienti del sistema lineare così ottenuto è la sottomatrice quadrata $A^{j_1,\dots,j_m}_{1,\dots,m}$, che è non singolare, in quanto le sue colonne sono indipendenti, e non dipende dalla scelta degli scalari $\xi_{k_1}, \dots, \xi_{k_q}$ che compaiono solo nella colonna dei termini noti. Quindi il sistema lineare (4.11) è di Cramer ed ammette come unica soluzione la m-pla $\begin{pmatrix} \xi_{j_1} \\ \vdots \\ \xi_{j_m} \end{pmatrix}$, che si determina con la regola di Cramer. Ma allora la n-pla $\begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}$ è una soluzione del sistema (4.1). Al variare di tutte le possibili scelte di $\xi_{k_1}, \dots, \xi_{k_q}$ si ottengono tutte e sole le soluzioni del sistema lineare (4.1). Infatti, se

$$\begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} \tag{4.12}$$

è una soluzione del sistema (4.1), sostituendo gli scalari $\lambda_{k_1}, \dots, \lambda_{k_q}$ alle incognite $x_{k_1}, \dots, x_{k_q}$ si ottiene un sistema di Cramer nelle incognite $x_{j_1}, \dots, x_{j_m}$ che dovrà necessariamente ammettere come unica soluzione la m-pla costituita dagli scalari $\lambda_{j_1}, \dots, \lambda_{j_m}$ e quindi anche la soluzione (4.12) si ottiene nel modo sopra descritto. Tale descrizione delle soluzioni di (4.1) può esprimersi anche dicendo che esiste una biezione

$$\omega : \mathbb{K}^q \longrightarrow \mathrm{Sol}\,(4.1) \subseteq \mathbb{K}^n \tag{4.13}$$

che associa all'arbitraria q-pla $(\xi_{k_1}, \dots, \xi_{k_q})$ la soluzione completa $(\xi_1, \dots, \xi_n)$ nel modo sopra indicato. Diremo che $\xi_{k_1}, \dots, \xi_{k_q}$ al loro variare in $\mathbb{K}$ parametrizzano Sol (4.1). Pertanto $\xi_{k_1}, \dots, \xi_{k_q}$ vengono detti talvolta *parametri*, e si dice che $\mathbb{K}^q$ funge da *spazio dei parametri*. La discussione appena svolta si sintetizza nel seguente enunciato:

4.15 2^o Teorema di unicità. *Consideriamo il sistema lineare* (4.1) *e supponiamo che esso sia già ridotto in forma normale e che si abbia* $m < n$. *Se* $\mathcal{S} = [A^{j_1}, \dots, A^{j_m}]$ *è un sistema massimale di colonne della prima (e quindi anche della seconda) matrice del sistema e* $k_1, \dots, k_q$ *sono come sopra, per ogni* q*-pla* $(\xi_{k_1}, \dots, \xi_{k_q}) \in \mathbb{K}^q$ *esiste un'unica soluzione* $\xi = (\xi_1, \dots, \xi_n) \in \mathbb{K}^n$ *del* *sistema, fornita dalla regola di Cramer applicata al sistema* (4.11) *alle incognite* $x_{j_1}, \dots, x_{j_m}$. *Inoltre, la* (4.13) *è una biezione.*

Consideriamo ora un sistema lineare omogeneo (4.2). Abbiamo già osservato che (4.2) è compatibile, poiché ammette sicuramente almeno la soluzione banale. Siamo quindi interessati alla ricerca di altre eventuali soluzioni (non banali) di (4.2). Supponiamo che il sistema lineare omogeneo (4.2) sia ridotto in forma normale, ovvero che si abbia $\rho(A) = m \leq n$. Se $m = n$, il sistema (4.2) è di Cramer e quindi ammette un'unica soluzione, e cioè quella banale.

Per semplicità, indichiamo, nel seguito, anche con il simbolo S l'insieme $\mathrm{Sol}\,(4.1)$ e poniamo anche $S_0 = \mathrm{Sol}\,(4.2)$.

4.16 Teorema. *L'insieme* S_0 *delle soluzioni del sistema lineare omogeneo* (4.2), *che supponiamo ridotto in forma normale, è un sottospazio vettoriale di* $\mathbb{K}^n$ *e si ha che* $\dim S_0 = n - m$.

Dimostrazione. Già sappiamo che $\mathbf{0} \in S_0$. Inoltre, se $\lambda \in \mathbb{K}$ e i vettori numerici colonna (di ordine n) ξ, ξ' sono in S_0, si ha che

$$A\xi = \mathbf{0} \ ; \quad A\xi' = \mathbf{0}$$

e quindi

$$A(\xi + \xi') = A\xi + A\xi' = \mathbf{0} \ ; \quad A(\lambda\xi) = \lambda(A\xi) = \mathbf{0}$$

e cioè $\xi + \xi'$ e $\lambda\xi$ sono ancora soluzioni di (4.2), ovvero elementi di S_0. Pertanto S_0 è un sottospazio di $\mathbb{K}^n$. Osserviamo poi che l'applicazione ω definita nella (4.13) per sistemi lineari arbitrari, risulta lineare quando (e solo quando) il sistema è omogeneo. Poiché poi ω è una biezione, ω è un isomorfismo tra $\mathbb{K}^q$ e S_0. Quindi $\dim S_0 = q = n - m$. □

In generale, se $S_0 \subseteq \mathbb{K}^n$ e $\xi \in \mathbb{K}^n$, l'insieme

$$S = \left\{ \xi' \in \mathbb{K}^n \mid \exists\, \xi'' \in S_0 \mid \xi' = \xi + \xi'' \right\}$$

si dice ξ-*traslato* di S_0 e si indica con il simbolo $\xi + S_0$. Se S_0 è un sottospazio vettoriale di $\mathbb{K}^n$, $\xi + S_0$ prende anche il nome di *laterale* di S_0 in $\mathbb{K}^n$ determinato da ξ. Per quanto detto in precedenza, se consideriamo un sistema lineare (4.1) ed il sistema lineare omogeneo (4.2) ad esso associato, ed indichiamo con S, S_0 i rispettivi insiemi di soluzioni e con ξ un qualunque elemento di S, abbiamo che $S = \xi + S_0$. Osserviamo che S_0 è un sottospazio vettoriale di $\mathbb{K}^n$, mentre S è un sottospazio vettoriale di $\mathbb{K}^n$ se e solo se il sistema (4.1) è omogeneo, ovvero se e solo se gli scalari $b_1, \dots, b_m$ sono tutti nulli, come è facile verificare notando, ad esempio, che $\mathbf{0} \in S$ se e solo se il sistema (4.1) è omogeneo. Un insieme S di questo tipo viene considerato un sottoinsieme notevole di $\mathbb{K}^n$ e prende il nome di *sottospazio affine* di $\mathbb{K}^n$, nel senso che verrà chiarito in un capitolo successivo. Anche $\mathbb{K}^n$ può vedersi come sottospazio affine (improprio) di se stesso, in quanto è l'insieme delle soluzioni del sistema lineare vuoto.

Studieremo ora un altro procedimento, che coinvolge l'uso delle operazioni elementari sulle righe di una matrice, per la risoluzione di un sistema lineare. Tale procedimento è noto come il *Metodo di Gauss-Jordan*.

4.17 Lemma. *Consideriamo il sistema lineare* (4.1) *e indichiamo con* A, A' *le matrici di tale sistema. Sia poi* P' *una matrice a scala equivalente ad* A'. *Indichiamo con* $P = (p_{i,j})$ *il blocco di* P' *costituito dalle prime* n *colonne, e con* $Q = \begin{pmatrix} q_1 \\ \vdots \\ q_m \end{pmatrix}$ *l'ultima colonna di* P'. *Il sistema lineare*

$$PX = Q \tag{4.14}$$

è equivalente al sistema (4.1).

Dimostrazione. È chiaro che occorre e basta provare l'asserto nel caso in cui P' si ottiene da A' mediante una sola operazione elementare. Se l'operazione è di tipo (i) oppure (ii), l'asserto è banale. Infatti se si effettua una operazione del tipo $R_i \rightsquigarrow \lambda R_i$ nel sistema lineare (4.14) compaiono le stesse equazioni del sistema lineare (4.1), tranne la i-ma, che risulta moltiplicata per λ, e questo non influenza l'insieme delle soluzioni. Se l'operazione è del tipo (ii), il sistema lineare (4.14) coincide con il sistema lineare (4.1), variando soltanto l'ordine in cui compaiono le equazioni. Supponiamo infine che P' si ottenga da A' mediante l'operazione $R_i \rightsquigarrow R_i+\lambda R_j$. Sia ad esempio $i < j$. I sistemi lineari (4.1) e (4.14) sono della forma

$$\begin{cases} A_1X = b_1 \\ \vdots \\ A_iX = b_i \\ \vdots \\ A_jX = b_j \\ \vdots \\ A_mX = b_m \end{cases} ; \qquad \begin{cases} A_1X = b_1 \\ \vdots \\ (A_i + \lambda A_j)X = b_i + \lambda b_j \\ \vdots \\ A_jX = b_j \\ \vdots \\ A_mX = b_m \end{cases}$$

Essi differiscono solo nella i-ma equazione. Se $Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ è una soluzione del primo sistema, allora $A_hY = b_h$ per ogni h e quindi

$$(A_i + \lambda A_j)Y = A_iY + \lambda A_jY = b_i + \lambda b_j \ .$$

Pertanto Y è anche soluzione del secondo sistema. Il viceversa è analogo. □

Il lemma precedente può essere usato per la risoluzione del sistema lineare (4.1). Infatti già sappiamo che esiste una matrice P' del tipo richiesto nel lemma. Pertanto possiamo studiare il sistema (4.14). La matrice P è anch'essa a scala. Abbiamo già osservato che $\rho(P)$ è il numero di righe non nulle di P e $\rho(P')$ è il numero di righe non nulle di P'. Pertanto il sistema lineare (4.14) sarà compatibile se e solo se l'ultima riga non banale di P è anche l'ultima riga non banale di P'. Il sistema lineare (4.14) sarà invece incompatibile se ciò non accade, ovvero se esiste un indice di riga h tale che

$$p_{h,1} = \cdots = p_{h,n} = 0 \; ; \quad q_h \neq 0 \; .$$

In tal caso infatti, la h-ma equazione sarà

$$0x_1 + \cdots + 0x_n = q_h \neq 0$$

e quindi è ovvio che (4.14) non ha soluzioni.

Se entrambe le matrici di un sistema lineare sono a scala, il sistema stesso si dirà a scala. Consideriamo ora un sistema a scala compatibile del tipo (4.14), trascurando le eventuali equazioni banali. Il sistema lineare (4.14) è del tipo

$$\begin{cases} p_{1,j_1}x_{j_1} + \cdots + p_{1,j_2}x_{j_2} + \cdots + p_{1,j_h}x_{j_h} + \cdots & = q_1 \\ \qquad\qquad p_{2,j_2}x_{j_2} + \cdots + p_{2,j_h}x_{j_h} + \cdots & = q_2 \\ \qquad\qquad\qquad\qquad \vdots & \;\;\vdots \\ \qquad\qquad\qquad\qquad p_{h,j_h}x_{j_h} + \cdots & = q_h \end{cases}$$

dove $p_{1,j_1}, \ldots, p_{h,j_h}$ sono i pivot. Poiché $p_{1,j_1}, \ldots, p_{h,j_h} \neq 0$, la sottomatrice $P^{j_1,\ldots,j_h}_{1,\ldots,h}$ è triangolare alta ed è non degenere, in quanto

$$\det P^{j_1,\ldots,j_h}_{1,\ldots,h} = p_{1,j_1} \cdot \cdots \cdot p_{h,j_h} \neq 0 \; .$$

Pertanto, usando il metodo dei determinanti, si ottengono le soluzioni di (4.14) scegliendo dei valori arbitrari per le incognite di posto diverso da $j_1, \ldots, j_h$ e risolvendo poi il sistema di Cramer nelle incognite $x_{j_1}, \ldots, x_{j_h}$ così ottenuto. Si opera, in pratica, come segue. L'ultima equazione è

$$p_{h,j_h}x_{j_h} + p_{h,j_h+1}x_{j_h+1} + \cdots + p_{h,n}x_n = q_h \; .$$

Poiché $p_{h,j_h} \neq 0$, si fissano degli scalari $\xi_{j_h+1}, \ldots \xi_n$ in modo arbitrario, si sostituiscono alle incognite $x_{j_h+1}, \ldots, x_n$ e si ottiene l'equazione

$$p_{h,j_h}x_{j_h} = q_h - p_{h,j_h+1}\xi_{j_h+1} - \cdots - p_{h,n}\xi_n$$

da cui si ricava un unico valore ξ_{j_h} da sostituire a x_{j_h}. Sostituendo nella penultima equazione si ottiene

$$\begin{aligned} p_{h-1,j_{h-1}}x_{j_{h-1}} &+ \cdots + p_{h-1,j_h-1}x_{j_h-1} \\ &= q_{h-1} - p_{h-1,j_h}\xi_{j_h} - p_{h-1,j_h+1}\xi_{j_h+1} - \cdots - p_{h-1,n}\xi_n \; . \end{aligned}$$

Si scelgono arbitrariamente degli scalari $\xi_{j_{h-1}+1}, \dots, \xi_{j_h-1}$ da sostituire alle incognite $x_{j_{h-1}+1}, \dots, x_{j_h-1}$ e si ricava un unico valore $\xi_{j_{h-1}}$ da sostiture all'incognita $x_{j_{h-1}}$. Si continua in modo analogo ad agire usando successivamente le equazioni di posto $h-2, h-3, \dots, 1$.

4.3 Alcuni esempi

4.18 Esempio. *Consideriamo il sistema lineare*

$$\begin{cases} x + y = 1 \\ x - y = 0 \end{cases} \tag{4.15}$$

di due equazioni nelle due incognite x, y. Le due matrici A, A' di tale sistema e il vettore colonna B dei termini noti sono

$$A = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \quad ; \quad B = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad ; \quad A' = (A|B) = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \end{pmatrix} .$$

Poiché $\det A = -2 \neq 0$, *si ha che $\rho(A) = \rho(A') = 2$. Pertanto il sistema* (4.15) *è di Cramer ed ammette un'unica soluzione $\xi = (\bar{x}, \bar{y})$, dove*

$$\bar{x} = \frac{\begin{vmatrix} 1 & 1 \\ 0 & -1 \end{vmatrix}}{-2} = \frac{1}{2} \quad ; \quad \bar{y} = \frac{\begin{vmatrix} 1 & 1 \\ 1 & 0 \end{vmatrix}}{-2} = \frac{1}{2} . \tag{4.16}$$

Risolviamo ora il sistema (4.15) *con il metodo di Gauss-Jordan. La matrice a scala P' che si ottiene da A' con l'algoritmo di Gauss è*

$$P' = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 1 \end{pmatrix} .$$

Pertanto il sistema (4.15) *è equivalente al seguente sistema a scala*

$$\begin{cases} x + y = 1 \\ \quad 2y = 1 \end{cases} \tag{4.17}$$

Dalla seconda delle (4.17) *si ricava $y = 1/2$. Sostituendo nella prima delle* (4.17) *si ottiene anche $x = 1/2$.*

4.19 Esempio. *Consideriamo il sistema lineare*

$$\begin{cases} x + y = 1 \\ x - y = 0 \\ x + 3y = 2 \end{cases} \tag{4.18}$$

di tre equazioni nelle due incognite x, y. Con le solite notazioni relative alle matrici del sistema, abbiamo che

$$A' = (A|B) = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & 3 & 2 \end{pmatrix} .$$

Poiché $\det A_{1,2}^{1,2} = -2 \neq 0$ e $\det A' = 0$, *si ha che $\rho(A) = \rho(A') = 2$. Pertanto il sistema* (4.18) *è compatibile (per il Teorema di Rouché-Capelli), ed è equivalente ad un suo sottosistema costituito da due equazioni. Poiché* $\det A_{1,2}^{1,2} \neq 0$, *possiamo selezionare le prime due equazioni e il sistema* (4.18) *è equivalente al sistema* (4.15). *Risolviamo ora il sistema* (4.18) *con il metodo di Gauss-Jordan. La matrice a scala P' che si ottiene da A' con l'algoritmo di Gauss è*

$$P' = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 0 \end{pmatrix} .$$

Pertanto il sistema (4.18) *è equivalente al sistema a scala* (4.17).

4.20 Esempio. *Consideriamo il sistema lineare*

$$\begin{cases} x & + & y & + & z & = & 1 \\ x & - & y & & & = & 0 \\ 2x & + & & & z & = & 2 \end{cases} \tag{4.19}$$

di tre equazioni nelle tre incognite x, y, z. *Con le solite notazioni relative alle matrici del sistema, abbiamo che*

$$A' = (A|B) = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 0 & 0 \\ 2 & 0 & 1 & 1 \end{pmatrix} .$$

Poiché $\det A^{1,3}_{1,2} = -1 \neq 0$ *e non ci sono in* A *e in* A' *minori non nulli di ordine 3, si ha che* $\rho(A) = \rho(A') = 2$. *Pertanto il sistema* (4.19) *è compatibile (per il Teorema di Rouché-Capelli), ed è equivalente ad un suo sottosistema costituito da due equazioni. Poiché* $\det A^{1,3}_{1,2} \neq 0$, *possiamo selezionare le prime due equazioni e il sistema* (4.19) *è equivalente al seguente sistema*

$$\begin{cases} x & + & y & + & z & = & 1 \\ x & - & y & & & = & 0 \end{cases} \tag{4.20}$$

Inoltre, avendo selezionato $A^{1,3}_{1,2}$ *come sottomatrice fondamentale, seguendo il procedimento indicato dal* 2^o *Teorema di unicità possiamo riscrivere il sistema* (4.20) *portando al secondo membro i termini in* y. *Otteniamo*

$$\begin{cases} x + z = 1 - y \\ x \qquad = y \end{cases} \tag{4.21}$$

Assegnamo valori arbitrari $\xi \in \mathbb{R}$ *ad* y *e risolviamo in* x, z *con la regola di Cramer, ovvero risolviamo, per ogni* $\xi \in \mathbb{R}$ *il sistema di Cramer*

$$\begin{cases} x + z = 1 - \xi \\ x \qquad = \xi \end{cases}$$

Otteniamo

$$x = \frac{\begin{vmatrix} 1-\xi & 1 \\ \xi & 0 \end{vmatrix}}{\begin{vmatrix} 1 & 1 \\ 1 & 0 \end{vmatrix}} = \xi \quad ; \quad y = \frac{\begin{vmatrix} 1 & 1-\xi \\ 1 & \xi \end{vmatrix}}{\begin{vmatrix} 1 & 1 \\ 1 & 0 \end{vmatrix}} = 1 - 2\xi \ .$$

Pertanto l'insieme S *delle soluzioni del sistema* (4.19) *è*

$$\mathrm{Sol}\,(4.19) = \left\{ (\xi, \xi, 1 - 2\xi) \mid \xi \in \mathbb{R} \right\} \subseteq \mathbb{R}^3 \ .$$

Si dice che Sol (4.19) *è descritto al variare del* parametro *reale* ξ. *Vediamo cosa succede se scegliamo come sottomatrice fondamentale di* A *(e anche di* A'*) la sottomatrice* $A^{1,2}_{1,2}$. *Si selezionano ancora le prime due equazioni del sistema* (4.19), *ma l'incognita usata come parametro è questa volta* z. *Pertanto per ogni* $\lambda \in \mathbb{R}$ *si pone* $z = \lambda$ *e si risolve, con la regola di Cramer, il sistema*

$$\begin{cases} x + y = 1 - \lambda \\ x - y = 0 \end{cases} \tag{4.22}$$

Si ottiene

$$x = \frac{\begin{vmatrix} 1-\lambda & 1 \\ 0 & -1 \end{vmatrix}}{\begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix}} = \frac{1-\lambda}{2} \quad ; \quad y = \frac{\begin{vmatrix} 1 & 1-\lambda \\ 1 & 0 \end{vmatrix}}{\begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix}} = \frac{1-\lambda}{2} \ .$$

Naturalmente si può anche dedurre direttamente dalla seconda delle (4.22) *che* $x = y$, *sostituire nella prima delle* (4.22) *e trovare che* $x = y = (1 - \lambda)/2$. *L'insieme delle soluzioni* S' *è pertanto*

$$S' = \left\{ ((1-\lambda)/2, (1-\lambda)/2, \lambda) \mid \lambda \in \mathbb{R} \right\} \subseteq \mathbb{R}^3 \ .$$

Verifichiamo che $S = S'$. *Fissato un elemento* $(\bar{\xi}, \bar{\xi}, 1 - 2\bar{\xi}) \in S$ *(relativo al valore* $\bar{\xi}$ *del parametro* ξ, *osserviamo che tale terna è anche un elemento di* S' *(relativo al valore* $\bar{\lambda} = 1 - 2\bar{\xi}$ *del parametro* λ), *e quindi* $S \subseteq S'$. *Analogamente si vede che* $S' \subseteq S$. *Risolviamo ora il sistema* (4.19) *con il metodo di Gauss-Jordan. La matrice a scala* P' *che si ottiene da* A' *con l'algoritmo di Gauss è*

$$P' = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 2 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} .$$

Pertanto il sistema (4.19) *è equivalente al seguente sistema a scala.*

$$\begin{cases} x + y + z = 1 \\ \quad 2y + z = 1 \end{cases}$$

Posto $y = \xi$, *dalla seconda delle* (4.20) *si ricava che* $z = 1 - 2\xi$, *e poi sostituendo nella prima si trova che* $x = \xi$.

4.21 Esempio. *Consideriamo il sistema lineare*

$$\begin{cases} x_1 & + & x_2 & + & x_3 & + & x_4 & = & 1 \\ 2x_1 & + & & & & & x_4 & = & 1 \\ x_1 & - & x_2 & - & x_3 & & & = & 0 \\ 3x_1 & + & 2x_2 & + & 2x_3 & + & 2x_4 & = & 2 \end{cases} \tag{4.23}$$

di quattro equazioni nelle quattro incognite x_1, x_2, x_3, x_4. *Con le solite notazioni relative alle matrici del sistema, abbiamo che*

$$A' = (A|B) = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 2 & 0 & 0 & 1 & 1 \\ 1 & -1 & -1 & 0 & 0 \\ 3 & 2 & 2 & 2 & 2 \end{pmatrix} .$$

Si ha che $\det A^{1,3,4}_{1,2,4} = 1 \neq 0$ *e* $A^{1,3,4}_{1,2,4}$ *è una sottomatrice fondamentale di* A *e di* A'. *Pertanto* $\rho(A) = \rho(A') = 3$ *e il sistema* (4.23) *risulta compatibile (per il Teorema di Rouché-Capelli) ed equivalente al suo sottosistema costituito dalla prima, seconda e quarta equazioni. Seguendo il metodo del* 2^o *Teorema di unicità possiamo usare l'incognita* x_2 *come parametro, ponendo ad esempio* $x_2 = \xi$, *e risolvere, per ogni* $\xi \in \mathbb{R}$, *il sistema di Cramer*

$$\begin{cases} x_1 & + & x_3 & + & x_4 & = & 1 - \xi \\ 2x_1 & + & & & x_4 & = & 1 \\ 3x_1 & + & 2x_3 & + & 2x_4 & = & 2 - 2\xi \end{cases}$$

Con la regola di Cramer otteniamo

$$x_1 = \frac{\begin{vmatrix} 1-\xi & 1 & 1 \\ 1 & 0 & 1 \\ 2-2\xi & 2 & 2 \end{vmatrix}}{\det A^{1,3,4}_{1,2,4}} = 0 \quad ; \quad x_3 = \frac{\begin{vmatrix} 1 & 1-\xi & 1 \\ 2 & 1 & 1 \\ 3 & 2-2\xi & 2 \end{vmatrix}}{\det A^{1,3,4}_{1,2,4}} = -\xi$$

$$x_4 = \frac{\begin{vmatrix} 1 & 1 & 1-\xi \\ 2 & 0 & 1 \\ 3 & 2 & 2-2\xi \end{vmatrix}}{\det A^{1,3,4}_{1,2,4}} = 1 .$$

L'insieme delle soluzioni S *del sistema (31), descritto parametricamente, è*

$$S = \{\, (0, \xi, -\xi, 1) \mid \xi \in \mathbb{R} \,\} .$$

Risolviamo ora il sistema (4.23) *con il metodo di Gauss-Jordan. La matrice a scala* P' *che si ottiene da* A' *con l'algoritmo di Gauss è*

$$P' = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 2 & 2 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} .$$

Pertanto il sistema (4.23) è equivalente al seguente sistema a scala.

$$\begin{cases} x_1 + x_2 + x_3 + x_4 = 1 \\ 2x_2 + 2x_3 + x_4 = 1 \\ x_4 = 1 \end{cases} \tag{4.24}$$

Dalla terza delle (4.24) si ricava che $x_4 = 1$. Sostituendo nella seconda delle (4.24) tale valore, ponendo $x_3 = \xi$ si ottiene $x_2 = -\xi$. Sostituendo nella prima delle (4.24) si trova che $x_1 = 0$.

4.22 Esempio. *Sia k un arbitrario numero reale e consideriamo il sistema*

$$\begin{cases} kx + y = 1 \\ x - y = 0 \end{cases} \tag{4.25}$$

Se si studia il sistema al variare di $k \in \mathbb{R}$, si dice che tale sistema è parametrico e k si dice parametro (reale). Con le solite notazioni relative alle matrici associate al sistema, abbiamo che

$$A' = (A|B) = \begin{pmatrix} k & 1 & 1 \\ 1 & -1 & 0 \end{pmatrix} .$$

Si ha che $\rho(A') = 2$ per ogni $k \in \mathbb{R}$ e una sua sottomatrice fondamentale è data da ${A'}^{2,3}_{1,2}$, poiché det ${A'}^{2,3}_{1,2} = 1 \neq 0$. *Il rango di A dipende invece dalla scelta di k. Infatti* $\det A = -k-1$ *e quindi $\rho(A) = 1$ se $k = -1$ e $\rho(A) = 2$ se $k \neq -1$. Pertanto, in base al Teorema di Rouché-Capelli, il sistema (4.25) è incompatibile se $k = -1$. Se invece $k \neq -1$, il sistema è compatibile (ed è di Cramer) e la sua unica soluzione è data da*

$$x = -\frac{\begin{vmatrix} 1 & 1 \\ 0 & -1 \end{vmatrix}}{k+1} = \frac{1}{k+1} \quad ; \quad y = -\frac{\begin{vmatrix} k & 1 \\ 1 & 0 \end{vmatrix}}{k+1} = \frac{1}{k+1} .$$

Risolviamo ora il sistema (4.25) usando il procedimento di Gauss-Jordan. Applicando l'operazione elementare $R_2 \rightsquigarrow R_2 - \frac{1}{k}R_1$ alla matrice A' si ottiene la matrice

$$P' = \begin{pmatrix} k & 1 & 1 \\ 0 & -\frac{1+k}{k} & -\frac{1}{k} \end{pmatrix} .$$

Tale operazione ha senso solo se $k \neq 0$, quindi il caso $k = 0$ va trattato a parte. Per $k \neq 0$ il sistema (4.25) è equivalente al seguente sistema a scala

$$\begin{cases} kx + y = 1 \\ \dfrac{k+1}{k} y = \dfrac{1}{k} \end{cases} \tag{4.26}$$

Moltiplicando la seconda equazione del sistema (4.25) per k si ottiene il sistema (ancora equivalente) a scala

$$\begin{cases} kx + y = 1 \\ (k+1)y = 1 \end{cases} \tag{4.27}$$

il quale è compatibile se e solo se $k \neq -1$. In tal caso dalla seconda delle (4.27) si ricava che $y = \frac{1}{k+1}$. Sostituendo tale valore della y nella prima delle (4.27) si ricava che $x = \frac{1}{k+1}$. Rimane da esaminare il caso $k = 0$, che è peraltro banale.

4.23 Esempio. *Consideriamo il seguente sistema*

$$\begin{cases} hx + ky = -1 \\ x - ky = h \end{cases} \tag{4.28}$$

Il sistema (4.28) è un sistema di 2 equazioni nelle 2 incognite x, y, con due parametri reali h, k. Con le solite notazioni relative alle matrici del sistema, abbiamo che

$$A' = (A|B) = \begin{pmatrix} h & k & -1 \\ 1 & -k & h \end{pmatrix} .$$

Studiamo i ranghi di A, A' al variare di h, k. Abbiamo che $\det A = -k(h+1)$. Pertanto $\rho(A) = 1$ se $k = 0$ oppure $h = -1$, $\rho(A) = 2$ se $k \neq 0$ e $h \neq -1$. Osserviamo che per ogni h, k la sottomatrice ${A'}^{1,3}_{1,2}$ è fondamentale, in quanto $\det {A'}^{1,3}_{1,2} = h^2 + 1 \neq 0$. Pertanto $\rho(A') = 2$ per ogni valore dei parametri e, in base al Teorema di Rouché-Capelli, il sistema (4.28) risulta compatibile (e di Cramer) se e solo se $k \neq 0$ e $h \neq -1$. In tal caso la soluzione, fornita dalla regola di Cramer, è

$$x = \frac{\begin{vmatrix} -1 & k \\ h & -k \end{vmatrix}}{\begin{vmatrix} h & k \\ 1 & -k \end{vmatrix}} = \frac{h-1}{h+1} \quad ; \quad y = \frac{\begin{vmatrix} h & -1 \\ 1 & h \end{vmatrix}}{\begin{vmatrix} h & k \\ 1 & -k \end{vmatrix}} = -\frac{h^2+1}{k(h+1)} \ .$$

Risolviamo ora il sistema (4.28) con il metodo di Gauss-Jordan. Applicando l'operazione elementare $R_2 \rightsquigarrow R_2 - \frac{1}{h} R_1$ alla matrice A' si ottiene la matrice a scala

$$P' = \begin{pmatrix} h & k & -1 \\ 0 & \frac{-kh+1}{h} & \frac{h^2+1}{h} \end{pmatrix} \ .$$

Affinché tale operazione abbia senso, si deve supporre $h \neq 0$. Pertanto il caso $h = 0$ va trattato a parte (ed è peraltro banale). Quindi, per $h \neq 0$ il sistema (4.28) è equivalente al seguente sistema a scala

$$\begin{cases} hx + ky = -1 \\ -\dfrac{k(h+1)}{h} y = \dfrac{h^2+1}{h} \end{cases} \tag{4.29}$$

o anche, moltiplicando la seconda delle (4.29) per h, al sistema

$$\begin{cases} hx + ky = -1 \\ -k(h+1)y = h^2 + 1 \end{cases} \tag{4.30}$$

Osserviamo che $h^2+1 \neq 0$ per ogni valore dei parametri, e quindi il sistema (4.30) è compatibile se e solo se $-k(h+1) \neq 0$, ovvero $k \neq 0$ e $h \neq -1$. In tal caso dalla seconda delle (4.30) si ricava che $y = -\frac{h^2+1}{k(h+1)}$. Sostituendo poi tale valore nella prima delle (4.30) si trova anche che $x = \frac{h-1}{h+1}$.

4.24 Esempio. *Consideriamo il seguente sistema sul campo $\mathbb{C}$.*

$$\begin{cases} x - ky = 1 \\ kx + y = 0 \end{cases} \tag{4.31}$$

Con le solite notazioni si ha che

$$A' = (A|B) = \begin{pmatrix} 1 & -k & 1 \\ k & 1 & 0 \end{pmatrix} \ .$$

Quindi $\det A = 1 + k^2 = 0$ se e solo se $k = \pm i$; d'altra parte

$$\det {A'}^{2,3}_{1,2} = \begin{vmatrix} -k & 1 \\ 1 & 0 \end{vmatrix} = -1 \neq 0 \quad \forall\, k \in \mathbb{C} \ .$$

Pertanto $\rho(A) = 2$ se $k \neq \pm i$ e $\rho(A) = 1$ se $k = \pm i$, mentre $\rho(A') = 2$ per ogni $k \in \mathbb{C}$. Il sistema (4.31) è dunque incompatibile per $k = \pm i$. Negli altri casi è compatibile, di Cramer, e la soluzione è

$$x = \frac{\begin{vmatrix} 1 & -k \\ 0 & 1 \end{vmatrix}}{k^2+1} = \frac{1}{k^2+1} \quad ; \quad y = \frac{\begin{vmatrix} 1 & 1 \\ k & 0 \end{vmatrix}}{k^2+1} = -\frac{k}{k^2+1} \ . \tag{4.32}$$

Naturalmente lo stesso sistema studiato sul campo $\mathbb{R}$ risulta sempre compatibile e la soluzione è ancora data dalle (4.32).

4.25 Esempio. *Consideriamo il seguente sistema sul campo $\mathbb{Q}$.*

$$\begin{cases} 2x + ky = |k| \\ kx + y = 1 \end{cases} \tag{4.33}$$

Con le solite notazioni si ha che

$$A' = (A|B) = \begin{pmatrix} 2 & k & |k| \\ k & 1 & 1 \end{pmatrix} .$$

Quindi $\det A = 2 - k^2 \neq 0$ *per ogni* $k \in \mathbb{Q}$ *e si ha che* $\rho(A) = \rho(A') = 2$ *e il sistema (4.33) è compatibile, di Cramer, per ogni* $k \in \mathbb{Q}$*. La soluzione è quindi*

$$x = \frac{\begin{vmatrix} |k| & k \\ 1 & 1 \end{vmatrix}}{2 - k^2} \quad ; \quad y = \frac{\begin{vmatrix} 2 & |k| \\ k & 1 \end{vmatrix}}{2 - k^2} .$$

Lo stesso sistema, studiato sul campo $\mathbb{R}$ *richiede una discussione più articolata. Infatti*

$$\det A = 2 - k^2 = 0 \quad \Longleftrightarrow \quad k = \pm\sqrt{2}$$

e quindi $\rho(A) = 1$ *se* $k = \pm\sqrt{2}$, $\rho(A) = 2$ *se* $k \neq \pm\sqrt{2}$*. Pertanto per* $k \neq \pm\sqrt{2}$ *si ha che che* $\rho(A') = 2$ *e il sistema (4.33) è compatibile, di Cramer, e la sua soluzione è come sopra. Per* $k = \pm\sqrt{2}$ *bisogna studiare il rango di* A'*. Nel caso* $k = -\sqrt{2}$ *si ha che* $\rho(A') = 2$ *e il sistema (4.33) è incompatibile. Nel caso* $k = \sqrt{2}$ *si ha che* $\rho(A') = 1$ *e il sistema (4.33) è compatibile ed equivalente ad un suo sottosistema costituito da una sola equazione, ad esempio la seconda, e si risolve con il metodo del* 2^o *Teorema di unicità ponendo* $x = \xi$ *e ricavando quindi che* $y = 1 - \sqrt{2}\xi$*. L'insieme delle soluzioni* S *è quindi descritto, al variare del parametro* ξ *come segue*

$$S = \left\{ (\xi, 1 - \sqrt{2}\xi) \mid \xi \in \mathbb{R} \right\} .$$

4.4 Rappresentazione di sottospazi vettoriali

Sia V uno spazio vettoriale di dimensione n sul campo $\mathbb{K}$, $\mathcal{B} = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$ una sua base ordinata, e $W \leq V$ un sottospazio vettoriale di dimensione h. Consideriamo inoltre un sistema lineare omogeneo del tipo (4.2), espresso quindi in forma matriciale da $AX = \mathbf{0}$, dove A è una matrice di tipo $m \times n$, e supponiamo che sia $\rho(A) = n - h$. Indichiamo con $\mathrm{Sol}(*)$ il corrispondente insieme di soluzioni. Come è noto, esso è un sottospazio vettoriale di $\mathbb{K}^n$, ha dimensione h ed i suoi elementi (cioè le soluzioni) sono descritti al variare di h parametri. Diremo che il sistema (4.2) è una rappresentazione (cartesiana) di W se accade che

$$\mathbf{u} \in W \quad \Longleftrightarrow \quad \phi_{\mathcal{B}}(\mathbf{u}) \in \mathrm{Sol}(*) ,$$

ovvero, $\mathbf{u} \in W$ se e solo se il vettore numerico delle sue componenti è una soluzione di $(*)$. In altri termini, l'isomorfismo coordinato

$$\phi_{\mathcal{B}} : V \longrightarrow \mathbb{K}^n$$

si restringe ad un isomorfismo

$$\bar{\phi}_{\mathcal{B}} : W \longrightarrow \mathrm{Sol}(*) .$$

4.26 Proposizione. *Ogni sistema omogeneo del tipo (4.2), con* $\rho(A) = n - h$*, rappresenta un sottospazio* $W \leq V$ *di dimensione* h*.*

Dimostrazione. Basta porre $W = \left(\phi_{\mathcal{B}}\right)^{-1}\left(\mathrm{Sol}(*)\right)$. □

Vale anche, in un certo senso, il viceversa: ogni sottospazio possiede (infinite) rappresentazioni cartesiane. Mostreremo ora come si costruisce una rappresentazione cartesiana. Supponiamo che il sottospazio W abbia una base $\mathcal{B}' = (\mathbf{u}_1, \ldots, \mathbf{u}_h)$. Un vettore $\mathbf{v} \in V$ appartiene a W se e solo se esso dipende da $\mathcal{B}'$, ovvero

$$\mathbf{v} \in W \iff \exists\, t_1, \ldots, t_h \in \mathbb{K} \mid \mathbf{v} = t_1\mathbf{u}_1 + \cdots + t_h\mathbf{u}_h \ .$$

Passando alle componenti, posto

$$\phi_{\mathcal{B}}(\mathbf{u}_1) = \begin{pmatrix} u_{1,1} \\ \vdots \\ u_{n,1} \end{pmatrix} \ ; \ \ldots \ ; \ \phi_{\mathcal{B}}(\mathbf{u}_h) = \begin{pmatrix} u_{1,h} \\ \vdots \\ u_{n,h} \end{pmatrix} \ ; \ \phi_{\mathcal{B}}(\mathbf{v}) = X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \ ,$$

ciò equivale a dire che

$$\mathbf{v} \in W \iff \exists\, t_1, \ldots, t_h \in \mathbb{K} \mid X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = t_1 \begin{pmatrix} u_{1,h} \\ \vdots \\ u_{n,h} \end{pmatrix} + \cdots + t_h \begin{pmatrix} u_{1,h} \\ \vdots \\ u_{n,h} \end{pmatrix} \ ,$$

o anche

$$\begin{cases} x_1 = t_1u_{1,1} + \cdots + t_hu_{1,h} \\ \qquad\vdots \\ x_n = t_1u_{n,1} + \cdots + t_hu_{n,h} \end{cases}$$

L'espressione

$$W : \begin{cases} x_1 = t_1u_{1,1} + \cdots + t_hu_{1,h} \\ \qquad\vdots \\ x_n = t_1u_{n,1} + \cdots + t_hu_{n,h} \end{cases}$$

si dice rappresentazione *parametrica* di W, con parametri $t_1, \ldots, t_h$.

Torniamo ora al punto in cui si richiede che $\mathbf{v}$ dipenda da $\mathcal{B}'$. Poiché $\mathcal{B}'$ è indipendente, ciò equivale a dire che il sistema $\mathcal{S} = [\mathbf{u}_1, \ldots, \mathbf{u}_h, \mathbf{v}]$ di vettori di V risulti dipendente, ovvero che sia dipendente il corrispondente sistema $\bar{\mathcal{S}}$ dei vettori numerici delle componenti dei vettori di $\mathcal{S}$:

$$\bar{\mathcal{S}} = \phi_{\mathcal{B}}(\mathcal{S}) = \left[\begin{pmatrix} u_{1,1} \\ \vdots \\ u_{n,1} \end{pmatrix}, \ldots, \begin{pmatrix} u_{1,h} \\ \vdots \\ u_{n,h} \end{pmatrix}, \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \right] .$$

In altri termini, posto

$$A = \begin{pmatrix} u_{1,1} & \ldots & u_{1,h} & x_1 \\ \vdots & & \vdots & \vdots \\ u_{n,1} & \ldots & u_{n,h} & x_n \end{pmatrix} ,$$

abbiamo che $\mathbf{v} \in W$ se e solo se $\rho(A) = h$. Poiché le prime h colonne di A sono indipendenti, possiamo determinare una sottomatrice individuata da tali colonne e da h opportune righe, con determinante non nullo. Supponiamo, per semplicità, che siano indipendenti proprio le prime h righe, ovvero che si abbia

$$\det\left(A^{1,\dots,h}_{1,\dots,h}\right) \neq 0 \ .$$

Allora la condizione $\rho(A) = h$ equivale, in base al Teorema degli Orlati, a richiedere che la sottomatrice $A^{1,\dots,h}_{1,\dots,h}$ sia fondamentale, ovvero che i suoi $n-h$ orlati siano degeneri. Tali orlati si ottengono da $A^{1,\dots,h}_{1,\dots,h}$ considerando in aggiunta una delle $n-h$ righe trascurate in $A^{1,\dots,h}_{1,\dots,h}$ e l'ultima colonna. L'annullarsi dei determinanti di tali orlati descrive il seguente sistema omogeneo di $n-h$ equazioni in n incognite $x_1, \dots, x_n$, che è la rappresentazione cercata:

$$W : \begin{cases} \det\left(A^{1,\dots,h,h+1}_{1,\dots,h,h+1}\right) = 0 \\ \det\left(A^{1,\dots,h,h+1}_{1,\dots,h,h+2}\right) = 0 \\ \vdots \\ \det\left(A^{1,\dots,h,h+1}_{1,\dots,h,n}\right) = 0 \end{cases}$$

Osserviamo che questa è una rappresentazione cartesiana minimale di W, nel senso che $n-h$ è il numero minimo di equazioni necessarie per rappresentare un sottospazio di dimensione h di uno spazio di dimensione n. Illustriamo tale costruzione un po' astratta con alcuni esempi.

4.27 Esempio. (Sottospazio di uno spazio vettoriale astratto). *Nello spazio tridimensionale dei vettori liberi ordinari V, si consideri il sottospazio W generato dai vettori $\mathbf{u}_1, \mathbf{u}_2$ che, in una fissata base $\mathcal{B} = (\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$, hanno componenti*

$$\mathbf{u}_1 =_{\mathcal{B}} (1, 0, 1) \ ; \ \mathbf{u}_2 =_{\mathcal{B}} (1, -1, 2) \ .$$

Poiché il sistema $\mathcal{S} = [\mathbf{u}_1, \mathbf{u}_2]$ è indipendente, W ha dimensione 2 ed $\mathcal{S}$ è una base di W. Se

$$X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

è il vettore numerico delle componenti del generico vettore $\mathbf{v} \in V$, abbiamo che

$$\mathbf{v} \in W \iff \mathbf{v} \text{ dipende da } \mathcal{S} \iff [\mathbf{u}_1, \mathbf{u}_2, \mathbf{v}] \text{ è dipendente} \iff \rho \begin{pmatrix} 1 & 1 & x_1 \\ 0 & -1 & x_2 \\ 1 & 2 & x_3 \end{pmatrix} = 2 \ .$$

Scelta la sottomatrice evidenziata in rosso, l'unico orlato possibile è dato dall'intera matrice, e quindi una rappresentazione di W è

$$W : \det \begin{pmatrix} 1 & 1 & x_1 \\ 0 & -1 & x_2 \\ 1 & 2 & x_3 \end{pmatrix} = 0 \ .$$

Sviluppando tale determinante con al regola di Laplace rispetto all'ultima colonna, otteniamo

$$W : \det \begin{pmatrix} 0 & -1 \\ 1 & 2 \end{pmatrix} x_1 - \det \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} x_2 + \det \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix} x_3 = 0 \ ,$$

ovvero

$$W : x_1 - x_2 - x_3 = 0$$

una rappresentazione di W *mediante un'unica equazione omogenea in* x_1, x_2, x_3*. Ricordiamo che* W *è isomorfo al sottospazio* $\mathrm{Sol}(*)$ *delle soluzioni del sistema omogeneo (in questo caso costituito da un'unica equazione) che rappresenta* W*.*

4.28 Esempio. (Sottospazio di uno spazio vettoriale numerico). *Consideriamo lo spazio vettoriale numerico di dimensione 4* $V = \mathbb{R}^4$*. Sia* W *il sottospazio generato dai vettori*

$$\mathbf{u}_1 = (1, 2, 0, 1) \ ; \ \mathbf{u}_2 = (2, 1, 1, -1) \ .$$

Indicato con $\mathbf{v} = (x_1, x_2, x_3, x_4)$ *il vettore generico, abbiamo che* $\mathbf{v} \in W$ *se e solo se*

$$\rho \begin{pmatrix} 1 & 2 & x_1 \\ \color{red}{2} & \color{red}{1} & x_2 \\ \color{red}{0} & \color{red}{1} & x_3 \\ 1 & -1 & x_4 \end{pmatrix} = 2 \ .$$

Utilizzando, ad esempio, la sottomatrice non degenere evidenziata in rosso, tale condizione sul rango equivale a dire che i suoi orlati

$$\begin{pmatrix} 1 & 2 & x_1 \\ \color{red}{2} & \color{red}{1} & x_2 \\ \color{red}{0} & \color{red}{1} & x_3 \end{pmatrix} \ ; \ \begin{pmatrix} \color{red}{2} & \color{red}{1} & x_2 \\ \color{red}{0} & \color{red}{1} & x_3 \\ 1 & -1 & x_4 \end{pmatrix}$$

sono degeneri, ovvero, sviluppando rispetto alla terza colonna,

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

Effettuando i calcoli abbiamo quindi la rappresentazione richiesta:

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

Ricordiamo che W*, in generale, è isomorfo al sottospazio* $\mathrm{Sol}(*)$ *delle soluzioni del sistema omogeneo che rappresenta* W*. Ma in questo caso (*$V = \mathbb{R}^4$*, uno spazio vettoriale numerico) si ha che* $W = \mathrm{Sol}(*)$*.*

4.29 Esempio. (Sottospazio corrispondente ad un sistema dato). *Consideriamo il sistema*

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

Si tratta di un sistema omogeneo di 3 equazioni in 4 incognite. Ma il rango della matrice associata è 2, e quindi il sistema si riduce ad un sistema di 2 equazioni in 4 incognite, e rappresenta un sottospazio isomorfo a quello dell'esempio precedente.

Capitolo 5

Endomorfismi

5.1 Matrici e applicazioni lineari

Consideriamo due spazi vettoriali V, V' di dimensione n, m rispettivamente, e sia $f : V \to V'$ un'applicazione lineare. Fissiamo inoltre delle basi ordinate

$$\mathcal{B} = (\mathbf{u}_1, \ldots, \mathbf{u}_n) \ ; \quad \mathcal{B}' = (\mathbf{u}'_1, \ldots, \mathbf{u}'_m)$$

degli spazi vettoriali V, V' rispettivamente. È possibile associare ad f una matrice $A \in \mathcal{M}_{m,n}$ rispetto alle basi $\mathcal{B}, \mathcal{B}'$ al modo seguente. La colonna j-ma della matrice A sarà il vettore numerico colonna (di ordine m)

$$A^j = \begin{pmatrix} a_{1,j} \\ \vdots \\ a_{m,j} \end{pmatrix}$$

delle componenti del vettore $f(\mathbf{u}_j)$ nella base $\mathcal{B}'$. In altre parole, essendo $\mathcal{B}'$ una base di V', per ogni j sono univocamente determinati degli scalari $a_{1,j}, \ldots, a_{m,j}$ tali che

$$f(\mathbf{u}_j) = a_{1,j}\mathbf{u}'_1 + \cdots + a_{m,j}\mathbf{u}'_m \ .$$

Tali scalari formano la colonna j-ma della matrice A. Sia ora $\mathbf{v} \in V$ e siano

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \ ; \quad Y = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}$$

i vettori numerici colonna delle componenti di $\mathbf{v}$ e di $f(\mathbf{v})$ rispetto alle basi $\mathcal{B}, \mathcal{B}'$. Si ha che

$$Y = AX \ .$$

Infatti, poiché

$$\mathbf{v} = \sum_{j=1}^{n} x_j \mathbf{u}_j$$

si ha che

$$\begin{aligned} f(\mathbf{v}) &= f\Big(\sum_{j=1}^{n} x_j \mathbf{u}_j\Big) \\ &= \sum_{j=1}^{n} x_j f(\mathbf{u}_j) \\ &= \sum_{j=1}^{n} x_j \Big(\sum_{i=1}^{m} a_{i,j} \mathbf{u}'_i\Big) \\ &= \sum_{i=1}^{m} \Big(\sum_{j=1}^{n} a_{i,j} x_j\Big) \mathbf{u}'_i \ . \end{aligned}$$

D'altra parte abbiamo anche che

$$f(\mathbf{v}) = \sum_{i=1}^{m} y_i \mathbf{u}'_i$$

e quindi, poiché ogni vettore si esprime in modo unico come combinazione lineare dei vettori di una base, deduciamo che

$$y_i = \sum_{j=1}^{n} a_{i,j} x_j$$

e cioè $Y = AX$. Viceversa, se A è una matrice tale che per ogni vettore $\mathbf{v} \in V$, detti X, Y i vettori numerici colonna delle componenti di $\mathbf{v}$ e di $f(\mathbf{v})$ in $\mathcal{B}, \mathcal{B}'$, si ha che $Y = AX$, allora A è la matrice associata ad f rispetto alle basi $\mathcal{B}, \mathcal{B}'$. Infatti, per ogni $j = 1, \ldots, n$ il vettore $\mathbf{u}_j$ ha come vettore numerico colonna delle componenti in $\mathcal{B}$ il vettore

$$\begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}$$

dove lo scalare 1 compare nella j-ma posizione. Pertanto il vettore numerico colonna delle componenti di $f(\mathbf{u}_j)$ sarà

$$A \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} = A^j \ .$$

Se indichiamo con Φ, Ψ gli isomorfismi coordinati di V, V' rispetto alle basi $\mathcal{B}, \mathcal{B}'$, quanto detto finora si riassume dicendo che A è la matrice associata ad f

rispetto alle basi $\mathcal{B}, \mathcal{B}'$ se e solo se

$$\Psi\big(f(\mathbf{v})\big) = A\big(\Phi(\mathbf{v})\big) \qquad \forall\, \mathbf{v} \in V\ .$$

Talvolta l'espressione $Y = AX$ è detta *rappresentazione* di f, e può anche scriversi come segue:

$$\begin{cases} y_1 = a_{1,1}x_1 + \cdots + a_{1,n}x_n \\ y_2 = a_{2,1}x_1 + \cdots + a_{2,n}x_n \\ \ \vdots \qquad \vdots \qquad\qquad\quad \vdots \\ y_m = a_{m,1}x_1 + \cdots + a_{m,n}x_n \end{cases}$$

5.1 ESEMPIO. *Sia $f : V \to V'$ l'applicazione nulla, definita ponendo $f(\mathbf{u}) = \mathbf{0}$ per ogni $\mathbf{u} \in V$. Fissati comunque dei riferimenti in V e V', si verifica immediatamente che la matrice associata ad f rispetto a tali riferimenti è quella nulla.*

5.2 ESEMPIO. *Sia $V = V'$, fissiamo una base $\mathcal{B}$ in V e consideriamo l'applicazione identica id_V (che è lineare). La matrice associata ad id_V rispetto alla base $\mathcal{B}$ è la matrice identica.*

5.3 ESEMPIO. *Definiamo un'applicazione $f : \mathbb{R}^2 \to \mathbb{R}^3$ ponendo $f(x, y) = (x - y, x + y, 2x)$. L'applicazione f è lineare, poiché tali sono le sue componenti. La matrice A associata ad f rispetto alle basi canoniche è*

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

come è agevole verificare calcolando $f(1, 0)$ e $f(0, 1)$.

5.4 ESEMPIO. *Siano $f, g : V \to V'$ due applicazioni lineari, e siano A, B le matrici ad esse associate rispetto a dei fissati riferimenti. Si verifica agevolmente che $f = g$ se e solo se $A = B$.*

Consideriamo ora un'applicazione lineare $\omega : \mathbb{K}^n \to \mathbb{K}^m$ tra spazi vettoriali numerici. Siano $\omega_1, \ldots, \omega_m$ le componenti di ω. In altri termini, se indichiamo, come già fatto nel Capitolo 2, con $\pi_i : \mathbb{K}^m \to \mathbb{K}$, $i = 1, \ldots, m$ la proiezione di $\mathbb{K}^m$ sull'i-mo fattore, le componenti di ω saranno le composte $\omega_i = \pi_i \circ \omega$, $(i = 1, \ldots, m)$. Fissiamo in $\mathbb{K}^n, \mathbb{K}^m$ le rispettive basi canoniche (ordinate) $\mathcal{B} = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$ ed $\mathcal{B}' = ($
‘$E_1, \ldots,$
‘$E_m)$ e consideriamo la matrice A associata ad ω rispetto a tali basi. Si ha, in particolare, che $\omega(\mathbf{e}_j) = A^j$, $j = 1, \ldots, n$.

5.5 LEMMA. *Nella situazione ora descritta* $\dim \operatorname{im} \omega = \rho(A)$.

Dimostrazione. Poiché $\mathcal{B}$ genera $\mathbb{K}^n$, il sistema $[A^1, \ldots, A^n]$ genera $\operatorname{im} \omega \leq \mathbb{K}^m$ e quindi una base di $\operatorname{im} \omega$ è fornita da un sistema massimale di colonne indipendenti si A. La cardinalità di tale base è quindi, per definizione, proprio $\rho(A)$. $\square$

5.6 Proposizione. *Sia $f : V \to V'$ un'applicazione lineare e sia $A \in \mathcal{M}_{m,n}$ la matrice associata ad f rispetto a delle basi ordinate fissate $\mathcal{B} = (\mathbf{u}_1, \ldots, \mathbf{u}_n)$, $\mathcal{B}' = (\mathbf{u}'_1, \ldots, \mathbf{u}'_m)$ di V, V'. Si ha che* $\dim \operatorname{im} f = \rho(A)$.

Dimostrazione. Consideriamo gli isomorfismi coordinati

$$\Phi : V \longrightarrow \mathbb{K}^n \ ; \quad \Psi : V' \longrightarrow \mathbb{K}^m$$

rispetto alle basi $\mathcal{B}, \mathcal{B}'$. Si verifica agevolmente che A è anche la matrice associata all'applicazione lineare

$$\omega = \Psi \circ f \circ \Phi^{-1} : \mathbb{K}^n \longrightarrow \mathbb{K}^m$$

rispetto alle basi canoniche. Osserviamo che $\Phi^{-1}(\mathbb{K}^n) = V$ e quindi

$$\operatorname{im}(f \circ \Phi^{-1}) = f \circ \Phi^{-1}(\mathbb{K}^n) = f(V) = \operatorname{im} f \ .$$

Inoltre

$$\operatorname{im} \omega = \operatorname{im}(\Psi \circ f \circ \Phi^{-1}) = \Psi\big(f \circ \Phi^{-1}(\mathbb{K}^n)\big) = \Psi(\operatorname{im} f \circ \Phi^{-1}) = \Psi(\operatorname{im} f)$$

e poiché Ψ è un isomorfismo, la sua restrizione

$$\Psi_{|\operatorname{im} f} : \operatorname{im} f \longrightarrow \operatorname{im} \omega$$

è ancora un isomorfismo. Pertanto, in base al lemma precedente, possiamo concludere che $\dim \operatorname{im} f = \dim \operatorname{im} \omega = \rho(A)$. $\square$

Consideriamo ora, oltre agli spazi vettoriali V, V' e alle loro basi ordinate $\mathcal{B}, \mathcal{B}'$, anche un terzo spazio vettoriale V'' e una sua base ordinata $\mathcal{B}'' = (\mathbf{u}''_1, \ldots, \mathbf{u}''_s)$. Siano poi $f : V \to V'$, $g : V' \to V''$ due applicazioni lineari e A, B le matrici ad esse associate rispetto alle basi fissate. Abbiamo che $A \in \mathcal{M}_{m,n}$, $B \in \mathcal{M}_{s,m}$, ed ha pertanto senso considerare la matrice $C = BA \in \mathcal{M}_{s,n}$.

5.7 Proposizione. *Nella situazione sopra descritta, la matrice associata all'applicazione lineare composta $g \circ f$, rispetto alle basi $\mathcal{B}, \mathcal{B}''$, è C.*

Dimostrazione. Per ogni $\mathbf{v} \in V$, detti X, Y, Z i vettori numerici colonna delle componenti di $\mathbf{v}, f(\mathbf{v}), gf(\mathbf{v})$ rispetto alle basi $\mathcal{B}, \mathcal{B}', \mathcal{B}''$, abbiamo che $Y = AX$ e $Z = BY$. Pertanto

$$Z = B(AX) = (BA)X$$

e quindi, posto $C = BA$, C è la matrice associata a $g \circ f$ rispetto alle basi $\mathcal{B}, \mathcal{B}''$. $\square$

5.8 Corollario. *Sia $f : V \to V'$ un isomorfismo. Fissate delle basi $\mathcal{B}, \mathcal{B}'$ in V, V' e dette A, B le matrici associate ad f, f^{-1} rispetto a tali basi si ha che A è una matrice invertibile e che $B = A^{-1}$.*

Dimostrazione. Osserviamo che se $V \cong V'$ allora V e V' hanno la stessa dimensione, diciamo n, e quindi A, B sono entrambe quadrate di ordine n. Poiché $f^{-1} \circ f = id_V$, in base ai risultati ed esempi precedenti si ha che la matrice associata ad id_V rispetto alla base $\mathcal{B}$, e cioè I_n, è proprio il prodotto righe per colonne di B ed A, ovvero $I_n = BA$. Ripetendo lo stesso ragionamento usando il fatto che $f \circ f^{-1} = id_{V'}$ si trova che $AB = I_n$, da cui l'asserto. $\square$

Osserviamo che si può procedere anche in modo opposto a quanto fatto finora: a partire da una matrice costruiremo delle applicazioni lineari (tra spazi vettoriali numerici o anche di tipo generale) che ammettono tale matrice come matrice ad esse associata rispetto a basi fissate. Sia dunque $A \in \mathcal{M}_{m,n}$ una matrice e definiamo un'applicazione

$$\omega_A : \mathbb{K}^n \to \mathbb{K}^m \tag{5.1}$$

al modo seguente. Per ogni vettore numerico di ordine n scritto in forma di colonna $\beta = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix}$, poniamo

$$\omega_A(\beta) = A \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} \in \mathbb{K}^m .$$

Si verifica che ω_A è lineare (esercizio: usare la distributività del prodotto righe per colonne rispetto alla somma di matrici) e che la matrice associata ad ω_A rispetto alle basi canoniche è proprio A.

Siano ora V, V' due spazi vettoriali di dimensione n, m e $\mathcal{B}, \mathcal{B}'$ due basi ordinate di V, V' rispettivamente. Fissata una matrice $A \in \mathcal{M}_{m,n}$ esiste un'unica applicazione lineare $f : V \to V'$ che ammette A come matrice associata rispetto a $\mathcal{B}, \mathcal{B}'$. f si costruisce come segue. Indichiamo con Φ, Ψ gli isomorfismi coordinati di V, V' rispetto a $\mathcal{B}, \mathcal{B}'$, consideriamo l'applicazione lineare ω_A e poniamo

$$f = \Psi \circ \omega_A \circ \Phi^{-1} .$$

Tale applicazione è lineare, in quanto composta di applicazioni lineari. È poi agevole verificare che la matrice associata ad f rispetto alle basi fissate è proprio A. Il Corollario 5.8 si inverte, nel seguente senso.

5.9 Corollario. *Sia $A \in \mathcal{M}_{n,n}$ una matrice invertibile. Siano V, V' due spazi vettoriali, di dimensione n e $\mathcal{B}, \mathcal{B}'$ delle basi ordinate fissate in tali spazi* *vettoriali. Sia inoltre $f : V \to V'$ l'applicazione lineare tale che A sia la matrice associata ad f rispetto alle basi $\mathcal{B}, \mathcal{B}'$. Allora f è un isomorfismo e A^{-1} è la matrice associata all'isomorfismo inverso f^{-1} rispetto a $\mathcal{B}', \mathcal{B}$.*

Dimostrazione. Sia $g : V' \to V$ tale che A^{-1} sia la matrice associata a g rispetto a $\mathcal{B}', \mathcal{B}$. Allora la matrice associata a $gf : V \to V$ è $A^{-1}A = I_n$ e quindi $gf = id_V$. Analogamente si vede che $fg = id_{V'}$. Quindi f è un isomorfismo, $g = f^{-1}$ e A^{-1} è la matrice associata a f^{-1} rispetto a $\mathcal{B}', \mathcal{B}$. $\square$

Il lettore potrà verificare, per esercizio, utilizzando le tecniche sopra descritte, che se $A \in \mathcal{M}_{m,n}$ e $B \in \mathcal{M}_{n,s}$ si ha che $\rho(A \cdot B) \leq \rho(A)$ e che $\rho(A \cdot B) \leq \rho(B)$.

Concludiamo questa sezione con una osservazione sul prodotto tra matrici. Come già sottolineato precedentemente, tale prodotto non è integro, nel senso che può accadere che $A \neq O$, $B \neq O$ e $AB = O$. Però se $A \in \mathcal{M}_{m,n}$ e supponiamo che per ogni vettore colonna $Y \in \mathbb{K}^n$ il prodotto righe per colonne AY sia nullo, vuol dire che ω_A è l'applicazione lineare nulla e quindi A è la matrice nulla.

5.2 Cambiamenti di riferimento

Ci chiediamo ora che relazione esiste tra le componenti di un vettore in due diverse basi. Consideriamo quindi la seguente situazione. Sia V un $\mathbb{K}$-spazio vettoriale e consideriamo due sue basi ordinate $\mathcal{B} = (\mathbf{u}_1, \ldots, \mathbf{u}_n)$, $\tilde{\mathcal{B}} = (\tilde{\mathbf{u}}_1, \ldots, \tilde{\mathbf{u}}_n)$. Siano inoltre $\Phi, \tilde{\Phi} : V \to \mathbb{K}^n$ gli isomorfismi coordinati di V rispetto a $\mathcal{B}$, $\tilde{\mathcal{B}}$. Per ogni vettore $\mathbf{v} \in V$ sono univocamente determinati gli scalari $\beta_1, \ldots, \beta_n$ e gli scalari $\tilde{\beta}_1, \ldots, \tilde{\beta}_n$ tali che

$$\mathbf{v} = \sum_{j=1}^{n} \beta_j \mathbf{u}_j = \sum_{j=1}^{n} \tilde{\beta}_j \tilde{\mathbf{u}}_j \tag{5.2}$$

ovvero

$$\Phi(\mathbf{v}) = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} \quad ; \quad \tilde{\Phi}(\mathbf{v}) = \begin{pmatrix} \tilde{\beta}_1 \\ \vdots \\ \tilde{\beta}_n \end{pmatrix} .$$

Costruiamo una matrice $B = (b_{i,j}) \in \mathcal{M}_{n,n}$ ponendo $B^j = \tilde{\Phi}(\mathbf{u}_j)$. In altri termini, la colonna j-ma di B è costituita dalle componenti $b_{1,j}, \ldots, b_{n,j}$ del vettore $\mathbf{u}_j$ in $\tilde{\mathcal{B}}$:

$$\mathbf{u}_j = b_{1,j}\tilde{\mathbf{u}}_1 + \cdots + b_{n,j}\tilde{\mathbf{u}}_n \ . \tag{5.3}$$

Dalle (5.2),(5.3) deduciamo che

$$\begin{aligned}\sum_{i=1}^{n} \tilde{\beta}_i \tilde{\mathbf{u}}_i &= \sum_{j=1}^{n} \beta_j \mathbf{u}_j \\ &= \sum_{j=1}^{n} \beta_j \Big(\sum_{i=1}^{n} b_{i,j} \tilde{\mathbf{u}}_i \Big) \\ &= \sum_{i=1}^{n} \Big(\sum_{j=1}^{n} b_{i,j} \beta_j \Big) \tilde{\mathbf{u}}_i\end{aligned}$$

e quindi

$$\tilde{\beta}_i = \sum_{j=1}^{n} b_{i,j} \beta_j \qquad \forall\ i\ . \tag{5.4}$$

Le relazioni scalari (5.4) corrispondono alla relazione vettoriale

$$\begin{pmatrix} \tilde{\beta}_1 \\ \vdots \\ \tilde{\beta}_n \end{pmatrix} = B \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} \ . \tag{5.5}$$

La (5.5) si dice formula di cambiamento delle componenti relativa al passaggio da $\mathcal{B}$ a $\tilde{\mathcal{B}}$ e B prende il nome di matrice di passaggio da $\mathcal{B}$ a $\tilde{\mathcal{B}}$.

Il lettore potrà provare, per esercizio, che la matrice associata all'applicazione lineare $id_V : V \to V$ rispetto alle basi $\mathcal{B}$, $\tilde{\mathcal{B}}$ è proprio B.

Dall'esercizio ora proposto si può dedurre che B è invertibile. Sia infatti B' la matrice associata a id_V rispetto alle basi $\tilde{\mathcal{B}}$, $\mathcal{B}$. Schematizziamo tale situazione indicando di fianco allo spazio vettoriale in questione anche la base fissata e sotto la freccia la matrice associata. Si ha che

$$V, \mathcal{B} \xrightarrow[B]{id_V} V, \tilde{\mathcal{B}} \xrightarrow[B']{id_V} V, \mathcal{B}\ .$$

In base ad un risultato già acquisito, la matrice associata alla composta

$$id_V = id_V \circ id_V : V, \mathcal{B} \longrightarrow V, \mathcal{B}$$

è $B' \cdot B$, ma sappiamo anche che essa coincide con I_n e quindi $B' \cdot B = I_n$. Analogamente si verifica che $B \cdot B' = I_n$. Pertanto B è invertibile e B' è la sua inversa. In modo analogo si prova che se $\bar{\mathcal{B}}$ è un'altra base e B'' è la matrice associata ad id_V rispetto alle basi $\tilde{\mathcal{B}}$ e $\bar{\mathcal{B}}$ allora la matrice associata ad id_V rispetto a $\mathcal{B}$ e $\bar{\mathcal{B}}$ è $B'' \cdot B$. Tutte le osservazioni fatte finora possono essere sintetizzate come segue. Indichiamo con $M_{\mathcal{B}}^{\tilde{\mathcal{B}}}$ la matrice associata ad id_V rispetto a $\mathcal{B}$, $\tilde{\mathcal{B}}$ (ovvero, per quanto detto la matrice di passaggio da $\mathcal{B}$ a $\tilde{\mathcal{B}}$). Valgono le seguenti relazioni

$$M_{\mathcal{B}}^{\mathcal{B}} = I_n\ ; \quad M_{\tilde{\mathcal{B}}}^{\bar{\mathcal{B}}} \cdot M_{\mathcal{B}}^{\tilde{\mathcal{B}}} = M_{\mathcal{B}}^{\bar{\mathcal{B}}}\ .$$

In particolare
$$M_{\tilde{\mathcal{B}}}^{\mathcal{B}} \cdot M_{\mathcal{B}}^{\tilde{\mathcal{B}}} = M_{\mathcal{B}}^{\mathcal{B}} = I_n = M_{\tilde{\mathcal{B}}}^{\tilde{\mathcal{B}}} = M_{\mathcal{B}}^{\tilde{\mathcal{B}}} \cdot M_{\tilde{\mathcal{B}}}^{\mathcal{B}}$$
e quindi
$$M_{\tilde{\mathcal{B}}}^{\mathcal{B}} = \left(M_{\mathcal{B}}^{\tilde{\mathcal{B}}}\right)^{-1} .$$

5.3 Alcune applicazioni dei determinanti

Usiamo ora la teoria dei determinanti per affrontare la seguente questione. Sia $f : V \to V$ un endomorfismo e siano $\mathcal{B}, \tilde{\mathcal{B}}$ delle basi ordinate di V. Siano inoltre
$$A = M_{\mathcal{B}}^{\mathcal{B}}(f) \ ; \quad \tilde{A} = M_{\tilde{\mathcal{B}}}^{\tilde{\mathcal{B}}}(f) \ .$$
In altri termini, A è la matrice associata ad f quando in V sia stata fissata la base $\mathcal{B}$ (sia quando ci riferiamo a V come il dominio di f che quando ci riferiamo a V come codominio), e $\tilde{A}$ è la matrice associata ad f quando in V sia stata fissata la base $\tilde{\mathcal{B}}$. Studiamo la relazione tra A e $\tilde{A}$. Posto $B = M_{\mathcal{B}}^{\tilde{\mathcal{B}}}$ abbiamo, graficamente, la seguente situazione
$$V, \tilde{\mathcal{B}} \xrightarrow[B^{-1}]{id_V} V, \mathcal{B} \xrightarrow[A]{f} V, \tilde{\mathcal{B}} \xrightarrow[B]{id_V} V, \tilde{\mathcal{B}} \ .$$
D'altra parte abbiamo anche che
$$V, \tilde{\mathcal{B}} \xrightarrow[\tilde{A}]{f = id_V \circ f \circ id_V} V, \tilde{\mathcal{B}} \ .$$
Poiché la matrice associata alla composta $f = id_V \circ f \circ id_V$ coincide con il prodotto delle matrici associate alle singole applicazioni id_V, f, id_V, si ha che
$$\tilde{A} = B \cdot A \cdot B^{-1} \ .$$
Pertanto, abbiamo che
$$M_{\tilde{\mathcal{B}}}^{\tilde{\mathcal{B}}}(f) = M_{\mathcal{B}}^{\tilde{\mathcal{B}}} \cdot M_{\mathcal{B}}^{\mathcal{B}}(f) \cdot M_{\tilde{\mathcal{B}}}^{\mathcal{B}} \ .$$

Questo procedimento suggerisce la seguente

5.10 Definizione. *Siano $A, \tilde{A} \in \mathcal{M}_{n,n}$. Diremo che A e $\tilde{A}$ sono coniugate (e scriveremo $A \sim \tilde{A}$) se esiste una matrice $B \in GL_n(\mathbb{K})$ tale che*
$$\tilde{A} = BAB^{-1} \ .$$

Il lettore potrà verificare agevolmente, per esercizio, che $\sim$ è una relazione di equivalenza in $\mathcal{M}_{n,n}$ e che $A \sim \tilde{A} \Rightarrow \rho(A) = \rho(\tilde{A})$.

Il ragionamento esposto all'inizio di questo paragrafo si inverte, nel senso che ora precisiamo.

5.11 TEOREMA. *Se V e $\mathcal{B} = (\mathbf{u}_1, \ldots, \mathbf{u}_n)$ sono come sopra e $B \in GL_n(\mathbb{K})$, esiste un'unica base ordinata $\tilde{\mathcal{B}} = (\tilde{\mathbf{u}}_1, \ldots, \tilde{\mathbf{u}}_n)$ tale che $B = M_{\mathcal{B}}^{\tilde{\mathcal{B}}}$.*

Dimostrazione. Sia $C = (c_{i,j})$ l'inversa di B e poniamo, per ogni $i = 1, \ldots, n$,

$$\tilde{\mathbf{u}}_i = \sum_{j=1}^{n} c_{j,i} \mathbf{u}_j \ .$$

Poniamo inoltre

$$\tilde{\mathcal{B}} = (\tilde{\mathbf{u}}_1, \ldots, \tilde{\mathbf{u}}_n) \ .$$

Osserviamo che, per ogni $h = 1, \ldots, n$,

$$\begin{aligned} \sum_{i=1}^{n} b_{i,h} \tilde{\mathbf{u}}_i &= \sum_{i=1}^{n} b_{i,h} \Big(\sum_{j=1}^{n} c_{j,i} \mathbf{u}_j \Big) \\ &= \sum_{j=1}^{n} \Big(\sum_{i=1}^{n} b_{i,h} c_{j,i} \Big) \mathbf{u}_j \\ &= \sum_{j=1}^{n} \delta_{j,h} \mathbf{u}_j \\ &= \mathbf{u}_h \end{aligned}$$

cioè

$$\mathbf{u}_h = \sum_{i=1}^{n} b_{i,h} \tilde{\mathbf{u}}_i \ . \tag{5.6}$$

Infatti

$$\begin{aligned} \sum_{i=1}^{n} b_{i,h} c_{j,i} &= \sum_{i=1}^{n} b_{i,h} \frac{(-1)^{i+j} \det B_{(i,j)}}{\det B} \\ &= \frac{1}{\det B} \sum_{i=1}^{n} (-1)^{i+j} b_{i,h} \det B_{(i,j)} \ . \end{aligned} \tag{5.7}$$

□

Se nella (5.7) poniamo $h = j$ si ha che

$$\sum_{i=1}^{n} (-1)^{i+j} b_{i,h} \det B_{(i,j)} = \underbrace{\sum_{i=1}^{n} (-1)^{i+j} b_{i,j} \det B_{(i,j)} = \det B}_{Laplace} \ .$$

Se invece $h \neq j$,

$$\sum_{i=1}^{n} (-1)^{i+j} b_{i,h} \det B_{(i,j)} = 0$$

in quanto tale espressione è il determinante della matrice

$$(B^1, \dots, B^h, \dots, B^h \dots, B^n)$$

dove la colonna B^h compare sia all'h-mo che al j-mo posto, sviluppato secondo la j-ma colonna (in pratica si applica il 2^o teorema di Laplace alle colonne). In ogni caso quindi

$$\sum_{i=1}^{n} b_{i,h} c_{j,i} = \delta_{j,h} \ .$$

La (5.6) ci assicura che i vettori di base $\mathbf{u}_h$ dipendono da $\tilde{\mathcal{B}}$. Quindi $\tilde{\mathcal{B}}$ genera V. Se il sistema $\tilde{\mathcal{B}}$ fosse dipendente, da esso si potrebbe estrarre una base di V di cardinalità minore di n. Ciò non può accadere, in quanto tutte le basi sono equipotenti e $\mathcal{B}$ ha n elementi. Pertanto $\tilde{\mathcal{B}}$ è indipendente e quindi è una base di V. Osserviamo infine che, per costruzione, $C = B^{-1} = M^{\mathcal{B}}_{\tilde{\mathcal{B}}}$ e quindi $B = M^{\tilde{\mathcal{B}}}_{\mathcal{B}}$.

Possiamo quindi concludere che due matrici quadrate dello stesso ordine sono coniugate se e solo se esse possono vedersi come le matrici associate ad una stessa applicazione lineare, rispetto a basi diverse (scelte opportunamente).

5.4 Autovettori e autovalori

Sia V uno spazio vettoriale su un campo $\mathbb{K}$, sia $\dim V = n$ e sia $f : V \to V$ un endomorfismo.

5.12 Definizione. *Un vettore non nullo $\mathbf{v} \in V$ si dice autovettore di f se esiste uno scalare λ tale che*

$$f(\mathbf{v}) = \lambda \mathbf{v} \ . \tag{5.8}$$

5.13 Definizione. *Uno scalare λ si dice autovalore di f se esiste un autovettore $\mathbf{v}$ tale che valga la* (5.8).

Se $\mathbf{v} \neq \mathbf{0}$ e vale la (5.8), diremo che $\mathbf{v}$ è un autovettore associato a λ e che λ è un autovalore associato a $\mathbf{v}$.

5.14 Lemma. *Sia $\mathbf{v}$ un autovettore. Esiste allora un unico autovalore ad esso associato.*

Dimostrazione. Siano λ, μ degli scalari tali che

$$f(\mathbf{v}) = \lambda \mathbf{v} = \mu \mathbf{v} \ .$$

Si ha allora che $(\lambda - \mu)\mathbf{v} = \mathbf{0}$ e quindi, essendo $\mathbf{v} \neq \mathbf{0}$, deve accadere che $\lambda - \mu = 0$, ovvero $\lambda = \mu$, e ciò prova l'unicità dell'autovalore associato all'autovettore $\mathbf{v}$. □

Sia ora $\lambda \in \mathbb{K}$ e definiamo il sottoinsieme $V_\lambda(f)$, o più semplicemente V_λ, ponendo

$$V_\lambda = \left\{ \mathbf{v} \in V \;\middle|\; f(\mathbf{v}) = \lambda \mathbf{v} \right\} \subseteq V \ .$$

5.15 Proposizione. *Per ogni scalare λ, il sottoinsieme V_λ di V è un sottospazio vettoriale.*

Dimostrazione. È chiaro che $\mathbf{0} \in V_\lambda$, in quanto

$$f(\mathbf{0}) = \mathbf{0} = \lambda \mathbf{0} \qquad \forall\, \lambda \in \mathbb{K} \ .$$

Siano ora $\alpha, \beta \in \mathbb{K}$, $\mathbf{u}, \mathbf{v} \in V_\lambda$. Allora $f(\mathbf{u}) = \lambda \mathbf{u}$ e $f(\mathbf{v}) = \lambda \mathbf{v}$. Pertanto

$$f(\alpha \mathbf{u} + \beta \mathbf{v}) = \alpha f(\mathbf{u}) + \beta f(\mathbf{v}) = \alpha \lambda \mathbf{u} + \beta \lambda \mathbf{v} = \lambda(\alpha \mathbf{u} + \beta \mathbf{v})$$

e quindi $\alpha \mathbf{u} + \beta \mathbf{v} \in V_\lambda$. □

Osserviamo che $V_\lambda \neq \{\mathbf{0}\}$ se e solo se esiste un vettore non nullo $\mathbf{v}$ tale che $f(\mathbf{v}) = \lambda \mathbf{v}$, e cioè se e solo se λ è un autovalore di f.

5.16 Definizione. *Se λ è un autovalore di f, il sottospazio vettoriale V_λ (necessariamente non banale) di V si dice autospazio di f associato all'autovalore λ.*

Pertanto l'autospazio V_λ associato ad un autovalore λ è costituito dal vettore nullo $\mathbf{0}$ e dagli autovettori associati a λ.

5.17 Definizione. *L'endomorfismo f si dice diagonalizzabile se esiste una base ordinata $\mathcal{D} = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$ costituita da autovettori. Una base siffatta si dice anche base spettrale.*

Osserviamo che se $\mathcal{D} = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$ è una base ordinata di autovettori e $\lambda_1, \ldots, \lambda_n$ sono gli autovalori associati agli autovettori $\mathbf{v}_1, \ldots, \mathbf{v}_n$, la matrice D associata ad f rispetto a tale base è la matrice diagonale

$$D = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix} \ . \tag{5.9}$$

Viceversa, se $\mathcal{D}$ è una base ordinata e la matrice D associata ad f rispetto a $\mathcal{D}$ è diagonale ed è data dalla (5.9), allora $\mathcal{D}$ è una base di autovettori e gli elementi della diagonale di D sono gli autovalori associati a tali autovettori.

Le osservazioni precedenti giustificano in qualche modo la terminologia usata e possono essere formalizzate nel seguente enunciato.

5.18 Teorema. *L'endomorfismo f è diagonalizzabile se e solo se esiste una base ordinata $\mathcal{D}$ tale che la matrice D associata ad f rispetto a $\mathcal{D}$ risulti diagonale.*

Tenendo presente il Teorema 5.11 e quanto detto a proposito delle matrici coniugate, gli endomorfismi diagonalizzabili possono anche essere caratterizzati come segue.

5.19 Teorema. *Sia $f : V \to V$ un endomorfismo e sia $\mathcal{B}$ una base ordinata di V. Allora f è diagonalizzabile se e solo se la matrice $A = M_{\mathcal{B}}^{\mathcal{B}}(f)$ è simile ad una matrice diagonale.*

Sia ora $\mathcal{B} = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$ una base ordinata, $A = (a_{i,j}) \in \mathcal{M}_{n,n}$ la matrice associata ad f rispetto a $\mathcal{B}$ e $\Phi : V \to \mathbb{K}^n$ l'isomorfismo coordinato associato a $\mathcal{B}$.

5.20 Teorema di caratterizzazione degli autovalori. *Uno scalare λ è un autovalore di f se e solo se* $\det(A - \lambda I_n) = 0$.

Dimostrazione. Ricordiamo che lo scalare λ è un autovalore se e solo se esiste un vettore non nullo $\mathbf{w}$ tale che $f(\mathbf{w}) = \lambda\mathbf{w}$. Sia dunque $\mathbf{w}$ un vettore non nullo e sia

$$W = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} = \Phi(\mathbf{w})$$

il suo vettore coordinato rispetto a $\mathcal{B}$. Se $f(\mathbf{w}) = \mathbf{w}'$ e

$$W' = \begin{pmatrix} w'_1 \\ \vdots \\ w'_n \end{pmatrix} = \Phi(\mathbf{w}')$$

abbiamo che $\mathbf{w}' = f(\mathbf{w}) = \lambda\mathbf{w} \iff W' = \lambda W$, come si deduce facilmente dal confronto delle componenti. D'altra parte, si è già osservato che la matrice A è tale che risulti $W' = AW$. Quindi $f(\mathbf{w}) = \lambda\mathbf{w}$ se e solo se $AW = \lambda W$. Ma λW può anche scriversi come $(\lambda I_n)W$ e quindi λ è un autovalore, e $\mathbf{w}$ è un autovettore ad esso associato, se e solo se

$$AW - (\lambda I_n)W = \mathbf{0}$$

ovvero se e solo se

$$(A - \lambda I_n)W = \mathbf{0}$$

ovvero ancora se e solo se la n-pla $W = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}$ è una soluzione non banale del sistema lineare omogeneo (scritto in forma matriciale)

$$(A - \lambda I_n)X = \mathbf{0} \ . \tag{5.10}$$

In definitiva λ è un autovalore se e solo se esiste una soluzione non banale W del sistema lineare omogeneo (5.10), e ciò avviene se e solo se

$$\det(A - \lambda I_n) = 0 \ .$$

$\square$

Osserviamo che

$$A - \lambda I_n = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \dots & a_{n,n} \end{pmatrix} - \begin{pmatrix} \lambda & 0 & \dots & 0 \\ 0 & \lambda & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda \end{pmatrix}$$
$$= \begin{pmatrix} a_{1,1} - \lambda & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} - \lambda & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \dots & a_{n,n} - \lambda \end{pmatrix}$$

e quindi $\det(A - \lambda I_n)$ è una espressione polinomiale in λ. Infatti si ha

$$\det(A - \lambda I_n) = (-1)^n \lambda^n + \alpha_{n-1}\lambda^{n-1} + \cdots + \alpha_0$$

dove $\alpha_0, \dots, \alpha_{n-1}$ sono scalari opportuni. Se poniamo allora

$$p_A = \alpha_0 + \alpha_1 x + \cdots + \alpha_{n-1}x^{n-1} + (-1)^n x^n$$

la discussione precedente ci assicura che λ è un autovalore di f se e solo se è una radice del polinomio p_A.

5.21 Esempio. *Sia V uno spazio vettoriale su $\mathbb{R}$ di dimensione 3 e sia $\mathcal{B} = (\mathbf{u}, \mathbf{v}, \mathbf{w})$ una sua base. Definiamo un endomorfismo $f : V \to V$ ponendo $f(\mathbf{u}) = \mathbf{u} + \mathbf{v}$, $f(\mathbf{v}) = \mathbf{u} + \mathbf{w}$, $f(\mathbf{w}) = -3\mathbf{u} - \mathbf{v}$. La matrice A associata ad f rispetto a $\mathcal{B}$ è*

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

Pertanto

$$\det(A - \lambda I_3) = \begin{vmatrix} 1-\lambda & 1 & -3 \\ 1 & -\lambda & -1 \\ 0 & 1 & -\lambda \end{vmatrix} = -\lambda^3 + \lambda^2 - 2$$

(cioè $\alpha_0 = -2,\ \alpha_1 = 0,\ \alpha_2 = 1$), e si ha che

$$p_A(x) = -x^3 + x^2 - 2 = -(x+1)(x^2 - 2x + 2)\ .$$

Da ciò si deduce che $\lambda = -1$ è un autovalore (l'unico peraltro) di f.

Poiché gli autovalori di un endomorfismo f dipendono unicamente da f stesso, non certo da eventuali basi fissate in V, si deduce che la costruzione del polinomio p_A *potrebbe* dipendere dalla base $\mathcal{B}$ fissata e dalla matrice A associata ad f rispetto a $\mathcal{B}$, ma le radici di p_A, che sono appunto gli autovalori di f, non possono dipendere dalla scelta di $\mathcal{B}$. Si può dire in effetti qualcosa di più: il polinomio p_A non dipende dalla base $\mathcal{B}$ scelta.

5.22 Teorema di invarianza del polinomio caratteristico. *Siano $\mathcal{B}, \mathcal{B}'$ due basi di V e siano A, A' le matrici associate all'endomorfismo f rispetto a tali basi. Abbiamo che $p_A = p_{A'}$.*

Dimostrazione. Se C è la matrice di passaggio da $\mathcal{B}$ a $\mathcal{B}'$ si ha che

$$A' = C^{-1}AC$$

e quindi

$$\begin{aligned}
\det(A' - xI_n) &= \det(C^{-1}AC - xI_n) \\
&= \det(C^{-1}AC - C^{-1}xI_nC) \\
&= \det(C^{-1}(A - xI_n)C) \\
&= \det(C^{-1})\det(A - xI_n)\det(C) \\
&= \det(A - xI_n)
\end{aligned}$$

e cioè i polinomi $p_A, p_{A'}$ coincidono. □

Pertanto il polinomio p_A non dipende dalla base scelta, ma solo da f.

5.23 Definizione. *Il polinomio p ottenuto fissando una qualunque base $\mathcal{B}$ di V e ponendo $p = p_A$ prende il nome di polinomio caratteristico dell'endomorfismo f.*

In base a tale definizione ed al teorema precedente, si ha che λ è un autovalore di f se e solo se è una radice del polinomio caratteristico p di f. In definitiva, in base alla discussione precedente, il polinomio caratteristico p di un endomorfismo f si può definire ponendo

$$p = \det(A - xI_n)$$

dove A è la matrice associata ad f rispetto ad una qualunque base fissata.

5.5 Diagonalizzazione

Sia λ un autovalore di f.

5.24 Definizione. *La molteplicità algebrica $m(\lambda)$ dell'autovalore λ è la molteplicità di λ come radice del polinomio caratteristico p.*

Pertanto $m(\lambda) = h$ se $(x-\lambda)^h$ divide $p(x)$ ma $(x-\lambda)^{h+1}$ non divide $p(x)$. L'autovalore λ si dirà semplice se $m(\lambda) = 1$, multiplo se $m(\lambda) > 1$, in particolare si dirà doppio se $m(\lambda) = 2$, triplo se $m(\lambda) = 3$, e così via.

5.25 Definizione. *La molteplicità geometrica dell'autovalore λ è la dimensione* $\dim V_\lambda$ *dell'autospazio V_λ.*

Osserviamo esplicitamente che la molteplicità geometrica di un autovalore λ è sempre strettamente positiva. Infatti poiché λ è un autovalore, il sottospazio V_λ è certamente non banale, e cioè $\dim V_\lambda \geq 1$.

5.26 Teorema sulle Molteplicità. *Se λ è un autovalore dell'endomorfismo f, si ha che*

$$\dim V_\lambda \leq m(\lambda) \ .$$

Dimostrazione. Sia $t = \dim V_\lambda$. È chiaro che $t \leq n$ e $m(\lambda) \leq n$. Sia $\mathcal{B}_\lambda = (\mathbf{v}_1, \ldots, \mathbf{v}_t)$ una base ordinata di V_λ. Si ha quindi che

$$f(\mathbf{v}_i) = \lambda \mathbf{v}_i \, , \qquad \forall \; i = 1, \ldots, t \ . \tag{5.11}$$

Sia $\mathcal{B} = (\mathbf{v}_1, \ldots, \mathbf{v}_t, \mathbf{v}_{t+1}, \ldots, \mathbf{v}_n)$ una base ordinata di V ottenuta completando la base $\mathcal{B}_\lambda$ di V_λ e A la matrice associata ad f rispetto a $\mathcal{B}$. Abbiamo che

$$A = \left(\begin{array}{cccc|c} \lambda & 0 & \ldots & 0 & \\ 0 & \lambda & \ldots & 0 & H \\ \vdots & \vdots & \ddots & \vdots & \\ 0 & 0 & \ldots & \lambda & \\ \hline & & O & & K \end{array} \right) \begin{array}{l} \left.\vphantom{\begin{array}{c}\lambda\\0\\\vdots\\0\end{array}}\right\} t \\ \}\, n-t \end{array}$$

dove $H \in \mathcal{M}_{t,n-t}$, $K \in \mathcal{M}_{n-t,n-t}$ sono opportune matrici e $O \in \mathcal{M}_{n-t,t}$ è la matrice nulla. Infatti la (5.11) ci assicura che i vettori coordinati dei vettori immagine $f(\mathbf{v}_1), \ldots, f(\mathbf{v}_t)$ sono rispettivamente i vettori numerici colonna di ordine n

$$\begin{pmatrix}\lambda\\0\\\vdots\\0\\0\\\vdots\\0\end{pmatrix} ; \begin{pmatrix}0\\\lambda\\\vdots\\0\\0\\\vdots\\0\end{pmatrix} ; \dots ; \begin{pmatrix}0\\\vdots\\0\\\lambda\\0\\\vdots\\0\end{pmatrix} \qquad \begin{matrix}\left.\vphantom{\begin{matrix}0\\0\\0\\0\end{matrix}}\right\} t\\ \left.\vphantom{\begin{matrix}0\\0\\0\end{matrix}}\right\} n-t\end{matrix}$$

che sono proprio le prime t colonne di A. Ma allora

$$\begin{aligned} p &= \det(A - xI_n) \\ &= \det \left(\begin{array}{cccc|c} \lambda - x & 0 & \dots & 0 & \\ 0 & \lambda - x & \dots & 0 & \\ \vdots & \vdots & \ddots & \vdots & H \\ 0 & 0 & \dots & \lambda - x & \\ \hline & & O & & K - xI_{n-t} \end{array}\right) \\ &= (\lambda - x)^t \cdot \det(K - xI_{n-t}) \end{aligned}$$

e quindi λ è una radice di p di molteplicità almeno t. $\square$

Sia ora λ un autovalore di f, $\mathcal{B}$ una base ordinata di V e $A = M_{\mathcal{B}}^{\mathcal{B}}(f)$. L'autospazio V_λ è l'insieme dei vettori $\mathbf{w}$ tali che il vettore delle componenti W di $\mathbf{w}$ in $\mathcal{B}$ sia una soluzione del sistema lineare omogeneo

$$(A - \lambda I_n)X = \mathbf{0} \ .$$

Detto $\mathrm{Sol}(\lambda)$ l'insieme delle soluzioni di tale sistema, abbiamo che $\mathrm{Sol}(\lambda) \leq \mathbb{K}^n$ e $\dim \mathrm{Sol}(\lambda) = n - \rho(A - \lambda I_n)$. Inoltre l'isomorfismo coordinato $\Phi_{\mathcal{B}} : V \to \mathbb{K}^n$ induce, per restrizione, un isomorfismo $\phi_\lambda : V_\lambda \longrightarrow \mathrm{Sol}(\lambda)$. Pertanto $\dim V_\lambda = n - \rho(A - \lambda I_n)$.

Al fine di fornire alcune caratterizzazioni della diagonalizzabilità di un endomorfismo, premettiamo il seguente risultato.

5.27 Lemma. *Sia f un endomorfismo e siano $\lambda_1, \dots, \lambda_t$ degli autovalori di f, a due a due distinti. Inoltre, per ogni $i = 1, \dots, t$, sia $\mathbf{v}_i$ un autovettore associato all'autovalore λ_i. Allora il sistema $[\mathbf{v}_1, \dots, \mathbf{v}_t]$ è indipendente.*

Dimostrazione. Se $t = 1$, l'enunciato è ovvio. Procediamo allora per induzione. Supponiamo che $m \geq 1$ e che l'enunciato sia vero per ogni $t \leq m$, e proviamo che l'enunciato è vero anche per $t = m+1$. Siano dunque $\lambda_1, \dots, \lambda_m, \lambda_{m+1}$ degli autovalori di f a due a due distinti e per ogni $i = 1, \dots, m + 1$ scegliamo un autovettore $\mathbf{v}_i$ associato a λ_i. Vogliamo provare che il sistema $[\mathbf{v}_1, \dots, \mathbf{v}_{m+1}]$ è indipendente. Supponiamo allora che $\alpha_1, \ldots, \alpha_m, \beta$ siano degli scalari tali che

$$\alpha_1 \mathbf{v}_1 + \cdots + \alpha_m \mathbf{v}_m + \beta \mathbf{v}_{m+1} = \mathbf{0} \ . \tag{5.12}$$

Bisogna verificare che tali scalari sono tutti nulli. Moltiplichiamo entrambi i membri della (5.12) per λ_{m+1} ed otteniamo

$$\alpha_1 \lambda_{m+1} \mathbf{v}_1 + \cdots + \alpha_m \lambda_{m+1} \mathbf{v}_m + \beta \lambda_{m+1} \mathbf{v}_{m+1} = \mathbf{0} \ .$$

Applichiamo ora f ad entrambi i membri della (5.12). Otteniamo

$$\alpha_1 \lambda_1 \mathbf{v}_1 + \cdots + \alpha_m \lambda_m \mathbf{v}_m + \beta \lambda_{m+1} \mathbf{v}_{m+1} = \mathbf{0} \ .$$

Sottraendo membro a membro otteniamo

$$\alpha_1 (\lambda_{m+1} - \lambda_1) \mathbf{v}_1 + \cdots + \alpha_m (\lambda_{m+1} - \lambda_m) \mathbf{v}_m = \mathbf{0} \ .$$

Pertanto, poiché per ipotesi induttiva i vettori $\mathbf{v}_1, \ldots, \mathbf{v}_m$ sono indipendenti, si avrà che $\alpha_1 = \cdots = \alpha_m = 0$. La (5.12) si riduce quindi a

$$\beta \mathbf{v}_{m+1} = \mathbf{0}$$

e poiché $\mathbf{v}_{m+1} \neq \mathbf{0}$ avremo anche che $\beta = 0$. $\square$

5.28 Corollario. *Siano $\lambda_1, \ldots, \lambda_t$ degli autovalori a due a due distinti dell'endomorfismo f. Allora il sottospazio congiungente $W = V_{\lambda_1} + \cdots + V_{\lambda_t} \leq V$ è una somma diretta.*

Dimostrazione. Sia $\mathbf{w} \in W$. Esistono allora dei vettori

$$\mathbf{v}_1 \in V_{\lambda_1}, \ldots, \mathbf{v}_t \in V_{\lambda_t}$$

tali che

$$\mathbf{w} = \mathbf{v}_1 + \cdots + \mathbf{v}_t \ .$$

Supponiamo che esistano anche dei vettori $\mathbf{u}_1 \in V_{\lambda_1}, \ldots, \mathbf{u}_t \in V_{\lambda_t}$ tali che

$$\mathbf{w} = \mathbf{u}_1 + \cdots + \mathbf{u}_t \ .$$

Sottraendo membro a membro abbiamo che

$$\mathbf{0} = (\mathbf{v}_1 - \mathbf{u}_1) + \cdots + (\mathbf{v}_t - \mathbf{u}_t) \ .$$

Applicando allora il lemma precedente deduciamo che

$$\mathbf{v}_1 - \mathbf{u}_1 = \mathbf{0} \ , \ \ldots \ , \ \mathbf{v}_t - \mathbf{u}_t = \mathbf{0} \ ,$$

e cioè $\mathbf{w}$ si esprime in unico modo come somma di vettori degli autospazi $V_{\lambda_1}, \ldots, V_{\lambda_t}$. $\square$

Il seguente enunciato esprime delle condizioni equivalenti per la diagonalizzabilità di un endomorfismo, ed è noto come il *Teorema Spettrale.*

5.29 TEOREMA SPETTRALE. *Sia V uno spazio vettoriale di dimensione n su un campo $\mathbb{K}$, sia $f: V \to V$ un endomorfismo e siano $\lambda_1, \dots, \lambda_m$ gli autovalori (a due a due distinti) di f. Le seguenti affermazioni sono equivalenti.*

(a) f *è diagonalizzabile;*

(b) $V = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_m}$*;*

(c) $\sum_{i=1}^m \dim V_{\lambda_i} = n$*;*

(d) $\sum_{i=1}^m m(\lambda_i) = n$ *ed inoltre* $\dim V_{\lambda_j} = m(\lambda_j)$ *per ogni* j*.*

Dimostrazione. Omessa. □

Poiché, come già osservato, fissata una base ordinata $\mathcal{B}$ e posto $A = M_{\mathcal{B}}^{\mathcal{B}}(f)$, si ha che $\dim V_{\lambda_i} = n - \rho(A - \lambda_i I_n)$, la (c) del teorema precedente si può anche scrivere $n = \sum_{i=1}^m \big(n - \rho(A - \lambda_i I_n)\big)$ e la (d) $\sum_{i=1}^m m(\lambda_i) = n$ ed inoltre $m(\lambda_j) = n - \rho(A - \lambda_j I_n) \quad (\forall\, j)$. Osserviamo che la condizione $\sum_{i=1}^m m(\lambda_i) = n$ equivale a dire che il polinomio caratteristico p è completamente riducibile, ovvero che p può esprimersi come prodotto di fattori lineari, o anche

$$p = (-1)^n (x - \lambda_1)^{m(\lambda_1)} \cdot \dots \cdot (x - \lambda_m)^{m(\lambda_m)} \; .$$

5.30 COROLLARIO DEL TEOREMA SPETTRALE. *Se $dimV = n$ e l'endomorfismo $f: V \to V$ ammette n autovalori distinti, allora f è diagonalizzabile.*

Dimostrazione. Siano $\lambda_1, \dots, \lambda_n$ gli autovalori di f. Poiché la dimensione di ogni autospazio è almeno 1, la somma delle dimensioni degli autospazi sarà almeno n. D'altra parte tale somma non può superare n, e quindi, per la (iii) del teorema precedente, f è diagonalizzabile. □

Il concetto di diagonalizzabilità può anche essere studiato considerando le matrici quadrate invece degli endomorfismi. Ricordiamo che se A è una matrice quadrata $n \times n$, è definito un endomorfismo $\omega_A : \mathbb{K}^n \longrightarrow \mathbb{K}^n$ che ammette A come matrice associata rispetto alla base canonica $\mathcal{B}$ di $\mathbb{K}^n$.

5.31 DEFINIZIONE. *Uno scalare λ (rispettivamente un vettore numerico W di $\mathbb{K}^n$) si dice autovalore (autovettore rispettivamente) di A se è un autovalore (autovettore rispettivamente) di ω_A.*

5.32 DEFINIZIONE. *La matrice A si dice diagonalizzabile se tale è ω_A.*

Abbiamo già osservato che l'endomorfismo ω_A è diagonalizzabile se e solo se esiste una base ordinata $\mathcal{D}$ tale che la matrice D associata ad ω_A rispetto a $\mathcal{D}$ sia diagonale. In tal caso, detta C la matrice di passaggio da $\mathcal{D}$ a $\mathcal{B}$, si avrà che C è invertibile e

$$D = C^{-1}AC \ . \tag{5.13}$$

Viceversa, se esistono una matrice invertibile C ed una matrice diagonale D tali che valga la (5.13), allora detta $\mathcal{D}$ la base di $\mathbb{K}^n$ tale che C sia la matrice di passaggio da $\mathcal{D}$ a $\mathcal{B}$, si ha che D è la matrice associata ad ω_A rispetto a $\mathcal{D}$ e ω_A risulta diagonalizzabile. Pertanto A è diagonalizzabile se e solo se esistono una matrice diagonale D e una martice invertibile C tali che valga la (5.13), ovvero se e solo se A è coniugata ad una matrice diagonale. In una tale situazione, si dice che C diagonalizza A. Osserviamo che le colonne di C (viste come vettori numerici colonna) costituiscono una base di autovettori di ω_A. Infatti, detti $\lambda_1, \dots, \lambda_n$ gli elementi (non necessariamente distinti) della diagonale di D, dalla (5.13) si deduce che $AC^j = \lambda_j C^j$, ovvero $\omega_A(C^j) = \lambda_j C_j$, per ogni j.

Concludiamo il capitolo con uno schema che riassume il modo di procedere per studiare la diagonalizzabilità di un endomorfismo $f : V \to V$. Sia $\mathcal{B}$ una base fissata e sia $A = M^{\mathcal{B}}_{\mathcal{B}}(f)$.

(i) Trovare gli autovalori di f, ovvero le radici $\lambda_1, \dots, \lambda_m$ del polinomio caratteristico $p = \det(A - xI_n)$. In altri termini, studiare l'equazione algebrica $\det(A - xI_n) = 0$;

(ii) Determinare le molteplicità algebriche $m(\lambda_1), \dots, m(\lambda_m)$ degli autovalori di f;

(iii) Sia $\sum_{i=1}^m m(\lambda_i) = t$. Se $t < n$ allora f non è diagonalizzabile. Se $t = n$, determinare la molteplicità geometrica $\dim V_{\lambda_i} = n - \rho(A - \lambda_i I_n)$ dell'autovalore λ_i. Se, per ogni i, essa coincide con $m(\lambda_i)$, allora f è diagonalizzabile.

Nel punto (iii), se $t = n$, basta verificare che $\dim V_{\lambda_i} = m(\lambda_i)$ solo per gli autovalori multipli di f. Infatti, se un autovalore λ è semplice, deve necessariamente accadere che $\dim V_\lambda = m(\lambda) = 1$. Nel caso f sia diagonalizzabile, se occorre determinare una base $\widetilde{\mathcal{B}}$ di autovettori di f, si deve determinare una base $\mathcal{B}_{\lambda_i}$ di V_{λ_i} per ogni i e porre

$$\widetilde{\mathcal{B}} = \mathcal{B}_{\lambda_1} \cup \dots \cup \mathcal{B}_{\lambda_m} \ .$$

In generale, fissata una qualunque base $\mathcal{B}$, e posto $A = M^{\mathcal{B}}_{\mathcal{B}}(f)$, se λ è un autovalore e $\mathbf{w} \in V$, detto W il vettore delle componenti di $\mathbf{w}$ in $\mathcal{B}$, avremo che $\mathbf{w} \in V_\lambda$ se e solo se W è soluzione del sistema lineare omogeneo

$$(A - \lambda I_n)X = \mathbf{0} \ . \tag{5.14}$$

Pertanto il sistema (5.14) è una rappresentazione cartesiana del sottospazio V_λ in $\mathcal{B}$.

Capitolo 6

Spazi vettoriali euclidei

In questo capitolo le dimostrazioni sono omesse.

6.1 Forme bilineari e prodotti scalari

Siano V, V', W degli spazi vettoriali su reali.

6.1 Definizione. *Un'applicazione $f : V \times V' \to W$ si dice bilineare se per ogni $\alpha, \beta \in \mathbb{R}$, $\mathbf{u}, \mathbf{v} \in V$, $\mathbf{u}', \mathbf{v}' \in V'$ si ha che*

(i) $f(\alpha\mathbf{u} + \beta\mathbf{v}, \mathbf{v}') = \alpha f(\mathbf{u}, \mathbf{v}') + \beta f(\mathbf{v}, \mathbf{v}')$;

(ii) $f(\mathbf{v}, \alpha\mathbf{u}' + \beta\mathbf{v}') = \alpha f(\mathbf{v}, \mathbf{u}') + \beta f(\mathbf{v}, \mathbf{v}')$.

Le (i) e (ii) si dicono proprietà di linearità sulla prima e sulla seconda componente rispettivamente. Osserviamo che se f è un'applicazione bilineare allora $f(\mathbf{0}, \mathbf{v}') = \mathbf{0} = f(\mathbf{v}, \mathbf{0})$, per ogni $\mathbf{v} \in V$, $\mathbf{v}' \in V'$, come il lettore potrà facilmente verificare. Noi siamo interessati al caso particolare in cui $V = V'$ e $W = \mathbb{R}$.

6.2 Definizione. *Si dice forma bilineare su V un'applicazione bilineare*

$$s : V \times V \longrightarrow \mathbb{R} \ . \tag{6.1}$$

6.3 Definizione. *La forma bilineare* (6.1) *si dice*

- *simmetrica se* $s(\mathbf{u}, \mathbf{v}) = s(\mathbf{v}, \mathbf{u}) \ \forall \ \mathbf{u}, \mathbf{v} \in V$,
- *antisimmetrica, o anche alternante, se* $s(\mathbf{u}, \mathbf{v}) = -s(\mathbf{v}, \mathbf{u}) \ \forall \ \mathbf{u}, \mathbf{v} \in V$.

Una forma bilineare simmetrica si dice anche *prodotto scalare* su V. Forniremo ora alcuni esempi.

6.4 Esempio. *Sia* $\dim V = 2$ *e sia* $\mathcal{B} = (\mathbf{e}_1, \mathbf{e}_2)$ *una base ordinata di* V. *Per ogni* $\mathbf{w}, \mathbf{z} \in V$, *detti* (w_1, w_2), (z_1, z_2) *i vettori delle componenti di* $\mathbf{w}, \mathbf{z}$ *in* $\mathcal{B}$, *poniamo*

$$s(\mathbf{w}, \mathbf{z}) = w_2 z_2 \ .$$

6.5 Esempio. *Sia* $\dim V = 2$ *e sia* $\mathcal{B} = (\mathbf{e}_1, \mathbf{e}_2)$ *una base ordinata di* V. *Per ogni* $\mathbf{w}, \mathbf{z} \in V$, *detti* (w_1, w_2), (z_1, z_2) *i vettori delle componenti di* $\mathbf{w}, \mathbf{z}$ *in* $\mathcal{B}$, *poniamo*

$$s(\mathbf{w}, \mathbf{z}) = w_2 z_1 + w_1 z_2 + w_2 z_2 \ .$$

6.6 Esempio. *Sia* $\dim V = 2$ *e sia* $\mathcal{B} = (\mathbf{e}_1, \mathbf{e}_2)$ *una base ordinata di* V. *Per ogni* $\mathbf{w}, \mathbf{z} \in V$, *detti* (w_1, w_2), (z_1, z_2) *i vettori delle componenti di* $\mathbf{w}, \mathbf{z}$ *in* $\mathcal{B}$, *poniamo*

$$s(\mathbf{w}, \mathbf{z}) = w_2 z_1 + w_2 z_2 \ .$$

6.7 Esempio. *Sia* $V = \mathbb{R}^n$. *Se* $X, Y \in \mathbb{R}^n$ *e si ha che*

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad ; \quad Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

poniamo

$$s(X, Y) = \sum_{i=1}^{n} x_i y_i \ .$$

6.8 Esempio. *Sia* $V = \mathbb{R}^3$ *e consideriamo la matrice*

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

Se X, Y *sono vettori numerici colonna di ordine tre, ovvero del tipo*

$$X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \ ; \ Y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$$

poniamo

$$s(X, Y) = ({}^t X) A Y = x_1 y_1 + x_3 y_1 + x_2 y_2 + x_1 y_3 \ .$$

6.9 Esempio. *Sia* $V = \mathbb{R}^3$ *e consideriamo la matrice*

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

Se X, Y *sono vettori numerici colonna di ordine tre, ovvero del tipo*

$$X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \ ; \ Y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$$

poniamo

$$s(X, Y) = ({}^t X) B Y = x_1 y_1 + x_3 y_1 + x_2 y_2 + x_1 y_3 + x_3 y_3 \ .$$

6.10 Esempio. *Sia* $V = \mathbb{R}^2$ *e definiamo* s *ponendo*

$$s\big((x_1, x_2), (y_1, y_2) \big) = x_1 y_1 - x_2 y_2 \ .$$

Il lettore potrà verificare per esercizio che s è bilineare in tutti gli esempi precedenti. Inoltre s è simmetrica in tutti gli esempi tranne il 6.6. Il prodotto scalare dell'Esempio 6.7 è noto come prodotto scalare standard in $\mathbb{R}^n$.

6.11 DEFINIZIONE. *Un prodotto scalare $s : V \times V \to \mathbb{R}$ si dice non degenere se*

$$s(\mathbf{u}, \mathbf{v}) = 0 \quad \forall\ \mathbf{u} \in V \ \Rightarrow\ \mathbf{v} = \mathbf{0} \tag{6.2}$$

o equivalentemente

$$\forall\ \mathbf{v} \neq \mathbf{0}\ \exists\ \mathbf{u} \in V \ \big|\ s(\mathbf{u}, \mathbf{v}) \neq 0\ . \tag{6.3}$$

Diremo che s è un prodotto scalare degenere quando ciò non accade, ovvero

$$\exists\ \mathbf{v} \neq \mathbf{0} \ \big|\ s(\mathbf{u}, \mathbf{v}) = 0 \quad \forall\ \mathbf{u} \in V\ . \tag{6.4}$$

I prodotti scalari degli Esempi 6.5, 6.7, 6.8, 6.10 sono non degeneri. Sono invece degeneri quelli degli Esempi 6.4, 6.9. Vedremo in seguito come verificare se un prodotto scalare è non degenere. Comunque osserviamo che nell'Esempio 6.4 per ogni $\mathbf{z} \in V$ si ha che

$$s(\mathbf{z}, \mathbf{e}_1) = 0$$

ma $\mathbf{e}_1 \neq \mathbf{0}$ (e quindi s è degenere). Analogamente, se nell'Esempio 6.9 consideriamo il vettore $Y = \left(\begin{smallmatrix} 1 \\ 0 \\ -1 \end{smallmatrix}\right)$, abbiamo che $s(X, Y) = 0$ per ogni $X \in \mathbb{R}^3$.

6.12 DEFINIZIONE. *Sia s un prodotto scalare in V. Un vettore non nullo $\mathbf{u} \in V$ si dice isotropo se $s(\mathbf{u}, \mathbf{u}) = 0$.*

Osserviamo che se non ci sono vettori isotropi allora s è non degenere. Altrimenti esisterebbe infatti un vettore non nullo $\mathbf{v}$ tale che $s(\mathbf{u}, \mathbf{v}) = 0$ per ogni $\mathbf{u} \in V$. In particolare si avrebbe quindi che $s(\mathbf{v}, \mathbf{v}) = 0$, e cioè $\mathbf{v}$ sarebbe un vettore isotropo. Viceversa, se s è non degenere possono esistere vettori isotropi. Ad esempio, se si considera il vettore $(1, -1) \in \mathbb{R}^2$ nell'Esempio 6.10, si ha che $s\big((1, -1), (1, -1)\big) = 0$.

Se in V è stata fissata una base, è possibile costruire una matrice a partire da una forma bilineare al modo seguente. Consideriamo la forma bilineare (6.1) e sia $\mathcal{B} = (\mathbf{e}_1, \dots, \mathbf{e}_n)$ una base ordinata di V. Definiamo una matrice $A = (a_{i,j}) \in \mathcal{M}_{n,n}(\mathbb{R})$ ponendo $a_{i,j} = s(\mathbf{e}_i, \mathbf{e}_j)$. Si dice che A è la matrice associata ad s rispetto a $\mathcal{B}$. Se $\mathbf{u}, \mathbf{v} \in V$ e

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad ; \quad Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

sono i vettori delle componenti di $\mathbf{u}, \mathbf{v}$ in $\mathcal{B}$, abbiamo che

$$\begin{aligned} s(\mathbf{u}, \mathbf{v}) &= s\Big(\sum_{i=1}^n x_i \mathbf{e}_i, \sum_{j=1}^n y_j \mathbf{e}_j\Big) \\ &= \sum_{i,j=1}^n x_i y_j s(\mathbf{e}_i, \mathbf{e}_j) \\ &= \sum_{i,j=1}^n x_i y_j a_{i,j} \\ &= ({}^tX)AY \end{aligned}$$

Viceversa, fissati V e $\mathcal{B}$ come sopra, e considerata una matrice $A \in \mathcal{M}_{n,n}(\mathbb{R})$, se per ogni coppia di vettori $(\mathbf{u}, \mathbf{v}) \in V \times V$ si pone $s(\mathbf{u}, \mathbf{v}) = ({}^tX)AY$ dove X, Y sono i vettori numerici colonna delle componenti di $\mathbf{u}, \mathbf{v}$ in $\mathcal{B}$, si verifica che l'applicazione s così definita è bilineare (usando la distributività e le altre proprietà del prodotto righe per colonne tra matrici). Ad esempio, le matrici associate alle forme bilineari degli Esempi 6.4, 6.5, 6.6 rispetto alla base $\mathcal{B} = (\mathbf{e}_1, \mathbf{e}_2)$, e alle forme bilineari degli Esempi 6.7, 6.8, 6.9, 6.10 rispetto alle basi canoniche, sono rispettivamente

$$\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \; ; \; \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \; ; \; \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} \; ; \; I_n$$

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

6.13 Esempio. Decomposizione di una forma bilineare. *Sia V uno spazio vettoriale reale e indichiamo con* $\mathrm{Bil}(V)$, $\mathrm{Sym}(V)$ *e* $\mathrm{Alt}(V)$ *rispettivamente l'insieme delle forme bilineari, simmetriche e alternanti su V. Possiamo dare a* $\mathrm{Bil}(V)$ *una struttura di spazio vettoriale introducendo una operazione interna (addizione) ed una esterna con operatori in* $\mathbb{R}$ *(moltiplicazione esterna) al modo seguente. Siano $f, g \in \mathrm{Bil}(V)$, $\lambda \in \mathbb{R}$ e definiamo due applicazioni $h, k : V \times V \to \mathbb{R}$ ponendo*

$$h(\mathbf{u}, \mathbf{v}) = f(\mathbf{u}, \mathbf{v}) + g(\mathbf{u}, \mathbf{v}) \quad ; \quad k(\mathbf{u}, \mathbf{v}) = \lambda f(\mathbf{u}, \mathbf{v}) .$$

Si verifica agevolmente che $h, k \in \mathrm{Bil}(V)$ e si pone

$$f + g = h \quad ; \quad \lambda \cdot f = k .$$

Il lettore potrà verificare che la struttura $(\mathrm{Bil}(V); +, \cdot)$ *è uno spazio vettoriale, che i sottoinsiemi* $\mathrm{Sym}(V)$ *e* $\mathrm{Alt}(V)$ *sono sottospazi di* $\mathrm{Bil}(V)$ *e che* $\mathrm{Sym}(V) \cap \mathrm{Alt}(V)$ *è il sottospazio banale. Se $f \in \mathrm{Bil}(V)$ definiamo $f_1, f_2 \in \mathrm{Bil}(V)$ ponendo*

$$f_1(\mathbf{u}, \mathbf{v}) = \frac{1}{2}\big(f(\mathbf{u}, \mathbf{v}) + f(\mathbf{v}, \mathbf{u})\big) \;\; ; \;\; f_2(\mathbf{u}, \mathbf{v}) = \frac{1}{2}\big(f(\mathbf{u}, \mathbf{v}) - f(\mathbf{v}, \mathbf{u})\big) .$$

Si verifica che $f_1 \in \mathrm{Sym}(V)$, e $f_2 \in \mathrm{Alt}(V)$. È poi evidente che $f = f_1 + f_2$. Pertanto

$$\mathrm{Bil}(V) = \mathrm{Sym}(V) \oplus \mathrm{Alt}(V) .$$

Definiamo due applicazioni $\omega_1 : \mathrm{Bil}(V) \to \mathrm{Sym}(V)$, $\omega_2 : \mathrm{Bil}(V) \to \mathrm{Alt}(V)$ ponendo $\omega_i(f) = f_i$ ($i = 1, 2$). Tali applicazioni sono epimorfismi e si dicono rispettivamente operatore di simmetrizzazione e di antisimmetrizzazione. Osserviamo che $\ker \omega_1 = \mathrm{Alt}(V) = \mathrm{im}\, \omega_2$ *e* $\ker \omega_2 = \mathrm{Sym}(V) = \mathrm{im}\, \omega_1$.

6.14 PROPOSIZIONE. *Sia V uno spazio vettoriale reale e sia $\mathcal{B} = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$ una sua base ordinata, s una forma bilineare su V e A la matrice associata ad s rispetto a $\mathcal{B}$. Allora s è un prodotto scalare se e solo se A è simmetrica.*

La seguente proposizione ci consente di riconoscere se un prodotto scalare è degenere.

6.15 PROPOSIZIONE. *Sia V uno spazio vettoriale reale e sia $\mathcal{B} = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$ una sua base ordinata, s un prodotto scalare su V e A la matrice associata ad s rispetto a $\mathcal{B}$. Allora s è non degenere se e solo se A è non degenere.*

Studiamo ora la relazione che intercorre tra due matrici associate ad uno stesso prodotto scalare rispetto a basi distinte. Sia dunque V uno spazio vettoriale dotato di un prodotto scalare s e siano $\mathcal{B} = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$, $\tilde{\mathcal{B}} = (\tilde{\mathbf{e}}_1, \ldots, \tilde{\mathbf{e}}_n)$ due basi ordinate di V. Siano inoltre $A, \tilde{A}$ le matrici associate ad s rispetto a tali basi, e poniamo $B = M^{\mathcal{B}}_{\tilde{\mathcal{B}}}$, la matrice di passaggio da $\tilde{\mathcal{B}}$ a $\mathcal{B}$. Per ogni $\mathbf{u}, \mathbf{v} \in V$, indicati con X, Y i vettori numerici colonna delle componenti di $\mathbf{u}, \mathbf{v}$ in $\mathcal{B}$ e con $\tilde{X}, \tilde{Y}$ i vettori numerici colonna delle componenti di $\mathbf{u}, \mathbf{v}$ in $\tilde{\mathcal{B}}$, si ha che

$$\begin{aligned} s(\mathbf{u}, \mathbf{v}) &= ({}^t\tilde{X})\tilde{A}\tilde{Y} \\ &= ({}^tX)AY \\ &= \big({}^t(B\tilde{X})\big)A(B\tilde{Y}) \\ &= ({}^t\tilde{X})({}^tB)AB\tilde{Y} \ . \end{aligned}$$

Tale relazione sussiste per ogni $\mathbf{u}, \mathbf{v} \in V$, cioè

$$({}^t\tilde{X})\tilde{A}\tilde{Y} = ({}^t\tilde{X})({}^tB)AB\tilde{Y} \qquad \forall\, \tilde{X}, \tilde{Y} \in \mathbb{R}^n$$

e da ciò si deduce che

$$\tilde{A} = ({}^tB)AB \ .$$

Quando in V è definito un prodotto scalare s, si può introdurre la nozione di ortogonalità in V (rispetto ad s).

6.16 DEFINIZIONE. *Siano $\mathbf{u}, \mathbf{v} \in V$. Diremo che $\mathbf{u}$ e $\mathbf{v}$ sono ortogonali, e scriveremo $\mathbf{u} \perp \mathbf{v}$, se $s(\mathbf{u}, \mathbf{v}) = 0$. Se poi $S \subseteq V$ diremo che $\mathbf{u}$ è ortogonale ad S, o anche che $\mathbf{u}$ è normale ad S, e scriveremo $\mathbf{u} \perp S$ se $\mathbf{u} \perp \mathbf{v}$ per ogni $\mathbf{v} \in S$.*

Se invece due vettori $\mathbf{u}, \mathbf{v}$ sono proporzionali, ovvero esiste uno scalare λ tale che $\mathbf{v} = \lambda\mathbf{u}$ oppure $\mathbf{u} = \lambda\mathbf{v}$ (in simboli $\mathbf{u} \propto \mathbf{v}$), si dice talvolta che $\mathbf{u}$ e $\mathbf{v}$ sono paralleli, e si scrive $\mathbf{u} \parallel \mathbf{v}$.

6.17 Definizione. *Sia $S \subseteq V$ una parte non vuota di V. Definiamo un altro sottoinsieme $S^\perp$ di V ponendo*

$$S^\perp = \{\, \mathbf{u} \in V \mid \mathbf{u} \perp S \,\} .$$

poniamo poi, per convenzione, $\emptyset^\perp = V$.

6.18 Proposizione. *Per ogni parte S di V si ha che $S^\perp$ è un sottospazio di V.*

6.19 Definizione. *Sia $W \leq V$. Il sottospazio $W^\perp \leq V$ si dice complemento ortogonale di W in V (rispetto ad s).*

Osserviamo che se $\mathbf{z} \in W \cap W^\perp$, $\mathbf{z} \neq \mathbf{0}$, allora, in particolare, $\mathbf{z} \perp \mathbf{z}$, cioè $s(\mathbf{z}, \mathbf{z}) = 0$ e $\mathbf{z}$ è isotropo. Pertanto, se non esistono vettori isotropi si ha che $W \cap W^\perp = \{\mathbf{0}\}$ e il sottospazio congiungente $W + W^\perp$ è una somma diretta.

6.20 Proposizione. *Sia $W \leq V$ e sia $\mathcal{B} = [\mathbf{w}_1, \dots, \mathbf{w}_r]$ una base di W. Allora un vettore $\mathbf{u}$ è ortogonale a W se e solo se $\mathbf{u} \perp \mathbf{w}_i$ per ogni $i = 1, \dots, r$.*

6.2 Spazi vettoriali euclidei

Restringiamo ora ulteriormente la nostra attenzione a particolari prodotti scalari non degeneri.

6.21 Definizione. *Sia V uno spazio vettoriale su $\mathbb{R}$ e sia s un prodotto scalare in V. Si dice che s è definito positivo se $s(\mathbf{u}, \mathbf{u}) > 0$, per ogni $\mathbf{u} \neq \mathbf{0}$.*

Il lettore potrà agevolmente verificare che se s è definito positivo allora s è anche non degenere, e non esistono vettori isotropi in V rispetto ad s. In particolare quindi, per ogni sottospazio W di V si ha che $W \cap W^\perp = \{\mathbf{0}\}$.

6.22 Definizione. *Uno spazio vettoriale euclideo è una coppia (V, s), dove V è uno spazio vettoriale su $\mathbb{R}$ ed s è un prodotto scalare definito positivo su V.*

D'ora in avanti considereremo solo spazi vettoriali euclidei finitamente generati, e scriveremo semplicemente V in luogo di (V, s). Il prodotto scalare tra due vettori $\mathbf{u}, \mathbf{v}$ in uno spazio vettoriale euclideo viene indicato, oltre che con il simbolo $s(\mathbf{u}, \mathbf{v})$, anche con i simboli $\langle \mathbf{u}, \mathbf{v} \rangle$ o $\mathbf{u} \cdot \mathbf{v}$.

6.23 Teorema (disuguaglianza di Schwarz). *Per ogni* $\mathbf{u}, \mathbf{v} \in V$ *si ha che*

$$(\mathbf{u} \cdot \mathbf{v})^2 \leq (\mathbf{u} \cdot \mathbf{u})(\mathbf{v} \cdot \mathbf{v}) \ . \tag{6.5}$$

Se poi i vettori $\mathbf{u}, \mathbf{v}$ *sono indipendenti, allora*

$$(\mathbf{u} \cdot \mathbf{v})^2 < (\mathbf{u} \cdot \mathbf{u})(\mathbf{v} \cdot \mathbf{v}) \ . \tag{6.6}$$

6.24 Definizione. *Per ogni* $\mathbf{u} \in V$ *poniamo* $\|\mathbf{u}\| = \mathbf{u} \cdot \mathbf{u}$, $|\mathbf{u}| = \sqrt{\mathbf{u} \cdot \mathbf{u}}$. *Gli scalari* $\|\mathbf{u}\|$, $|\mathbf{u}|$ *si dicono rispettivamente norma e lunghezza (o anche modulo) di* $\mathbf{u}$.

La (6.5) può quindi anche scriversi come $(\mathbf{u} \cdot \mathbf{v})^2 \leq \|\mathbf{u}\| \cdot \|\mathbf{v}\|$ o anche

$$|\mathbf{u} \cdot \mathbf{v}| \leq |\mathbf{u}| \cdot |\mathbf{v}| \ . \tag{6.7}$$

In questa ultima formulazione della disuguaglianza di Schwarz si è usato il simbolo $| \ |$ sia per indicare la lunghezza di un vettore che per indicare il valore assoluto di un numero reale.

6.25 Teorema. *Per ogni* $\mathbf{u}, \mathbf{v} \in V$ *e per ogni* $\lambda \in \mathbb{R}$ *si ha che*

(a) $|\mathbf{u}| \geq 0$*;*

(b) $|\mathbf{u}| = 0 \iff \mathbf{u} = \mathbf{0}$*;*

(c) $|\lambda \mathbf{u}| = |\lambda| \, |\mathbf{u}|$*;*

(d) $|\mathbf{u} + \mathbf{v}| \leq |\mathbf{u}| + |\mathbf{v}|$.

Siano ora dati due vettori $\mathbf{u}, \mathbf{v}$. Vogliamo introdurre la nozione di angolo (non orientato) $\theta = \widehat{\mathbf{uv}}$ tra $\mathbf{u}$ e $\mathbf{v}$. Se $\mathbf{u} = \mathbf{0}$ oppure $\mathbf{v} = \mathbf{0}$ poniamo $\theta = 0$. Se invece $\mathbf{u}, \mathbf{v}$ sono entrambi vettori non nulli, θ è l'unico scalare, nell'intervallo $[0, \pi]$, tale che

$$\cos\theta = \frac{\mathbf{u} \cdot \mathbf{v}}{|\mathbf{u}| \, |\mathbf{v}|} \in [-1, 1] \ . \tag{6.8}$$

La (6.8) ha senso, come si deduce dalla disuguaglianza di Schwarz nella formulazione (6.7). Osserviamo che $\mathbf{u} \cdot \mathbf{v} = |\mathbf{u}| \, |\mathbf{v}| \cos\theta$.

Consideriamo ora un sistema $\mathcal{S} = [\mathbf{v}_1, \dots, \mathbf{v}_m]$ di vettori di uno spazio euclideo V.

6.26 Definizione. *Il sistema* $\mathcal{S}$ *si dice ortogonale se* $\mathbf{v}_i \cdot \mathbf{v}_j = 0$, *per ogni coppia* (i, j) *di indici distinti. Diremo invece che* $\mathcal{S}$ *è ortonormale se* $\mathbf{v}_i \cdot \mathbf{v}_j = \delta_{i,j}$, *per ogni coppia* (i, j) *di indici.*

6.27 ESEMPIO. *In $\mathbb{R}^2$ con il prodotto scalare standard, il sistema $\mathcal{S} = [(1,1),(1,-1)]$ è ortogonale, ma non ortonormale. Il sistema $\mathcal{S}' = \left[\,(1/\sqrt{2},1/\sqrt{2}),\left(1/\sqrt{2},-1/\sqrt{2}\right)\,\right]$ è invece ortonormale. In $\mathbb{R}^3$ con il prodotto scalare standard, per ogni $t \in \mathbb{R}$ il sistema $\mathcal{S} = [(-1,\operatorname{sen} t,\cos t),(1,\operatorname{sen} t,\cos t)]$ è ortogonale, ma non ortonormale. Il sistema*

$$\mathcal{S}' = \left[\left(-\frac{1}{\sqrt{2}},\frac{\operatorname{sen} t}{\sqrt{2}},\frac{\cos t}{\sqrt{2}}\right),\left(\frac{1}{\sqrt{2}},\frac{\operatorname{sen} t}{\sqrt{2}},\frac{\cos t}{\sqrt{2}}\right)\right]$$

è invece ortonormale.

6.28 PROPOSIZIONE. *Sia $\mathcal{S} = [\mathbf{v}_1,\dots,\mathbf{v}_m]$ un sistema ortogonale di vettori non nulli di uno spazio vettoriale euclideo V. Allora $\mathcal{S}$ è indipendente.*

6.29 PROPOSIZIONE. *Sia $\mathcal{S} = [\mathbf{v}_1,\dots,\mathbf{v}_m]$ un sistema ortonormale di vettori di uno spazio vettoriale euclideo V, e supponiamo che il vettore $\mathbf{u} \in V$ dipenda da $\mathcal{S}$. Allora*

$$\mathbf{u} = \sum_{i=1}^{m}(\mathbf{u}\cdot\mathbf{v}_i)\mathbf{v}_i \ .$$

6.30 COROLLARIO. *Sia $\mathcal{B} = (\mathbf{v}_1,\dots,\mathbf{v}_n)$ una base ordinata ortonormale dello spazio vettoriale euclideo V. Allora per ogni $\mathbf{u} \in V$ si ha che*

$$\mathbf{u} = \sum_{i=1}^{n}(\mathbf{u}\cdot\mathbf{v}_i)\mathbf{v}_i \ .$$

In generale, dati due vettori $\mathbf{u},\mathbf{w} \in V$, con $\mathbf{w} \neq \mathbf{0}$, il vettore $\frac{\mathbf{u}\cdot\mathbf{w}}{\|\mathbf{w}\|}\mathbf{w}$ si dice *proiezione ortogonale* di $\mathbf{u}$ su $\mathbf{w}$ e lo scalare $\frac{\mathbf{u}\cdot\mathbf{w}}{\|\mathbf{w}\|}$ prende il nome di *coefficiente di Fourier* di $\mathbf{u}$ rispetto a $\mathbf{w}$. Il corollario precedente ci dice quindi che le componenti di un vettore $\mathbf{u}$ in una base ortonormale sono proprio i coefficienti di Fourier di $\mathbf{u}$ rispetto ai vettori di tale base. Osserviamo che $\mathbf{u}$ può scriversi come

$$\mathbf{u} = \frac{\mathbf{u}\cdot\mathbf{w}}{\|\mathbf{w}\|}\mathbf{w} + \left(\mathbf{u} - \frac{\mathbf{u}\cdot\mathbf{w}}{\|\mathbf{w}\|}\mathbf{w}\right) \ . \tag{6.9}$$

Il primo addendo della (6.9) è proporzionale a $\mathbf{w}$. Il secondo addendo invece è ortogonale a $\mathbf{w}$. Infatti

$$\frac{\mathbf{u}\cdot\mathbf{w}}{\|\mathbf{w}\|}\mathbf{w}\cdot\left(\mathbf{u} - \frac{\mathbf{u}\cdot\mathbf{w}}{\|\mathbf{w}\|}\mathbf{w}\right) = \frac{\mathbf{u}\cdot\mathbf{w}}{\|\mathbf{w}\|}(\mathbf{w}\cdot\mathbf{u}) - \frac{\mathbf{u}\cdot\mathbf{w}}{\|\mathbf{w}\|}\frac{\mathbf{u}\cdot\mathbf{w}}{\|\mathbf{w}\|}(\mathbf{w}\cdot\mathbf{w}) = 0 \ .$$

Pertanto la (6.9) si dice *decomposizione ortogonale* di $\mathbf{u}$ rispetto a $\mathbf{w}$. Si vede facilmente che una decomposizione siffatta è unica. Infatti, se

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

con $\mathbf{w}',\mathbf{w}''$ proporzionali a $\mathbf{w}$ e $\mathbf{z}',\mathbf{z}''$ ortogonali a $\mathbf{w}$, allora

$$\mathbf{w}' - \mathbf{w}'' + \mathbf{z}' - \mathbf{z}'' = \mathbf{0} \ .$$

Poiché $\mathbf{w}' - \mathbf{w}'' \perp \mathbf{z}' - \mathbf{z}''$, se fosse $\mathbf{w}' \neq \mathbf{w}''$ (e quindi anche $\mathbf{z}' \neq \mathbf{z}''$), avremmo che la somma di due vettori non nulli ortogonali (e quindi indipendenti) sarebbe banale, e ciò è assurdo.

6.31 ESEMPIO. *Sia $A \in GL_n(\mathbb{R})$. Consideriamo la spazio vettoriale euclideo $\mathbb{R}^n$ (con il prodotto scalare standard). Ricordiamo che $A \in O_n(\mathbb{R})$ (ovvero A è una matrice ortogonale) quando $A^{-1} = A_t$. Abbiamo che $A \in O_n(\mathbb{R})$ se e solo se le sue righe costituiscono una base ortonormale di $\mathbb{R}^n$. Infatti, se $A \in O_n(\mathbb{R})$ si ha che*

$$A \cdot A_t = A \cdot A^{-1} = I_n = (\delta_{i,j}) \ .$$

Quindi $A_i \cdot (A_t)^j = \delta_{i,j}$. Ma $(A_t)^j = A_j$ e dunque il sistema delle righe di A è ortonormale, e quindi è anche indipendente ed è una base, poiché ha ordine n. Il viceversa è analogo. Un simile ragionamento può essere fatto anche con le colonne di A.

6.3 Il procedimento di Gram-Schmidt

Vogliamo ora studiare un metodo per costruire una base ortonormale a partire da una base ordinata arbitraria. Tale metodo prende il nome di *procedimento di ortonormalizzazione di Gram–Schmidt*. Sia dunque $\mathcal{B} = (\mathbf{u}_1, \ldots, \mathbf{u}_n)$ una base ordinata dello spazio vettoriale euclideo V. Costruiamo una base ortogonale $\mathcal{B}' = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$ al modo seguente. Poniamo $\mathbf{v}_1 = \mathbf{u}_1$. Se $n = 1$, abbiamo già costruito una base ortogonale di V. Se invece $n > 1$, supponiamo induttivamente di aver già costruito un sistema ortogonale $\mathbf{v}_1, \ldots, \mathbf{v}_m$ ($m \geq 1$, $m < n$) in modo tale che

$$\mathcal{L}(\mathbf{u}_1, \ldots, \mathbf{u}_m) = \mathcal{L}(\mathbf{v}_1, \ldots, \mathbf{v}_m) \ .$$

Poniamo

$$\mathbf{v}_{m+1} = \mathbf{u}_{m+1} - \sum_{i=1}^{m} \frac{\mathbf{u}_{m+1} \cdot \mathbf{v}_i}{\|\mathbf{v}_i\|} \mathbf{v}_i \ .$$

Il sistema $\mathcal{S} = [\mathbf{v}_1, \ldots, \mathbf{v}_m, \mathbf{v}_{m+1}]$ è ancora ortogonale. Infatti, per ogni $j = 1, \ldots, m$

$$\begin{aligned}
\mathbf{v}_{m+1} \cdot \mathbf{v}_j &= \left(\mathbf{u}_{m+1} - \sum_{i=1}^{m} \frac{\mathbf{u}_{m+1} \cdot \mathbf{v}_i}{\|\mathbf{v}_i\|} \mathbf{v}_i \right) \cdot \mathbf{v}_j \\
&= \mathbf{u}_{m+1} \cdot \mathbf{v}_j - \sum_{i=1}^{m} \frac{\mathbf{u}_{m+1} \cdot \mathbf{v}_i}{\|\mathbf{v}_i\|} (\mathbf{v}_i \cdot \mathbf{v}_j) \\
&= \mathbf{u}_{m+1} \cdot \mathbf{v}_j - \frac{\mathbf{u}_{m+1} \cdot \mathbf{v}_j}{\|\mathbf{v}_j\|} (\mathbf{v}_j \cdot \mathbf{v}_j) \\
&= 0
\end{aligned}$$

D'altra parte è chiaro che $\mathbf{v}_{m+1} \in \mathcal{L}(\mathbf{u}_1, \ldots, \mathbf{u}_{m+1})$ e quindi

$$\mathcal{L}(\mathbf{v}_1, \ldots, \mathbf{v}_{m+1}) \leq \mathcal{L}(\mathbf{u}_1, \ldots, \mathbf{u}_{m+1}) \ .$$

Inoltre $\mathcal{S}$ è indipendente e quindi è una base di $\mathcal{L}(\mathbf{v}_1, \ldots, \mathbf{v}_{m+1})$. Pertanto

$$\dim \big(\mathcal{L}(\mathbf{v}_1, \ldots, \mathbf{v}_{m+1})\big) = m + 1 = \dim \big(\mathcal{L}(\mathbf{u}_1, \ldots, \mathbf{u}_{m+1})\big)$$

e dunque

$$\mathcal{L}(\mathbf{v}_1, \ldots, \mathbf{v}_{m+1}) = \mathcal{L}(\mathbf{u}_1, \ldots, \mathbf{u}_{m+1}) \; .$$

Ciò conclude il procedimento induttivo. Possiamo concludere che è determinato in tal modo un sistema ortogonale $\mathcal{B}' = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$ che, essendo indipendente, è una base (ortogonale) di V. Osserviamo che la costruzione effettuata è tale che

$$\mathcal{L}(\mathbf{v}_1, \ldots, \mathbf{v}_i) = \mathcal{L}(\mathbf{u}_1, \ldots, \mathbf{u}_i) \qquad \forall\ i \; .$$

In termini più espliciti, la base $\mathcal{B}'$si ottiene ponendo

$$\begin{aligned}
\mathbf{v}_1 &= \mathbf{u}_1 \; ; \\
\mathbf{v}_2 &= \mathbf{u}_2 - \frac{\mathbf{u}_2 \cdot \mathbf{v}_1}{\|\mathbf{v}_1\|}\, \mathbf{v}_1 \; ; \\
\mathbf{v}_3 &= \mathbf{u}_3 - \frac{\mathbf{u}_3 \cdot \mathbf{v}_1}{\|\mathbf{v}_1\|}\, \mathbf{v}_1 - \frac{\mathbf{u}_3 \cdot \mathbf{v}_2}{\|\mathbf{v}_2\|}\, \mathbf{v}_2 \; ; \\
&\vdots \\
\mathbf{v}_n &= \mathbf{u}_n - \frac{\mathbf{u}_n \cdot \mathbf{v}_1}{\|\mathbf{v}_1\|}\, \mathbf{v}_1 - \cdots - \frac{\mathbf{u}_n \cdot \mathbf{v}_{n-1}}{\|\mathbf{v}_{n-1}\|}\, \mathbf{v}_{n-1} \; .
\end{aligned}$$

Poniamo infine

$$\mathbf{w}_i = \frac{\mathbf{v}_i}{|\mathbf{v}_i|} \qquad \forall\ i \; .$$

È chiaro che $\mathcal{B}'' = (\mathbf{w}_1, \ldots, \mathbf{w}_n)$ è una base ortonormale di V, come il lettore potrà facilmente verificare.

6.32 Esempio. *Consideriamo la base ordinata $\mathcal{B} = (\mathbf{u}_1, \mathbf{u}_2)$ di $\mathbb{R}^2$, con $\mathbf{u}_1 = (1,1)$, $\mathbf{u}_2 = (1,2)$. Poniamo*

$$\mathbf{v}_1 = \mathbf{u}_1 = (1,1) \; ;$$

$$\mathbf{v}_2 = \mathbf{u}_2 - \frac{\mathbf{u}_2 \cdot \mathbf{v}_1}{\|\mathbf{v}_1\|}\, \mathbf{v}_1 = (1,2) - \frac{(1,2)\cdot(1,1)}{\|(1,1)\|}\,(1,1) = \Big(-\frac{1}{2}, \frac{1}{2} \Big) \; .$$

Abbiamo che $\mathcal{B}' = (\mathbf{v}_1, \mathbf{v}_2)$ è una base ortogonale (ma non ortonormale) di $\mathbb{R}^2$. Poniamo dunque

$$\mathbf{w}_1 = \frac{\mathbf{v}_1}{|\mathbf{v}_1|} = \frac{1}{\sqrt{2}}(1,1) = \Big(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \Big) \; ;$$

$$\mathbf{w}_2 = \frac{\mathbf{v}_2}{|\mathbf{v}_2|} = \sqrt{2}\Big(-\frac{1}{2}, \frac{1}{2} \Big) = \Big(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \Big)$$

e otteniamo la base ortonormale $\mathcal{B}'' = (\mathbf{w}_1, \mathbf{w}_2)$.

6.33 Esempio. *Consideriamo la base ordinata $\mathcal{B} = (\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3)$ di $\mathbb{R}^3$, con*

$$\mathbf{u}_1 = (1,0,1) \; ; \quad \mathbf{u}_2 = (1,1,0) \; ; \quad \mathbf{u}_3 = (0,0,1) \; .$$

Poniamo

$$
\begin{aligned}
\mathbf{v}_1 &= \mathbf{u}_1 = (1,0,1)\ ;\\
\mathbf{v}_2 &= \mathbf{u}_2 - \frac{\mathbf{u}_2\cdot\mathbf{v}_1}{\|\mathbf{v}_1\|}\,\mathbf{v}_1\\
&= (1,1,0) - \frac{(1,1,0)\cdot(1,0,1)}{\|(1,0,1)\|}(1,0,1)\\
&= (1,1,0) - \frac{1}{2}(1,0,1)\\
&= \left(\frac{1}{2},1,-\frac{1}{2}\right)\ ;\\
\mathbf{v}_3 &= \mathbf{u}_3 - \left(\frac{\mathbf{u}_3\cdot\mathbf{v}_1}{\|\mathbf{v}_1\|}\,\mathbf{v}_1 + \frac{\mathbf{u}_3\cdot\mathbf{v}_2}{\|\mathbf{v}_2\|}\,\mathbf{v}_2\right)\\
&= (0,0,1) - \left(\frac{(0,0,1)\cdot(1,0,1)}{\|(1,0,1)\|}(1,0,1)\right.\\
&\qquad \left. + \frac{(0,0,1)\cdot(1/2,1,-1/2)}{\|(1/2,1,-1/2)\|}(1/2,1,-1/2)\right)\\
&= (0,0,1) - \left(\frac{1}{2}(1,0,1) + \frac{-1/2}{3/2}\left(\frac{1}{2},1,-\frac{1}{2}\right)\right)\\
&= \left(-\frac{1}{3},\frac{1}{3},\frac{1}{3}\right)\ .
\end{aligned}
$$

Si verifica agevolmente che $\mathcal{B}'$ è una base ortogonale, ma non ortonormale. Si pone poi

$$
\begin{aligned}
\mathbf{w}_1 &= \frac{\mathbf{v}_1}{|\mathbf{v}_1|} = \frac{1}{\sqrt{2}}(1,0,1) = \left(\frac{1}{\sqrt{2}},0,\frac{1}{\sqrt{2}}\right)\ ;\\
\mathbf{w}_2 &= \frac{\mathbf{v}_2}{|\mathbf{v}_2|} = \sqrt{\frac{2}{3}}\left(\frac{1}{2},1,-\frac{1}{2}\right) = \left(\frac{1}{\sqrt{6}},\frac{\sqrt{2}}{\sqrt{3}},-\frac{1}{\sqrt{6}}\right)\ ;\\
\mathbf{w}_3 &= \frac{\mathbf{v}_3}{|\mathbf{v}_3|} = \sqrt{3}\left(-\frac{1}{3},\frac{1}{3},\frac{1}{3}\right) = \left(-\frac{1}{\sqrt{3}},\frac{1}{\sqrt{3}},\frac{1}{\sqrt{3}}\right)
\end{aligned}
$$

e si ottiene una base ortonormale $\mathcal{B}'' = (\mathbf{w}_1,\mathbf{w}_2,\mathbf{w}_3)$ di $\mathbb{R}^3$.

Poiché già sappiamo che ogni spazio vettoriale finitamente generato possiede basi, dal procedimento di ortonormalizzazione di Gram–Schmidt deduciamo che ogni spazio vettoriale euclideo finitamente generato possiede basi ortonormali.

6.34 Teorema. *Sia W un sottospazio dello spazio vettoriale euclideo V e sia $\mathbf{u} \in V$. Esistono allora, e sono unici, dei vettori $\mathbf{w} \in W$ e $\mathbf{z} \in W^\perp$ tali che $\mathbf{u} = \mathbf{w} + \mathbf{z}$.*

6.35 Corollario. *Sia W un sottospazio dello spazio vettoriale euclideo V. Si ha allora che $V = W \oplus W^\perp$ e quindi* $\dim V = \dim W + \dim W^\perp$.

In particolare si ha che $\dim W^\perp = \dim V - \dim W$.

6.36 Corollario. *Sia W un sottospazio dello spazio vettoriale euclideo V. Si ha allora che $\left(W^\perp\right)^\perp = W$.*

6.37 Definizione. *Siano $W_1, W_2 \leq V$ due sottospazi tali che $W_1 \not\subseteq W_2$ e $W_2 \not\subseteq W_1$. Diremo che W_1 è ortogonale a W_2, e scriveremo $W_1 \perp W_2$, se W_1 contiene il complemento ortogonale di W_2 in $W_1 + W_2$.*

Il lettore potrà verificare per esercizio che tale relazione è simmetrica.

6.38 Esempio. Ortogonalità tra sottospazi. *Consideriamo uno spazio euclideo V di dimensione n e due suoi sottospazi W_1, W_2 tali che $W_1 \not\subseteq W_2$ e $W_2 \not\subseteq W_1$.*

(i) *Sia* $n \geq 2$, $\dim W_1 = 1$, $\dim W_2 = n-1$. *Allora $W_1 + W_2 = V$ e quindi W_1 è ortogonale a W_2 se e solo se W_1 contiene $W_2^\perp$, ovvero, per motivi dimensionali, coincide con esso. In formule*
$$W_1 \perp W_2 \iff W_1 = W_2^\perp \ .$$

(ii) *Sia* $n \geq 2$, $\dim W_1 = \dim W_2 = 1$. *Allora* $\dim(W_1 + W_2) = 2$ *e quindi il complemento ortogonale di W' di W_2 in $W_1 + W_2$ ha dimensione 1. Pertanto $W_1 \perp W_2$ se e solo se $W_1 = W'$.*

(iii) *Sia* $\dim V > 2$ *e sia* $\dim W_1 = \dim W_2 = 2$. *Supponiamo inoltre che* $\dim(W_1 \cap W_2) = 1$ *(e quindi* $\dim(W_1 + W_2) = 3$*). Osserviamo che tale situazione si realizza, in particolare, nel caso in cui $n = 3$ (in tal caso $V = W_1 + W_2$). La dimensione del complemento ortogonale di W_2 in $W_1 + W_2$ è 1. Detti $\mathbf{n}_1, \mathbf{n}_2$ dei vettori di $W_1 + W_2$ ortogonali a W_1, W_2 rispettivamente, si ha che*
$$W_1 \perp W_2 \iff \mathbf{n}_2 \in W_1 \iff \mathbf{n}_1 \perp \mathbf{n}_2 \ .$$

Consideriamo una base ordinata ortonormale $\mathcal{B} = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$ dello spazio vettoriale euclideo V, sia $\mathcal{B}' = (\mathbf{e}'_1, \ldots, \mathbf{e}'_n)$ un'altra base ordinata e poniamo $B = M^{\mathcal{B}}_{\mathcal{B}'} = (b_{i,j})$.

6.39 Teorema. *La base $\mathcal{B}'$ è ortonormale se e solo se la matrice B è ortogonale.*

6.4 Diagonalizzazione ortogonale

Per concludere il capitolo, affrontiamo il problema della diagonalizzazione negli spazi vettoriali euclidei. D'ora in poi sia V uno spazio vettoriale euclideo di dimensione n, s il suo prodotto scalare ed $f : V \to V$ un endomorfismo. Fissata una base ordinata $\mathcal{B} = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$ di V possiamo considerare due matrici: la matrice M associata ad s e la matrice $A = M^{\mathcal{B}}_{\mathcal{B}}(f)$ associata ad f, entrambe rispetto a $\mathcal{B}$. Si verifica agevolmente che M è diagonale se e solo se $\mathcal{B}$ è ortogonale, e si ha che $M = I_n$ se e solo se $\mathcal{B}$ è ortonormale. D'altra parte, già sappiamo che A è diagonale se e solo se $\mathcal{B}$ è costituita da autovettori. Ci si pone quindi il problema di cercare una base ortonormale di autovettori. Ovviamente non sempre esiste una tale base, e se esiste essa non è, in generale, unica.

6.40 Definizione. *L'endomorfismo f si dice ortogonalmente diagonalizzabile se esiste una base ortonormale di autovettori di f.*

6.41 Definizione. *L'endomorfismo f si dice simmetrico (o anche autoaggiunto) se si ha che $s\big(\mathbf{u}, f(\mathbf{v})\big) = s\big(f(\mathbf{u}), \mathbf{v}\big)$, per ogni $\mathbf{u}, \mathbf{v} \in V$.*

6.42 Proposizione. *Sia f un endomorfismo simmetrico e sia $\mathcal{B}$ ortonormale. Allora $A = M^{\mathcal{B}}_{\mathcal{B}}(f)$ è simmetrica.*

6.43 Proposizione. *Sia f un endomorfismo e sia $\mathcal{B}$ ortonormale. Se $A = M^{\mathcal{B}}_{\mathcal{B}}(f)$ è simmetrica, allora f è simmetrico.*

Pertanto f è simmetrico se e solo se è simmetrica la matrice ad esso associata rispetto ad una base ortonormale (e quindi anche rispetto ad ogni altra base ortonormale).

6.44 Teorema. *Sia f un endomorfismo simmetrico e sia $p \in \mathbb{R}[x]$ il suo polinomio caratteristico. Allora p è completamente riducibile, ovvero esistono degli scalari $\alpha_1, \ldots, \alpha_t$ e degli interi positivi $\kappa_1, \ldots, \kappa_t$ tali che*

$$p = (-1)^n (x - \alpha_1)^{\kappa_1} \cdot \cdots \cdot (x - \alpha_t)^{\kappa_t} \ . \tag{6.10}$$

Siamo ora in grado di caratterizzare gli endomorfismi ortogonalmente diagonalizzabili.

6.45 Teorema. *Sia $f : V \to V$ un endomorfismo. Allora f è ortogonalmente diagonalizzabile se e solo se f è simmetrico.*

Allo scopo di fornire un metodo per costruire una base ortonormale di autovettori dell'endomorfismo simmetrico f, premettiamo il seguente

6.46 Lemma. *Sia f un endomorfismo simmetrico di V. Siano λ, μ due autovalori distinti di f e $\mathbf{u}, \mathbf{v}$ degli autovettori associati a tali autovalori. Allora $\mathbf{u} \perp \mathbf{v}$.*

La costruzione di una base ortonormale $\mathcal{B}$ di autovettori dell'endomorfismo simmetrico f può essere effettuata come segue. f è ortogonalmente diagonalizzabile, e quindi anche diagonalizzabile. Se $\lambda_1, \ldots, \lambda_t$ sono gli autovalori di f avremo che

$$\dim V_{\lambda_1} + \cdots + \dim V_{\lambda_t} = n \ .$$

Possono dunque essere determinate delle basi $\mathcal{B}'_1, \ldots, \mathcal{B}'_t$ associate rispettivamente agli autospazi $V_{\lambda_1}, \ldots, V_{\lambda_t}$. Con il procedimento di Gram–Schmidt si ottengono poi delle basi ortonormali $\mathcal{B}_1, \ldots, \mathcal{B}_t$ di tali autospazi. Si pone poi

$$\mathcal{B} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_t \ .$$

Dalla costruzione, in base al lemma precedente, è agevole verificare che $\mathcal{B}$ è una base ortonormale di autovettori di f.

Concludiamo questo paragrafo con alcune osservazioni sulla diagonalizzazione ortogonale di matrici. Sia $A \in \mathcal{M}_{n,n}(\mathbb{R})$ e sia $\mathcal{B}$ la base canonica di $\mathbb{R}^n$. Se A è simmetrica, l'endomorfismo

$$\omega_A : \mathbb{R}^n \longrightarrow \mathbb{R}^n$$

è simmetrico e quindi anche ortogonalmente diagonalizzabile. Esiste pertanto una base spettrale ortonormale $\mathcal{B}'$ di $\mathbb{R}^n$ ed una matrice $P \in GL_n(\mathbb{R})$ tale che $P^{-1} \cdot A \cdot P = D$ risulta diagonale (e sulla diagonale di D compaiono gli autovalori di A). Le colonne di P sono i vettori numerici delle componenti degli autovettori di cui è costituita la base spettrale $\mathcal{B}'$ in $\mathcal{B}$, e cioè $P = M^{\mathcal{B}}_{\mathcal{B}'}$. Quindi $P \in O_n(\mathbb{R})$, essendo anche $\mathcal{B}$ ortonormale, e dunque si ha che $P^{-1} = P_t$ e

$$D = P_t \cdot A \cdot P \ .$$

6.5 Forme quadratiche

Sia V uno spazio vettoriale reale.

6.47 Definizione. *Un'applicazione $q : V \to \mathbb{R}$ si dice forma quadratica su V se esiste una forma bilineare $\tilde{q}$ su V tale che $q(\mathbf{v}) = \tilde{q}(\mathbf{v}, \mathbf{v})$, per ogni $\mathbf{v} \in V$.*

Diremo che q è la forma quadratica associata a $\tilde{q}$. È chiaro che q può essere associata a varie forme bilineari, mentre ogni forma bilineare ammette un'unica forma quadratica associata. In altri termini, indicato con $\mathrm{Q}(V)$ l'insieme delle forme quadratiche su V, l'applicazione $\omega : \mathrm{Bil}(V) \to \mathrm{Q}(V)$ che associa ad ogni forma bilineare la sua forma quadratica associata risulta suriettiva. Se $q \in \mathrm{Q}(V)$, definiamo un'applicazione $\bar{q} : V \times V \to \mathbb{R}$ ponendo

$$\bar{q}(\mathbf{u}, \mathbf{v}) = \frac{1}{2}\big(q(\mathbf{u} + \mathbf{v}) - q(\mathbf{u}) - q(\mathbf{v})\big) \ .$$

Il lettore potrà verificare che $\bar{q} \in \mathrm{Sym}(V) \leq \mathrm{Bil}(V)$ e che $\omega(\bar{q}) = q$, quindi ogni forma quadratica è associata ad (almeno) una forma bilineare simmetrica. Definiamo ora un'applicazione $\bar{\omega} : \mathrm{Q}(V) \to \mathrm{Bil}(V)$ ponendo $\bar{\omega}(q) = \bar{q}$. Si ha che $\mathrm{im}\, \bar{\omega} \subseteq \mathrm{Sym}(V)$. Inoltre $\omega \circ \bar{\omega} = id_{\mathrm{Q}(V)}$. Il lettore potrà verificare che $\bar{\omega} \circ \omega(f) = \omega_1(f)$ per ogni $f \in \mathrm{Bil}(V)$, dove ω_1 è l'operatore di simmetrizzazione dell'Esempio 6.13. La restrizione

$$(\bar{\omega} \circ \omega)_{|} : \mathrm{Sym}(V) \longrightarrow \mathrm{Sym}(V)$$

è un isomorfismo (l'isomorfismo identico) e quindi le restrizioni

$$\omega_| : \mathrm{Sym}(V) \longrightarrow \mathrm{Q}(V) \quad ; \quad \bar{\omega}_| : \mathrm{Q}(V) \longrightarrow \mathrm{Sym}(V)$$

sono l'una l'inversa dell'altra. L'insieme $\mathrm{Q}(V)$ ha una struttura vettoriale, e $\omega_|$ è un isomorfismo. Pertanto ogni forma quadratica è associata ad un'unica forma bilineare simmetrica.

Supponiamo ora che $\dim V = n$ e sia $\mathcal{B}''$ una base ordinata di V. Consideriamo una forma bilineare simmetrica f e la forma quadratica q ad essa associata. La matrice A associata ad f in $\mathcal{B}''$ si dice anche associata a q in $\mathcal{B}''$ (ed è simmetrica). Per ogni $\mathbf{u} \in V$, indicato con $X'' = \begin{pmatrix} x''_1 \\ \vdots \\ x''_n \end{pmatrix}$ il vettore coordinato di $\mathbf{u}$ rispetto a $\mathcal{B}''$, abbiamo che

$$q(\mathbf{u}) = f(\mathbf{u}, \mathbf{u}) = X''_t \cdot A \cdot X'' = \sum_{i,j} a_{i,j} x''_i x''_j \ .$$

Poiché A è simmetrica, esiste una matrice ortogonale $P \in O_n(\mathbb{R})$ tale che

$$D = P^{-1} \cdot A \cdot P = P_t \cdot A \cdot P = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}$$

dove $\lambda_1, \dots, \lambda_n$ sono gli autovalori di A. La matrice P può essere scelta in modo tale che nella n-pla $(\lambda_1, \dots, \lambda_n)$ compaiano prima gli (eventuali) scalari positivi, diciamo $\lambda_1, \dots, \lambda_s$, poi gli (eventuali) scalari negativi, diciamo $\lambda_{s+1}, \dots, \lambda_{s+t}$, ed infine, eventualmente, degli zeri. Osserviamo che $s + t = \rho(A) = \rho(D)$ (essendo P invertibile). Sappiamo che esiste un'unica base ordinata $\mathcal{B}'$ tale che P sia la matrice di passaggio da $\mathcal{B}'$ a $\mathcal{B}''$. Indicato con $X' = \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}$ il vettore coordinato di $\mathbf{u}$ in $\mathcal{B}'$, si ha che $X'' = PX'$ e quindi

$$\begin{aligned} q(\mathbf{u}) &= (P \cdot X')_t \cdot A \cdot P \cdot X' = X'_t \cdot P_t \cdot A \cdot P \cdot X' \\ &= X'_t \cdot (P_t \cdot A \cdot P) \cdot X' = X'_t \cdot D \cdot X' = \sum_{h=1}^{n} \lambda_h (x'_h)^2 \ . \end{aligned}$$

Consideriamo ora la matrice diagonale B definita come segue

$$b_{h,h} = \begin{cases} \lambda_h^{-\frac{1}{2}} & \text{se } 1 \le h \le s \\ (-\lambda_h)^{-\frac{1}{2}} & \text{se } s < h \le s + t \\ 1 & \text{se } s + t < h \end{cases}$$

Posto $C = B_t D B$ si vede facilmente che

$$C = \left(\begin{array}{c|c|c} I_s & O & O \\ \hline O & -I_t & O \\ \hline O & O & O \end{array}\right)$$

Sia $\mathcal{B}$ la base ordinata tale che la matrice invertibile B sia la matrice di passaggio da $\mathcal{B}$ a $\mathcal{B}'$. Indicato con $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ il vettore coordinato di $\mathbf{u}$ in $\mathcal{B}$ si ha che $X' = B \cdot X$ e quindi

$$\begin{aligned} q(\mathbf{u}) &= X'_t \cdot D \cdot X' = (B \cdot X)_t \cdot D \cdot B \cdot X = X_t \cdot B_t \cdot D \cdot B \cdot X = X_t \cdot C \cdot X \\ &= x_1^2 + \cdots + x_s^2 - \left(x_{s+1}^2 + \cdots + x_{s+t}^2\right) . \end{aligned}$$

Abbiamo quindi determinato una base ordinata $\mathcal{B}$ tale che, con le solite notazioni

$$q(\mathbf{u}) = x_1^2 + \cdots + x_s^2 - \left(x_{s+1}^2 + \cdots + x_{s+t}^2\right) . \tag{6.11}$$

Supponiamo che $\tilde{\mathcal{B}}$ sia una base ordinata e che, per ogni $\mathbf{u} \in V$, detto

$$Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

il vettore coordinato di $\mathbf{u}$ in $\tilde{\mathcal{B}}$, si abbia

$$q(\mathbf{u}) = y_1^2 + \cdots + y_p^2 - \left(y_{p+1}^2 + \cdots + y_{p+r}^2\right) . \tag{6.12}$$

Se $\tilde{A}$ è la matrice di q in $\tilde{\mathcal{B}}$ sappiamo che $\rho(A) = \rho(\tilde{A})$ e quindi $s + t = p + r$. Vogliamo provare che $s = p$ (e quindi anche $t = r$). Sia, per assurdo $p < s$. Sia W il sottospazio di V rappresentato in $\mathcal{B}$ (in forma cartesiana) dal sistema

$$\begin{cases} x_{s+1} = 0 \\ \quad\vdots \\ x_n = 0 \end{cases}$$

e W' il sottospazio di V rappresentato in $\tilde{\mathcal{B}}$ (in forma cartesiana) dal sistema

$$\begin{cases} x_1 = 0 \\ \quad\vdots \\ x_p = 0 \end{cases}$$

Abbiamo che $\dim W = s$, $\dim W' = n - p$ e $V = W + W'$. Quindi, per la formula di Grassmann,

$$\dim(W \cap W') = \dim W + \dim W' - \dim V = s - p > 0 .$$

Esiste pertanto un vettore non nullo $\mathbf{w} \in W \cap W'$. Poiché $\mathbf{w} \in W$, le sue componenti non banali in $\mathcal{B}$ sono tra le prime s e quindi, in base alla (6.11), $q(\mathbf{w}) > 0$. D'altra parte, essendo $\mathbf{w} \in W'$, le sue componenti non banali in $\tilde{\mathcal{B}}$ sono tra le ultime $n - p$ e quindi, per la (6.12), $q(\mathbf{w}) \leq 0$, e questo è assurdo. In definitiva, gli interi non negativi s, t sono caratteristici di q; si dice allora che q è in *forma canonica* in $\mathcal{B}$ e la (6.11) si dice forma canonica di q. La coppia (s, t) si dice *segnatura* di q. A posteriori vediamo quindi che la segnatura di q si ricava, fissato un qualunque riferimento, considerando la matrice A di q in tale riferimento e contando gli autovalori positivi e quelli negativi di A.

Capitolo 7

Spazi affini

7.1 Generalità su spazi e sottospazi affini

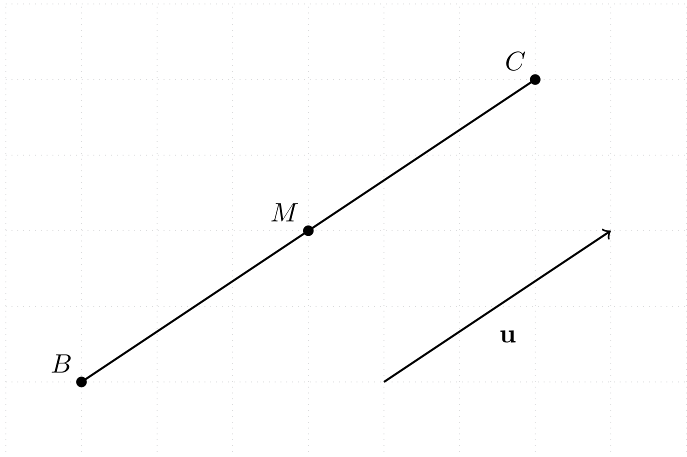
Sia $\mathbb{K}$ un campo, V un $\mathbb{K}$-spazio vettoriale e $\mathbb{A}$ un insieme non vuoto. Sia inoltre $\pi : \mathbb{A} \times \mathbb{A} \longrightarrow V$ un'applicazione.

7.1 Definizione. *La terna $(\mathbb{A}, V, \pi)$ si dice spazio affine sul campo $\mathbb{K}$, se*

(i) Per ogni $P \in \mathbb{A}$, $\mathbf{u} \in V$ esiste un unico $Q \in \mathbb{A}$ tale che $\pi(P, Q) = \mathbf{u}$;

(ii) Per ogni $P, Q, R \in \mathbb{A}$, si ha che $\pi(P, Q) + \pi(Q, R) = \pi(P, R)$.

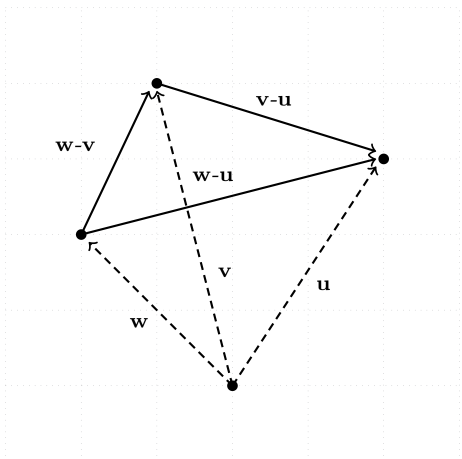
L'insieme $\mathbb{A}$ è il sostegno di tale struttura e, con abuso di notazione, si scriverà spesso $\mathbb{A}$ in luogo di $(\mathbb{A}, V, \pi)$. L'assioma (ii) è noto come *relazione di Chasles*. È d'uso comune indicare con il simbolo $\overrightarrow{PQ}$ il vettore $\pi(P, Q)$. Gli elementi di $\mathbb{A}$ si dicono *punti*, quelli di V *vettori liberi*. L'assioma (i) implica che l'applicazione π è suriettiva, ed in generale, per ogni $\mathbf{u} \in V$ la controimmagine $\pi^{-1}(\mathbf{u})$, che è un sottoinsieme di coppie in $\mathbb{A} \times \mathbb{A}$ si dice anch'esso vettore libero. Una coppia $(P, Q) \in \mathbb{A} \times \mathbb{A}$, si dice talvolta segmento orientato di estremi P, Q, o anche *vettore applicato*, con punto di applicazione P e punto finale Q. Per ogni punto $P \in \mathbb{A}$, l'applicazione indotta $\pi_P : \mathbb{A} \to V$ definita ponendo $\pi_P(Q) = \overrightarrow{PQ}$ è una biezione ed induce su $\mathbb{A}$ una struttura vettoriale isomorfa a quella di V, che prende il nome di spazio vettoriale dei vettori applicati in P. Se lo spazio V è finitamente generato, come sarà implicitamente assunto nel seguito, si pone $\dim \mathbb{A} = \dim V$. Dati tre punti $B, C, M \in \mathbb{A}$, diremo che M è il punto medio di B, C (o anche del segmento di estremi B, C), se $\overrightarrow{BM} = \overrightarrow{MC}$. Si verifica agevolmente, nel caso in cui il campo $\mathbb{K}$ sia ad esempio $\mathbb{R}$, $\mathbb{Q}$ 0 $\mathbb{C}$, che, comunque scelti B, C, esiste un unico punto M con tale proprietà.



$$\mathbf{u} = \pi(B, M) = \pi(M, C)$$

7.2 Esempio. Piano e spazio ordinari. *Sia $\mathbb{A}$ il piano (o anche lo spazio) della geometria elementare e sia V lo spazio vettoriale dei vettori liberi del piano (dello spazio rispettivamente). Definiamo $\pi : \mathbb{A} \times \mathbb{A} \to V$ ponendo $\pi(P, Q) = \overrightarrow{PQ}$ (il vettore libero rappresentato dal segmento orientato (P, Q)). È facile verificare che sono soddisfatte le proprietà assiomatiche (i), (ii).*

7.3 Esempio. Struttura affine naturale di uno spazio vettoriale. *Sia V uno spazio vettoriale. Definiamo $\pi : V \times V \to V$ ponendo $\pi(\mathbf{u}, \mathbf{v}) = \mathbf{v} - \mathbf{u}$, per ogni $\mathbf{u}, \mathbf{v} \in V$.*



$$\mathbf{w} - \mathbf{u} = (\mathbf{v} - \mathbf{u}) + (\mathbf{w} - \mathbf{v})$$

ovvero

$$\pi(\mathbf{u}, \mathbf{w}) = \pi(\mathbf{u}, \mathbf{v}) + \pi(\mathbf{v}, \mathbf{w})$$

Anche in questo caso è immediata la verifica delle proprietà assiomatiche (i), (ii).

Sia ora $\mathbb{B} \subseteq \mathbb{A}$ e sia $U = \pi(\mathbb{B} \times \mathbb{B})$. In altri termini, poniamo

$$U = \left\{ \mathbf{u} \in V \;\middle|\; \exists\ P, Q \in \mathbb{B} \;\middle|\; \pi(P,Q) = \mathbf{u} \right\} .$$

Consideriamo la restrizione $\bar{\pi} : \mathbb{B} \times \mathbb{B} \to U$ dell'applicazione π.

7.4 Definizione. *Se $U \leq V$ e la terna $(\mathbb{B}, U, \bar{\pi})$ è uno spazio affine, si dice che $\mathbb{B}$ è un sottospazio affine di $\mathbb{A}$; il sottospazio vettoriale U di V si dice sottospazio direttore di $\mathbb{B}$ e si indica anche con il simbolo $\overrightarrow{\mathbb{B}}$.*

7.5 Osservazione. *Sia $\mathbb{B}$ un sottospazio affine dello spazio affine $\mathbb{A}$.*

(i) *Fissato un qualunque punto $P \in \mathbb{B}$, l'applicazione indotta $\bar{\pi}_P$ è una biezione tra $\mathbb{B}$ e $\overrightarrow{\mathbb{B}}$. In altri termini*

$$\mathbb{B} = \left\{ Q \in \mathbb{A} \;\middle|\; \overrightarrow{PQ} \in \overrightarrow{\mathbb{B}} \right\} .$$

Se $\overrightarrow{PQ} = \mathbf{u}$, si scrive anche $Q = P + \mathbf{u}$. Si tratta di un simbolo formale di addizione, giustificato in seguito da questioni di coordinate e componenti. In modo analogo, si scrive talvolta $\mathbb{B} = P + \overrightarrow{\mathbb{B}}$, ovvero $\mathbb{B} = P + U$.

(ii) *Per ogni punto $P \in \mathbb{A}$ e per ogni sottospazio vettoriale $U \leq V$, l'insieme*

$$P + U = \left\{ Q \in \mathbb{A} \;\middle|\; \exists\ \mathbf{u} \in U \;\middle|\; Q = P + \mathbf{u} \right\}$$

è un sottospazio affine di $\mathbb{A}$, ed U è il suo sottospazio direttore, ovvero $\overrightarrow{P+U} = U$.

(iii) *Se $\dim \mathbb{B} = 0$, $\mathbb{B}$ si riduce ad un unico punto ed il suo sottospazio direttore è lo spazio vettoriale banale.*

(iv) *Se $\dim \mathbb{B} = 1$, $\mathbb{B}$ si dice* retta (affine) *e U ed il suo sottospazio direttore U si dice direttrice della retta. Essendo U uno spazio vettoriale di dimensione 1, ogni suo elemento non nullo è un generatore e si dice vettore direzionale della retta.*

(v) *Se $\dim \mathbb{B} = 2$, $\mathbb{B}$ si dice* piano (affine) *ed il suo sottospazio direttore U si dice giacitura del piano.*

(vi) *Se $\dim \mathbb{B} = n - 1$, $\mathbb{B}$ si dice* iperpiano (affine).

7.6 Definizione. *Siano $\mathbb{B} = P + U$, $\mathbb{B}' = P' + U'$ due sottospazi affini di $\mathbb{A}$. Diremo che $\mathbb{B}$ e $\mathbb{B}'$ sono paralleli, e scriveremo $\mathcal{B}B \parallel \mathbb{B}'$, se $U \leq U'$ oppure $U' \leq U$. Se, in particolare, si ha che $\dim \mathbb{B} = \dim \mathbb{B}'$, allora $\mathbb{B} \parallel \mathbb{B}'$ se e solo se $U = U'$.*

7.7 Osservazione. *(i)* *La relazione di parallelismo tra sottospazi è simmetrica e riflessiva, ma non transitiva.*

(ii) *La relazione di parallelismo tra sottospazi di dimensione fissata è anche transitiva, e quindi è una equivalenza.*

Consideriamo ora un vettore $\mathbf{u} \in V$ ed un sottospazio affine $\mathbb{B} = P + U$ di $\mathbb{A}$. Sottolineiamo il fatto che è d'uso comune dire che $\mathbf{u}$ è parallelo al sottospazio $\mathbb{B}$

quando $\mathbf{u} \in U$. Scriveremo allora $\mathbf{u} \parallel \mathbb{B}$. Siano ora $(\mathbb{A}, A, \pi)$ e $(\mathbb{A}', V', \pi')$ due spazi affini sul campo $\mathbb{K}$, e sia $f : \mathbb{A} \to \mathbb{A}'$ una applicazione. Per ogni punto $P \in \mathbb{A}$, posto $P' = f(P) \in \mathbb{A}'$, f induce un'altra applicazione $\overrightarrow{f_P} : V \to V'$ definita come segue. Per ogni $\mathbf{u} \in V$ esiste un unico punto $Q \in \mathbb{A}$ tale che $\overrightarrow{PQ} = \mathbf{u}$. Si considera allora il punto $Q' = f(Q) \in \mathbb{A}'$ ed il vettore libero $\mathbf{u}' = \overrightarrow{P'Q'} \in V'$ e si pone $\overrightarrow{f_P}(\mathbf{u}) = \mathbf{u}'$.

7.8 Definizione. *Se $\overrightarrow{f_P}$ è lineare e non dipende dalla scelta del punto P, si dice che f è una applicazione, o trasformazione, affine. Se poi $\overrightarrow{f_P}$ è un isomorfismo, f è una affinità.*

7.9 Teorema. *Per ogni punto $P \in \mathbb{A}$, per ogni punto $P' \in \mathbb{A}'$ e per ogni applicazione lineare $\alpha : V \to V'$ esiste un'unica applicazione affine $f : \mathbb{A} \to \mathbb{A}'$ tale che $f(P) = P'$ e $\overrightarrow{f_P} = \alpha$.*

Dimostrazione. Omessa. □

7.10 Proposizione. *Sia $f : \mathbb{A} \to \mathbb{A}'$ una applicazione affine e sia $\mathbb{B}$ un sottospazio affine di $\mathbb{A}$. Allora $f(\mathbb{B})$ è anch'esso un sottospazio di $\mathbb{A}'$ e $\dim f(\mathbb{B}) \leq \dim \mathbb{B}$.*

Dimostrazione. Omessa. □

7.2 Riferimenti e coordinate

Sia $\mathbb{A}$ uno spazio affine di dimensione n, $O \in \mathbb{A}$, $\mathcal{B} = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$ una base ordinata di V.

7.11 Definizione. *La coppia $\mathcal{R} = (O; \mathcal{B})$, ovvero la $(n+1)$-pla $(O; \mathbf{e}_1, \ldots, \mathbf{e}_n)$, si dice riferimento (affine) di $\mathbb{A}$, di origine O, associato a $\mathcal{B}$.*

Ricordiamo che in relazione alla base $\mathcal{B}$ di V possiamo considerare l'isomorfismo coordinato $\Phi_{\mathcal{B}} : V \to \mathbb{K}^n$ definito ponendo, per ogni $\mathbf{u} \in V$,

$$\Phi_{\mathcal{B}}(\mathbf{u}) = X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

dove $x_1, \ldots, x_n$ sono le componenti del vettore $\mathbf{u}$ nella base $\mathcal{B}$. In altri termini, si ha che $\mathbf{u} = \sum_{i=1}^{n} x_i \mathbf{e}_i$.

In modo analogo definiamo una applicazione

$$c_{\mathcal{R}} : \mathbb{A} \longrightarrow \mathbb{K}^n$$

al modo seguente: Per ogni $P \in \mathbb{A}$ poniamo

$$c_{\mathcal{R}}(P) = \Phi_{\mathcal{B}}(\overrightarrow{OP}) \ .$$

7.12 OSSERVAZIONE. *1.* $c_{\mathcal{R}}(O) = \mathbf{0}$.

2. *Se* $c_{\mathcal{R}}(P) = X$, $c_{\mathcal{R}}(Q) = Y$, *allora* $\Phi_{\mathcal{B}}(\overrightarrow{PQ}) = Y - X$. *Per tale motivo si scrive talvolta* $Q - P$ *invece di* $\overrightarrow{PQ}$.
3. *Se si considera la struttura affine standard di* $\mathbb{K}^n$, *l'applicazione* $c_{\mathcal{R}}$ *è una affinità.*
4. *Dati due punti* B, C, *il loro punto medio ha coordinate* $c_{\mathcal{R}}(M) = \frac{1}{2}\big(c_{\mathcal{R}}(B) + c_{\mathcal{R}}(C)\big)$.

Se $c_{\mathcal{R}} = X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, scriveremo anche

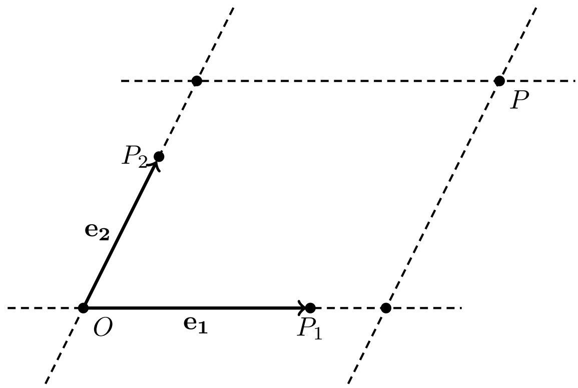
$$P \equiv_{\mathcal{R}} (x_1, \dots, x_n)$$

ed ometteremo l'indicazione del riferimento se ciò non darà adito ad equivoci. Il vettore X prende il nome di vettore coordinato del punto P in $\mathcal{R}$. Spesso si dirà semplicemente che il punto P ha coordinate $x_1, \dots, x_n$. Ad esempio, nell'osservazione precedente, il punto 4 può anche esprimersi dicendo che se $B \equiv_{\mathcal{R}} (b_1, \dots, b_n)$ e $C \equiv_{\mathcal{R}} (c_1, \dots, c_n)$ si ha che $M \equiv_{\mathcal{R}} \left(\frac{b_1+c_1}{2}, \dots, \frac{b_n+c_n}{2}\right)$.

Consideriamo ora $h{+}1$ punti $P_0, P_1, \dots, P_h$ ordinati di $\mathbb{A}$, ovvero la $(h{+}1)$-pla $(P_0, P_1, \dots, P_h) \in \mathbb{A}^{h+1}$.

7.13 DEFINIZIONE. *I punti* $P_0, P_1, \dots, P_h$ *sono indipendenti (in senso affine) se i corrispondenti vettori* $\overrightarrow{P_0P_1}, \dots, \overrightarrow{P_0P_h}$ *risultano indipendenti (in senso vettoriale).*

Osserviamo esplicitamente che, come è facile verificare, tale definizione non dipende dall'ordine in cui i punti compaiono. Se $\dim V = n$, possiamo trovare al più n vettori indipendenti in V, e quindi esisteranno al più $n+1$ punti indipendenti in $\mathbb{A}$. Una $(n+1)$-pla $\bar{\mathcal{R}} = (P_0, P_1, \dots, P_n)$ di punti indipendenti di $\mathbb{A}$ prende il nome di *riferimento baricentrico.* Posto $\mathbf{e}_i = \overrightarrow{P_0P_i}$ per ogni $i = 1, \dots, n$, i vettori $\mathbf{e}_1, \dots, \mathbf{e}_n$ costituiscono una base $\mathcal{B}$ di V e il riferimento affine $\mathcal{R} = (P_0; \mathcal{B})$ si dice associato al riferimento baricentrico $\bar{\mathcal{R}}$.



$$\mathcal{R} = \big(O; \mathbf{e_1}, \mathbf{e_2}\big)$$

$$\bar{\mathcal{R}} = \big(O, P_1, P_2\big)$$

$$P \equiv \left(\tfrac{4}{3}, \tfrac{3}{2}\right)$$

Siano ora $\mathcal{R} = (O; \mathcal{B})$, $\mathcal{R}' = (O'; \mathcal{B}')$ due riferimenti affini di $\mathbb{A}$. In relazione a tali riferimenti, possiamo considerare le affinità $c_{\mathcal{R}} : \mathbb{A} \to \mathbb{K}^n$ e $c_{\mathcal{R}'} : \mathbb{A} \to \mathbb{K}^n$. Pertanto, dato un punto $P \in \mathbb{A}$, poniamo $P \equiv_{\mathcal{R}} (x_1, \ldots, x_n)$, $P \equiv_{\mathcal{R}'} (x'_1, \ldots, x'_n)$ L'applicazione composta

$$\psi_{\mathcal{R},\mathcal{R}'} =: c_{\mathcal{R}'} \circ (c_{\mathcal{R}})^{-1} : \mathbb{K}^n \longrightarrow \mathbb{A} \longrightarrow \mathbb{K}^n$$

trasforma il vettore coordinato X di P in $\mathcal{R}$ nel vettore coordinato X' di P in $\mathcal{R}'$ e per tale motivo prende il nome di *cambiamento di riferimento* da $\mathcal{R}$ ad $\mathcal{R}'$.

7.14 Proposizione. *Nella situazione sopra descritta, si ha che $X' = AX + B$ dove $A = M_{\mathcal{B}}^{\mathcal{B}'}$ è la matrice del cambiamento delle basi da $\mathcal{B}$ a $\mathcal{B}'$ e $B = c_{\mathcal{R}'}(O)$ è il vettore coordinato dell'origine O del riferimento $\mathcal{R}$ nel riferimento $\mathcal{R}'$.*

Dimostrazione. Omessa. □

7.3 Rappresentazione di sottospazi affini

Sia $\mathbb{A}$ uno spazio affine di dimensione n sul campo $\mathbb{K}$, $\mathcal{R} = (O; \mathbf{e}_1, \ldots, \mathbf{e}_n)$ un suo riferimento e $\mathbb{B}$ un sottospazio affine di $\mathbb{A}$ di dimensione h. Consideriamo un sistema lineare compatibile

$$AX = B \tag{7.1}$$

dove A è una matrice di tipo $m \times n$ e supponiamo che, posto $A' = (A|B)$, si abbia $\rho(A) = \rho(A') = n - h$. Indichiamo con il simbolo Sol (7.1) il sottoinsieme di $\mathbb{K}^n$ delle soluzioni del sistema. Come è noto esso dipende da h parametri. Diremo che il sistema (7.1) è una rappresentazione cartesiana del sottospazio $\mathbb{B}$ rispetto al riferimento affine $\mathcal{R}$ se accade che

$$P \in \mathbb{B} \iff c_{\mathcal{R}}(P) \in \mathrm{Sol}\,(7.1)$$

ovvero se e solo se l'applicazione $c_{\mathcal{R}}$ si restringe ad una biezione

$$\bar{c}_{\mathcal{R}} : \mathbb{B} \longrightarrow \mathrm{Sol}\,(7.1)\ .$$

Vale il seguente enunciato.

7.15 PROPOSIZIONE. *Ogni sistema compatibile del tipo* (7.1) *rappresenta un sottospazio affine (di dimensione h). Inoltre il sistema omogeneo associato rappresenta il sottospazio direttore di tale sottospazio affine.*

Dimostrazione. Omessa. □

In generale possiamo rappresentare un sottospazio affine in due modi:

(i) mediante la costruzione di un sistema del tipo (7.1) che rappresenti il sottospazio (rappresentazione cartesiana);

(ii) mediante la descrizione dell'insieme delle soluzioni di un tale sistema, utilizzando h parametri (rappresentazione parametrica).

Vediamo come si costruiscono tali rappresentazioni. Il sottospazio $\mathbb{B}$ è individuato, come abbiamo visto, da un suo punto P_0 e dal suo sottospazio direttore $U = \overrightarrow{\mathbb{B}}$. Sia $\mathcal{B}' = (\mathbf{u}_1, \dots, \mathbf{u}_h)$ una base ordinata di U. In tale situazione la coppia $\mathcal{R}' = (P_0; \mathcal{B}')$ è un riferimento di $\mathbb{B}$. Un punto $P \in \mathbb{A}$ appartiene al sottospazio $\mathbb{B}$ se e solo se il vettore $\overrightarrow{P_0P}$ appartiene ad U, ovvero se e solo se $\overrightarrow{P_0P}$ dipende da $\mathcal{B}'$, e cioè esistono degli scalari $t_1, \dots, t_h$ tali che

$$\overrightarrow{P_0P} = \sum_{i=1}^{h} t_i \mathbf{e}_i \ . \tag{7.2}$$

Passando alle coordinate dei punti e dei vettori in questione, posto

$$c_{\mathcal{R}}(P_0) = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} \ ; \ \Phi_{\mathcal{B}}(\mathbf{u}_1) = \begin{pmatrix} u_{1,1} \\ \vdots \\ u_{n,1} \end{pmatrix} \ ; \ \dots \ ; \ \Phi_{\mathcal{B}}(\mathbf{u}_h) = \begin{pmatrix} u_{1,h} \\ \vdots \\ u_{n,h} \end{pmatrix}$$

ed indicate con $x_1, \dots, x_n$ le coordinate del punto generico P, abbiamo che

$$\Phi_{\mathcal{B}}(\overrightarrow{P_0P}) = \begin{pmatrix} x_1 - z_1 \\ \vdots \\ x_n - z_n \end{pmatrix}$$

e la relazione (7.2) equivale, mediante l'isomorfismo coordinato, alla relazione

$$\begin{pmatrix} x_1 - z_1 \\ \vdots \\ x_n - z_n \end{pmatrix} = t_1 \begin{pmatrix} u_{1,1} \\ \vdots \\ u_{n,1} \end{pmatrix} + \cdots + t_h \begin{pmatrix} u_{1,h} \\ \vdots \\ u_{n,h} \end{pmatrix}$$

Tale relazione tra vettori numerici equivale alle relazioni scalari

$$\begin{cases} x_1 - z_1 = t_1 u_{1,1} + \cdots + t_h u_{1,h} \\ \qquad\vdots \\ x_n - z_n = t_1 u_{n,1} + \cdots + t_h u_{n,h} \end{cases}$$

ovvero

$$\begin{cases} x_1 = z_1 + t_1 u_{1,1} + \cdots + t_h u_{1,h} \\ \quad\vdots \\ x_n = z_n + t_1 u_{n,1} + \cdots + t_h u_{n,h} \end{cases}$$

Si scrive allora

$$\mathbb{B} : \begin{cases} x_1 = z_1 + t_1 u_{1,1} + \cdots + t_h u_{1,h} \\ \quad\vdots \\ x_n = z_n + t_1 u_{n,1} + \cdots + t_h u_{n,h} \end{cases}$$

e si dice che questa è una rappresentazione parametrica del sottospazio $\mathbb{B}$ nel riferimento $\mathcal{R}$.

Torniamo ora al punto in cui si richiede che $\overrightarrow{P_0 P}$ dipenda da $\mathcal{B}'$, ovvero che il sistema $\mathcal{S} = \left[\mathbf{u}_1, \ldots, \mathbf{u}_h, \overrightarrow{P_0 P}\right]$ risulti dipendente. Ciò equivale a dire che sia dipendente il corrispondente sistema dei vettori numerici delle componenti dei vettori di $\mathcal{S}$, ovvero il sistema

$$\bar{\mathcal{S}} = \left[\begin{pmatrix} u_{1,1} \\ \vdots \\ u_{n,1} \end{pmatrix}, \ldots, \begin{pmatrix} u_{1,h} \\ \vdots \\ u_{n,h} \end{pmatrix}, \begin{pmatrix} x_1 - z_1 \\ \vdots \\ x_n - z_n \end{pmatrix} \right] .$$

Poiché già sappiamo che i primi h vettori di $\bar{\mathcal{S}}$ sono indipendenti, la dipendenza di $\bar{\mathcal{S}}$ equivale al fatto che la matrice

$$A = \begin{pmatrix} u_{1,1} & \cdots & u_{1,h} & z_1 - x_1 \\ \vdots & & \vdots & \vdots \\ u_{n,1} & \cdots & u_{n,h} & z_n - x_n \end{pmatrix}$$

(di tipo $n \times (h+1)$) abbia rango esattamente h. Poiché le prime h colonne di A sono indipendenti, in base al Teorema degli orlati possiamo trovare una sottomatrice fondamentale che coinvolge le prime h colonne, i cui orlati siano tutti nulli. Supponiamo, per semplicità, che il minore non nullo di A di ordine h sia individuato, oltre che dalle prime h colonne, anche dalle prime h righe, e cioè si abbia

$$\det A^{1,\ldots,h}_{1,\ldots,h} \neq 0 .$$

L'annullarsi di tutti i possibili orlati di tale minore descrive un sistema di $n-h$ equazioni nelle incognite $x_1, \ldots, x_n$, che è la rappresentazione cartesiana cercata del sottospazio $\mathbb{B}$:

$$\mathbb{B}: \begin{cases} \det A^{1,\ldots,h,h+1}_{1,\ldots,h,h+1} = 0 \\ \det A^{1,\ldots,h,h+1}_{1,\ldots,h,h+2} = 0 \\ \qquad\vdots \\ \det A^{1,\ldots,h,h+1}_{1,\ldots,h,n} = 0 \end{cases}$$

La costruzione della rappresentazione parametrica o cartesiana di un sottospazio $\mathbb{B}$ di dimensione h si può effettuare anche se il sottospazio è individuato da $h+1$ suoi punti indipendenti $P_0, P_1, \ldots, P_h$; basta infatti considerare i vettori $\mathbf{u}_1 = \overrightarrow{P_0P_1}\ ;\ \ldots\ ; \mathbf{u}_h = \overrightarrow{P_0P_h}$ e procedere come sopra.

7.4 Rette ed iperpiani

Sia $\mathbb{A}$ uno spazio affine di dimensione n sul campo $\mathbb{K}$, $\mathcal{R} = (O; \mathbf{e}_1, \ldots, \mathbf{e}_n)$ un suo riferimento e $\mathbb{B}$ una retta affine. Per comodità e per tradizione, indichiamo tale retta con il simbolo r. Allora la sua direttrice $\overrightarrow{r} = U$ è un sottospazio di V di dimensione 1. Un qualunque vettore non nullo $\mathbf{u} \in U$ genera quindi U. Un tale vettore si dice *vettore direzionale* di r, e le sue componenti (ordinate) $\ell_1, \ldots, \ell_n$ si dicono *numeri (o parametri) direttori* di r. Essi sono definiti a meno di proporzionalità.

Ripercorriamo ora il ragionamento fatto in generale per i sottospazi affini, adattandolo al caso in esame: se $P_0 \in r$ e $\mathbf{u}$ è un vettore non nullo di $\overrightarrow{r}$, un punto $P \in \mathbb{A}$ appartiene alla retta r se e solo se il vettore $\overrightarrow{P_0P}$ appartiene ad $\overrightarrow{r}$; poiché $\mathbf{u}$ genera $\overrightarrow{r}$, ciò vuol dire che esiste uno scalare $t \in \mathbb{K}$ tale che $\overrightarrow{P_0P} = t\mathbf{u}$. Passando alle componenti, posto

$$\Phi_{\mathcal{B}}(\mathbf{u}) = \begin{pmatrix} \ell_1 \\ \vdots \\ \ell_n \end{pmatrix} \quad ; \quad c_{\mathcal{R}}(P) = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad ; \quad c_{\mathcal{R}}(P_0) = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix}$$

abbiamo che

$$\Phi_{\mathcal{B}}(\overrightarrow{P_0P}) = \begin{pmatrix} x_1 - z_1 \\ \vdots \\ x_n - z_n \end{pmatrix}$$

e quindi la relazione (vettoriale) $\overrightarrow{P_0P} = t\mathbf{u}$ equivale alle relazioni scalari

$$\begin{cases} x_1 = z_1 + t\ell_1 \\ \quad\vdots \\ x_n = z_n + t\ell_n \end{cases} \tag{7.3}$$

che prendono il nome di *rappresentazione paremetrica* della retta affine r. Se $\bar{t} \in \mathbb{K}$, indicheremo con $P_{\bar{t}}$ il punto di r relativo al valore $\bar{t}$ del parametro, ovvero

$$P_{\bar{t}} = (z_1 + \bar{t}\ell_1, \ldots, z_n\bar{t}\ell_n)\ .$$

Per ottenere una *rappresentazione cartesiana* (in cui non compaiono parametri), si può eliminare il parametro dalle (7.3). Altrimenti si procede come segue. Il fatto che il vettore $\overrightarrow{P_0P}$ appartenga ad $\overrightarrow{r}$ equivale ad imporre che il sistema $\mathcal{S} = [\mathbf{u}, \overrightarrow{P_0P}]$ sia dipendente, ovvero, passando alle componenti, che la matrice

$$A = \begin{pmatrix} x_1 - z_1 & \cdots & x_n - z_n \\ \ell_1 & \cdots & \ell_n \end{pmatrix}$$

abbia rango 1. Si sceglie allora una componente non nulla di $\mathbf{u}$ (certamente esistente in quanto $\mathbf{u} \neq \mathbf{0}$), sia essa ad esempio ℓ_1, e si impone che tutti gli orlati della sottomatrice (ℓ_1) in A siano nulli, cioè

$$\begin{vmatrix} x_1 - z_1 & x_2 - z_2 \\ \ell_1 & \ell_2 \end{vmatrix} = 0 \quad ; \quad \ldots \quad ; \quad \begin{vmatrix} x_1 - z_1 & x_n - z_n \\ \ell_1 & \ell_n \end{vmatrix} = 0 \; .$$

Si tratta, come si vede, di $n-1$ equazioni di primo grado nelle incognite $x_1, x_2, \ldots, x_n$.

Esaminiamo ora il caso degli iperpiani di uno spazio affine. Sia $\mathbb{B}$ un iperpiano, ovvero un sottospazio di dimensione $n-1$. Per comodità e per tradizione, indichiamo tale sottospazio con il simbolo π. Sia P_0 un punto di π. Allora $\pi = P_0 + \overrightarrow{\pi}$. Una base ordinata di $\overrightarrow{\pi}$ è del tipo

$$\mathcal{B}' = (\mathbf{u}_1, \ldots, \mathbf{u}_{n-1}) \; .$$

Ripercorrendo la costruzione della rappresentazione cartesiana di un sottospazio in questo caso particolare, otteniamo che gli orlati da studiare si riducono ad un solo orlato e la rappresentazione cartesiana dell'iperpiano π è costituita da un'unica equazione lineare, che sarà del tipo

$$\pi : a_1x_1 + \cdots + a_nx_n + a = 0$$

dove non tutti gli scalari $a_1, \ldots, a_n$ sono nulli.

Osserviamo esplicitamente che se $n = 2$ gli iperpiani coincidono con le rette, se $n = 3$ essi sono proprio i piani.

Siano ora π', π'' due iperpiani di rappresentazione cartesiana

$$\pi' : a'_1x_1 + \cdots + a'_nx_n + a' = 0 \quad ; \quad \pi'' : a''_1x_1 + \cdots + a''_nx_n + a'' = 0 \; .$$

7.16 Proposizione. *Siano π', π'' due iperpiani distinti. Se π', π'' non sono paralleli, l'intersezione $\pi' \cap \pi''$ è un sottospazio affine di dimensione $n-2$ con rappresentazione cartesiana*

$$\pi' \cap \pi'' : \begin{cases} a'_1x_1 + \cdots + a'_nx_n + a' = 0 \\ a''_1x_1 + \cdots + a''_nx_n + a'' = 0 \end{cases}$$

Dimostrazione. Omessa. $\square$

7.17 Definizione. *Siano π', π'' due iperpiani distinti. Se π', π'' non sono paralleli, l'insieme $\mathcal{F}'$ degli iperpiani che contengono $\pi' \cap \pi''$ si dice fascio (proprio) di iperpiani individuato da π', π''. Se invece $\pi' || \pi''$, e cioè $\overrightarrow{\pi}' = \overrightarrow{\pi}''$, l'insieme $\mathcal{F}''$ degli iperpiani paralleli a π' (nonché a π''), ovvero l'insieme degli iperpiani che ammettono $\overrightarrow{\pi}'$ come sottospazio direttore, si dice fascio improprio di iperpiani di giacitura $\overrightarrow{\pi}'$.*

Osserviamo che gli iperpiani che appartengono al fascio (proprio o improprio) individuato da due iperpiani distinti π', π'' sono tutti e soli quelli del tipo

$$\bar{\pi} : \lambda(a'_1 x_1 + \cdots + a'_n x_n + a') + \mu(a''_1 x_1 + \cdots + a''_n x_n + a'') = 0$$

dove $\lambda, \mu \in \mathbb{K}$ non sono entrambi nulli. Indichiamo con $\mathcal{F}$, o anche $\mathcal{F}_{\pi',\pi''}$, il fascio (proprio o improprio) individuato da π', π''.

Se π', π'' non sono paralleli, e quindi gli iperpiani di $\mathcal{F}$ sono tutti e soli quelli che contengono il sottospazio $\pi' \cap \pi''$, tale sottospazio si dice asse del fascio. Se invece π', π'' sono paralleli, gli iperpiani di $\mathcal{F}$ sono tutti e soli quelli paralleli a π' (e anche a π''), e inoltre il generico iperpiano $\bar{\pi}$ del fascio $\mathcal{F}$ ammette una rappresentazione cartesiana del tipo

$$\bar{\pi} : a'_1 x_1 + \cdots + a'_n x_n + k = 0$$

dove k è uno scalare.

7.5 Spazi affini euclidei

Sia $(V, \mathbb{A}, \pi)$ uno spazio affine reale (ovvero sia $\mathbb{K} = \mathbb{R}$). Se V è uno spazio vettoriale euclideo e, dati due suoi vettori $\mathbf{u}, \mathbf{v}$, indichiamo con $\mathbf{u} \cdot \mathbf{v}$ il loro prodotto scalare, diremo che anche $\mathbb{A}$ è uno spazio (affine) euclideo. Le nozioni di norma, lunghezza o modulo, angolo ed ortogonalità riguardanti i vettori di V si riportano ai punti ed i sottospazi di $\mathbb{A}$. Ad esempio, se $P, Q, R \in \mathbb{A}$, si definisce la distanza $d(P, Q)$ ponendo

$$d(P, Q) = \left|\overrightarrow{PQ}\right| ,$$

e l'angolo $\widehat{PQR}$ ponendo

$$\widehat{PQR} = \widehat{\overrightarrow{QP}\overrightarrow{QR}} .$$

Se poi $\mathbb{B}$ è un sottospazio affine di $\mathbb{A}$, si pone

$$d(P, \mathbb{B}) = \min \left\{ d(P, Q) \mid Q \in \mathbb{B} \right\}$$

e, se $\mathbb{B}'$, $\mathbb{B}''$ sono due sottospazi affini di $\mathbb{A}$, si pone

$$d(\mathbb{B}', \mathbb{B}'') = \min \left\{ d(P, Q) \mid P \in \mathbb{B}',\ Q \in \mathbb{B}'' \right\} ,$$

potendosi infatti dimostrare l'esistenza di tali minimi. Sia ora $\mathbf{u}$ un vettore. Diremo che $\mathbf{u}$ è ortogonale al sottospazio affine $\mathbb{B}$, e scriveremo $\mathbf{u} \perp \mathbb{B}$, se $\mathbf{u} \perp \overrightarrow{\mathbb{B}}$, ovvero $\mathbf{u} \perp \mathbf{w}$ per ogni $\mathbf{w} \in \overrightarrow{\mathbb{B}}$.

Consideriamo un iperpiano $\pi : a_1x_1+\cdots+a_nx_n+a = 0$. Si verifica facilmente che il vettore $\mathbf{n}_\pi = (a_1, \ldots, a_n)$ è ortogonale a π ($\mathbf{n}_\pi \perp \pi$) e che $V = \mathcal{L}(\mathbf{n}_\pi) \oplus \overrightarrow{\pi}$. Pertanto ogni vettore ortogonale a π è proporzionale a $\mathbf{n}_\pi$. Per tale motivo si usa dire che $\mathbf{n}_\pi$ è *il* vettore normale di π, e che è definito a meno di proporzionalità. Osserviamo inoltre che ogni vettore libero $\mathbf{u} \in V$ può esprimersi, in unico modo, come somma $\mathbf{u} = h\mathbf{n}_\pi + \mathbf{w}$, dove $\mathbf{w}$ è un opportuno vettore di $\overrightarrow{\pi}$ e h è un opportuno scalare. Si dice che $\mathbf{w}$ è la proiezione ortogonale di $\mathbf{u}$ su π (o più precisamente su $\overrightarrow{\pi}$). Il fascio improprio $\mathcal{F}''$ costituito dagli iperpiani paralleli a π si indica talvolta anche con il simbolo $\mathcal{F}''_{\mathbf{n}_\pi}$, in quanto può vedersi anche come l'insieme dei piani che ammettono $\mathbf{n}_\pi$ come vettore normale.

Ci occuperemo ora degli spazi euclidei di dimensione 2 e 3, ed il lettore potrà fare riferimento, per visualizzare intuitivamente le situazioni che descriveremo, al piano ed allo spazio della geometria elementare.

7.6 Il piano affine ed euclideo reale

Poniamo dunque $n = 2$. Osserviamo esplicitamente che il modello classico per il piano affine $(\mathbb{A}, V, \mathbb{R})$ con $\dim \mathbb{A} = 2$ è il piano euclideo della geometria elementare. In tale modello, gli elementi di $\mathbb{A}$ sono i punti del piano e V è lo spazio vettoriale dei vettori liberi ordinari del piano. Laddove occorre considerare questioni metriche, angolari o di ortogonalità, si assume che sia dato in V un prodotto scalare definito positivo. Indichiamo tale spazio affine euclideo con il simbolo $\mathbb{E}^2$. I sottospazi di $\mathbb{E}^2$ sono i singleton (o impropriamente i punti) che hanno dimensione 0, le rette, di dimensione 1, ed infine $\mathbb{E}^2$ stesso, di dimensione 2. Abbiamo già affrontato il problema di rappresentare, in forma cartesiana o parametrica, le rette in uno spazio affine di dimensione arbitraria. Vediamo cosa succede in dettaglio in $\mathbb{E}^2$. Come osservato in generale, una retta r ha direttrice $\overrightarrow{r} \leq V$ e per ogni $P_0 \in r$ si ha che

$$r = P_0 + \overrightarrow{r}$$

ovvero

$$r = \left\{ P \in \mathbb{E}^2 \mid \overrightarrow{P_0P} \in \overrightarrow{r} \right\} .$$

Pertanto, fissato un vettore non nullo $\mathbf{u}_r \in \overrightarrow{r}$, ovvero un vettore $\mathbf{u}_r$ parallelo ad r, si ha che $\mathbf{u}_r$ genera $\overrightarrow{r}$ (più precisamente, il sistema $[\mathbf{u}_r]$ è una base di $\overrightarrow{r}$), e quindi

$$r = \left\{ P \in \mathbb{E}^2 \mid \exists\, t \in \mathbb{R} \mid \overrightarrow{P_0P} = t\mathbf{u}_r \right\} \tag{7.4}$$

o anche

$$r = \left\{ P \in \mathbb{E}^2 \mid [\mathbf{u}_r, \overrightarrow{P_0P}] \text{ è dipendente } \right\} . \tag{7.5}$$

Passando alle componenti, dalla (7.4) si deducono le equazioni parametriche di r, dalla (7.5) l'equazione cartesiana. Consideriamo infatti un riferimento

$\mathcal{R} = \big(O; \mathbf{e}_1, \mathbf{e}_2\big)$, che per comodità assumiamo essere monometrico ortogonale (ovvero $\mathbf{e}_1 \cdot \mathbf{e}_2 = 0$ ed $\mathbf{e}_1 \cdot \mathbf{e}_1 = 1 = \mathbf{e}_2 \cdot \mathbf{e}_2$) anche se in molte situazioni che affronteremo tale assunzione sarà superflua. Fissato un punto $P_0 \in \mathbb{E}^2$ e, scelto un vettore non nullo $\mathbf{u}$, vogliamo descrivere la retta r tale che $P_0 \in r$ ed $\overrightarrow{r} = \mathcal{L}(\mathbf{u})$. Indicato con P il generico punto del piano, e posto $P \equiv (x, y)$, $P_0 \equiv (x_0, y_0)$, $\mathbf{u} = (l, m)$, abbiamo che $\overrightarrow{P_0P} = (x - x_0, y - y_0)$ e quindi, in base alla (7.4),

$$P \in r \iff \exists\, t \in \mathbb{R} \mid \overrightarrow{P_0P} = t\mathbf{u}\ ,$$

ovvero, passando alle componenti,

$$(x - x_0, y - y_0) = t(l, m) = (tl, tm)\ .$$

Si ottiene quindi la seguente rappresentazione parametrica di r

$$r : \begin{cases} x - x_0 = tl \\ y - y_0 = tm \end{cases}$$

ovvero

$$r : \begin{cases} x = x_0 + tl \\ y = y_0 + tm \end{cases} \qquad . \tag{7.6}$$

Osserviamo esplicitamente che nella rappresentazione (7.6) i coefficienti del parametro t sono le componenti di un vettore direzionale di r. Utilizzando invece la (7.5), deve risultare dipendente il sistema

$$\big[\,(l, m), (x - x_0, y - y_0)\,\big]$$

e quindi

$$P \in r \iff \rho \begin{pmatrix} l & m \\ x - x_0 & y - y_0 \end{pmatrix} = 1 \iff \det \begin{pmatrix} l & m \\ x - x_0 & y - y_0 \end{pmatrix} = 0\ .$$

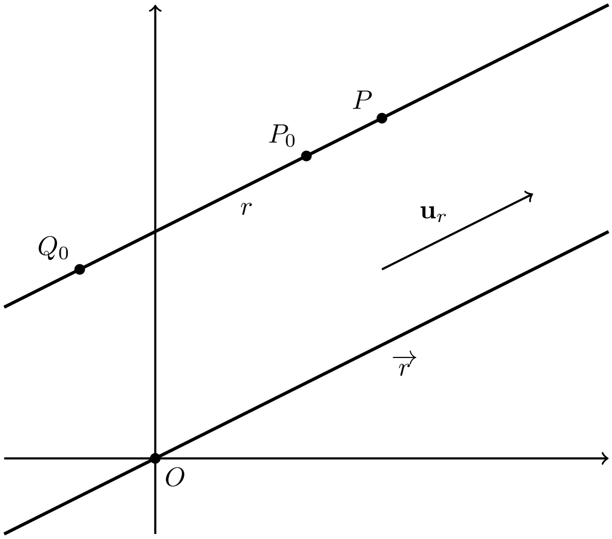
Una rappresentazione cartesiana della retta r è dunque la seguente:

$$r : -mx + ly + (mx_0 - ly_0) = 0\ .$$

Riassumendo, una rappresentazione cartesiana (o anche *implicita*) di r è del tipo

$$r : ax + by + c = 0 \tag{7.7}$$

dove a, b non sono entrambi nulli. In una tale rappresentazione si riconosce immediatamente che un vettore direzionale di r è dato da $\mathbf{u}_r = (-b, a)$. Si potrebbe dimostrare che rappresentano r tutte e sole le equazioni di primo grado in x, y proporzionali alla (7.7).



$$
\begin{aligned}
r &= P_0 + \overrightarrow{r} \\
&= Q_0 + \overrightarrow{r}
\end{aligned}
$$

$$
\begin{aligned}
\overrightarrow{P_0P} &= \tfrac{1}{2}\mathbf{u}_r \\
\overrightarrow{Q_0P} &= 2\mathbf{u}_r
\end{aligned}
$$

In figura, la direttrice $\overrightarrow{r}$ della retta r, che è il sottospazio vettoriale dello spazio vettoriale dei vettori liberi del piano paralleli ad r (come ad esempio $\overrightarrow{P_0P}$, $\overrightarrow{Q_0P}$,$\overrightarrow{P_0Q_0}$), è stata rappresentata graficamente come la parallela ad r per l'origine.

Siamo quindi in grado di descrivere una retta nel piano a partire da un suo punto e da un vettore non nullo ad essa parallelo. Possiamo anche partire da due punti di una retta ed ottenere una analoga descrizione. In effetti nella geometria euclidea, che rappresenta il nostro modello, sappiamo che per due punti distinti passa una ed una sola retta. Nel nostro contesto possiamo procedere come segue. Dati due punti P_0, P_1 distinti, possiamo porre $\mathbf{u} = \overrightarrow{P_0P_1}$ e descrivere la retta a partire dal punto P_0 e dal vettore $\mathbf{u}$.

Date ora due rette r', r'', vogliamo studiare le loro possibili posizioni reciproche. A tale scopo, siano esse rappresentate come segue:

$$
r' : a'x + b'y + c' = 0 \quad ; \quad r'' : a''x + b''y + c'' = 0 \; . \tag{7.8}
$$

È chiaro che l'intersezione $r' \cap r''$ è il luogo rappresentato dal sistema

$$
r' \cap r'' : \begin{cases} a'x + b'y + c' = 0 \\ a''x + b''y + c'' = 0 \end{cases} \tag{7.9}
$$

Posto

$$
A' = \left(\underbrace{\begin{matrix} a' & b' \\ a'' & b'' \end{matrix}}_{A} \left| \begin{matrix} -c' \\ -c'' \end{matrix} \right. \right)
$$

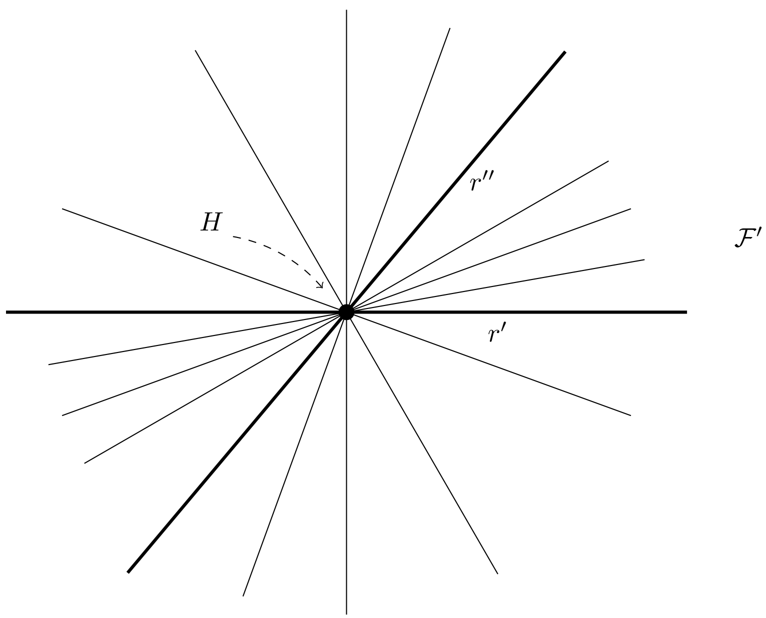
il sistema (7.9), di cui A ed A' sono rispettivamente la matrice incompleta e quella completa, è di Cramer se e solo se $\det A \neq 0$. In tal caso esiste un'unica soluzione $(\bar{x}, \bar{y})$ ed il punto $H \equiv (\bar{x}, \bar{y})$ è l'unico punto in cui le due rette si intersecano. Si dice allora che r', r'' sono incidenti, e H è il punto d'incidenza. Supponiamo ora che $\det A = 0$; le due rette sono allora parallele (infatti $\mathbf{u}_{r'} = (-b', a')$ e $\mathbf{u}_{r''} = (-b'', a'')$ sono proporzionali e quindi generano la stessa direttrice). In particolare, se $\rho(A') = 2$ il sistema (7.9) è incompatibile e le due rette hanno intersezione vuota. Diremo allora che r', r'' sono propriamente parallele, e scriveremo $r' \parallel r''$. Se invece $\rho(A') = 1$ il sistema (7.9) è compatibile (ma non di Cramer) e ammette infinite soluzioni, e le due rette coincidono. Infatti la circostanza che $\rho(A') = 1$ ci dice che le equazioni cartesiane delle due rette sono tra loro proporzionali, e dunque descrivono lo stesso luogo geometrico.

Consideriamo ora due rette distinte r', r'', rappresentate come sopra. Si ha pertanto che $\rho(A') = 2$ e che

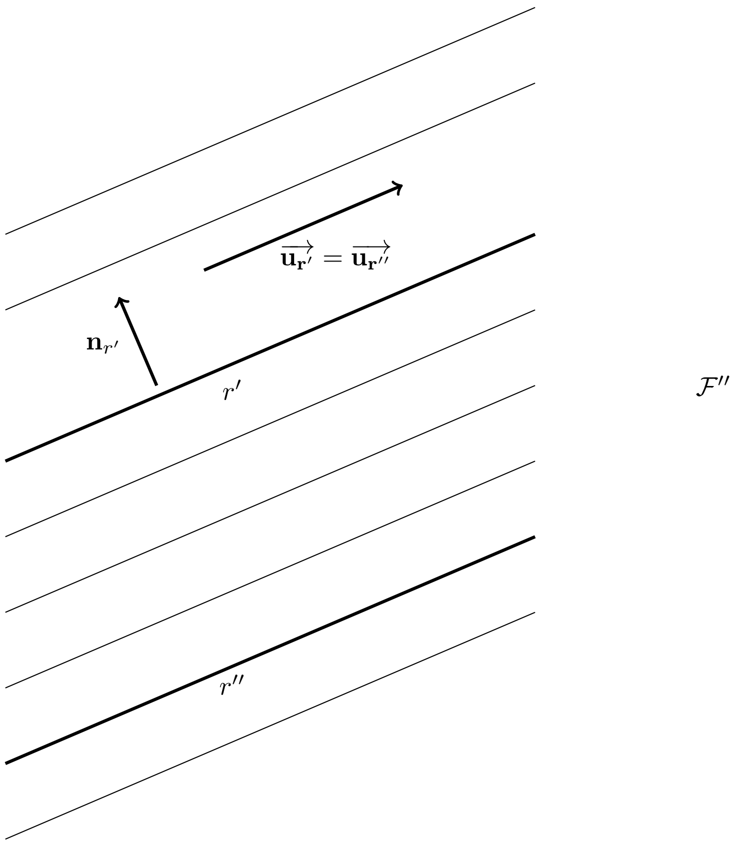
$$r' \parallel r'' \iff \rho(A) = 1 \iff \det A = 0 \ .$$

7.18 Definizione. *Se r', r'' non sono parallele, l'insieme $\mathcal{F}'$ delle rette che contengono $r' \cap r''$ si dice fascio proprio di rette individuato da r', r'. Se invece $r' || r''$, l'insieme $\mathcal{F}''$ delle rette parallele a r' (nonché a r''), ovvero l'insieme delle rette che ammettono $\overrightarrow{r'}$ come direttrice, si dice fascio improprio di rette di direttrice $\overrightarrow{r'}$.*

Nella seguente figura si fornisce un esempio di fascio proprio, e si indica con H il punto di intersezione di r' ed r''.



Nella prossima figura è invece rappresentato un fascio improprio.



Osserviamo che le rette che appartengono al fascio (proprio o improprio) individuato da r', r'' sono tutte e sole quelle del tipo

$$\bar{r} : \lambda(a'x + b'y + c') + \mu(a''x + b''y + c'') = 0$$

dove $\lambda, \mu \in \mathbb{R}$ non sono entrambi nulli.

Nel caso del fascio proprio, detto $\bar{P} \equiv (\bar{x}, \bar{y})$ il punto di incidenza di r', r'', la generica retta di $\mathcal{F}$ è del tipo

$$\bar{r} : a(x - \bar{x}) + b(y - \bar{y}) = 0 \ ,$$

dove a, b sono scalari arbitrari, non entrambi nulli. Nel caso del fascio improprio, la generica retta di $\mathcal{G}$ è del tipo

$$\bar{r} : a'x + b'y + k = 0 \ ,$$

dove k è un arbitrario scalare. Sia nel caso del fascio proprio che di quello improprio, la retta r rappresentata da

$$r : ax + by + c = 0$$

appartiene al fascio individuato da r', r'' se e solo se, posto

$$A = \begin{pmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \end{pmatrix} ,$$

si ha che $\rho(A) = 2$, ovvero $\det A = 0$ in quanto già sappiamo che il rango di A è almeno 2. Siano ora r', r'' come in (7.8). Indicati con $\mathbf{u}_{r'}, \mathbf{u}_{r''}$ i loro rispettivi vettori direzionali, si è già osservato che $\mathbf{u}_{r'} = (-b', a')$, $\mathbf{u}_{r''} = (-b'', a'')$.

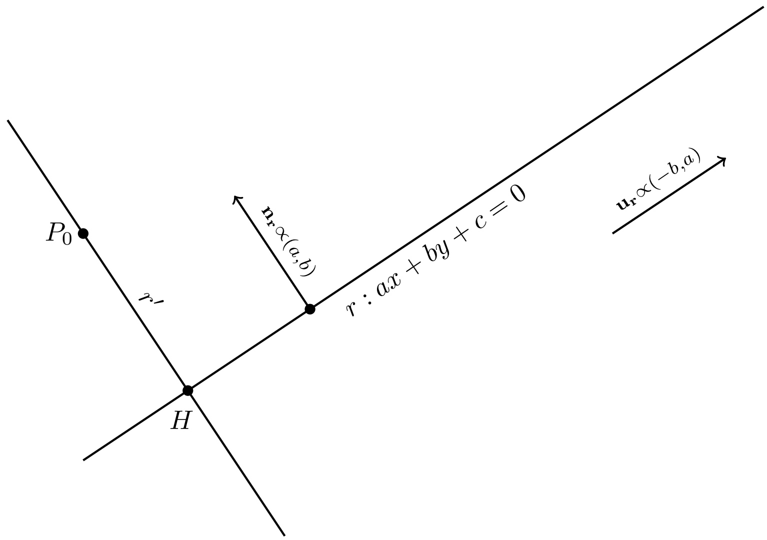
7.19 Definizione. *Le rette r', r'' si dicono ortogonali, e si scrive $r' \perp r''$, se $\mathbf{u}_{r'} \perp \mathbf{u}_{r''}$, ovvero $\mathbf{u}_{r'} \cdot \mathbf{u}_{r''} = 0$, e cioè $a'a'' + b'b'' = 0$.*

Osserviamo che se $r' \perp r''$, sicuramente r', r'' sono incidenti. Si dice anche che r', r'' sono perpendicolari. Ricordiamo che se $P_0 \equiv (x_0, y_0)$ e $P_1 \equiv (x_1, y_1)$ sono due punti, la loro distanza $d(P_0, P_1)$ è definita ponendo

$$d(P_0, P_1) = \left|\overrightarrow{P_0P_1}\right| = \sqrt{(x_1 - x_0)^2 + (y_1 - y_0)^2} \; .$$

Se consideriamo la retta r rappresentata come in (7.7), si verifica agevolmente che esiste un unico punto $H \in r$ tale che $d(P_0, r) = d(P_0, H)$. Si osserva infatti che esiste un'unica retta r' tale che $P_0 \in r'$ e $r' \perp r$. Basta porre $\mathbf{u}_{r'} = (a, b)$ (e così $\mathbf{u}_r \perp \mathbf{u}_{r'}$) e scrivere r' in forma parametrica. Il punto H si trova allora come intersezione di r ed r'. Si verifica facilmente che

$$d(P_0, r) = \frac{\left|ax_0 + by_0 + c\right|}{\sqrt{a^2 + b^2}} \; .$$



Sia ancora r come in (7.7) e consideriamo il vettore $\mathbf{n}_r = (a, b)$, che si dice *vettore normale* di r. Osserviamo che $\mathbf{n}_r \perp r$, ovvero $\mathbf{n}_r \perp \overrightarrow{r}$, e cioè $\mathbf{n}_r \perp \mathbf{w}$ per ogni $\mathbf{w} \in \overrightarrow{r}$. Infatti, poiché $\mathbf{u}_r$ genera $\overrightarrow{r}$, ogni $\mathbf{w} \in \overrightarrow{r}$ è del tipo $\mathbf{w} = k\mathbf{u}_r$, dove $\mathbf{u}_r = (-b, a)$ è il vettore direzionale di r e k è un opportuno scalare, ed è chiaro che $\mathbf{n}_r \cdot \mathbf{u}_r = (a, b) \cdot (-b, a) = -ab + ab = 0$, cioè $\mathbf{n}_r \perp \mathbf{u}_r$. Pertanto, considerate ancora le rette r', r'' come in (7.8) e i loro vettori normali $\mathbf{n}_{r'} = (a', b')$, $\mathbf{n}_{r''} = (a'', b'')$, la condizione di parallelismo tra tali rette del piano può anche esprimersi come segue:

$$r' \parallel r'' \quad \iff \quad \mathbf{u}_{r'} \parallel \mathbf{u}_{r''} \quad \iff \quad \mathbf{n}_{r'} \parallel \mathbf{n}_{r''} \quad \iff \quad (a', b') \propto (a'', b'') \ ,$$

o anche

$$r' \parallel r'' \quad \iff \quad \rho \begin{pmatrix} a' & b' \\ a'' & b'' \end{pmatrix} = 1 \quad \iff \quad \begin{vmatrix} a' & b' \\ a'' & b'' \end{vmatrix} = 0 \ .$$

7.7 Lo spazio affine ed euclideo reale

Sia ora $n = 3$. Osserviamo esplicitamente che il modello classico per lo spazio affine $(\mathbb{A}, V, \mathbb{R})$ con $\dim \mathbb{A} = 3$ è lo spazio euclideo della geometria elementare. In tale modello, gli elementi di $\mathbb{A}$ sono i punti dello spazio e V è lo spazio vettoriale dei vettori liberi ordinari dello spazio. Laddove occorre considerare questioni metriche, angolari o di ortogonalità, si assume che sia dato in V un prodotto scalare definito positivo. Indichiamo tale spazio affine euclideo con il simbolo $\mathbb{E}^3$. I sottospazi di $\mathbb{E}^3$ sono i singleton (o impropriamente i punti) che hanno dimensione 0, le rette, di dimensione 1, i piani, di dimensione 2, ed infine $\mathbb{E}^3$ stesso, di dimensione 3.

In questa esposizione supponiamo di aver fissato un riferimento

$$\mathcal{R} = \big(O; \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\big)$$

che per comodità assumiamo essere monometrico ortogonale (ovvero $\mathbf{e}_h \cdot \mathbf{e}_k = \delta_{h,k}$ per ogni $h, k = 1, 2, 3$) anche se in molte situazioni che affronteremo tale assunzione risulterà superflua. In particolare, dati due vettori $\mathbf{u}' = (x', y', z')$; $\mathbf{u}'' = (x'', y'', z'')$ di V, il loro prodotto scalare è dato dalla formula $\mathbf{u}' \cdot \mathbf{u}'' = x'x'' + y'y'' + z'z''$. Vogliamo definire una nuova operazione interna nello spazio dei vettori liberi ordinari, detto prodotto vettoriale. Dati due vettori $\mathbf{u}'$ e $\mathbf{u}''$, il loro prodotto vettoriale, che sarà indicato con il simbolo $\mathbf{u}' \times \mathbf{u}''$ o anche $\mathbf{u}' \wedge \mathbf{u}''$, è il vettore $\mathbf{w} = (l, m, n)$, ovvero $\mathbf{w} = l\mathbf{e}_1 + m\mathbf{e}_2 + n\mathbf{e}_3$, dove

$$l = \begin{vmatrix} y' & z' \\ y'' & z'' \end{vmatrix} \quad ; \quad m = -\begin{vmatrix} x' & z' \\ x'' & z'' \end{vmatrix} \quad ; \quad n = \begin{vmatrix} x' & y' \\ x'' & y'' \end{vmatrix} \ .$$

Osserviamo che l, m, n sono i minori di ordine 2, presi a segni alterni, della matrice

$$\begin{pmatrix} x' & y' & z' \\ x'' & y'' & z'' \end{pmatrix} \ .$$

7.20 Proposizione. *Valgono le seguenti proprietà del prodotto vettoriale:*

(i) $\mathbf{u}' \wedge \mathbf{u}'' = -\mathbf{u}'' \wedge \mathbf{u}'$;

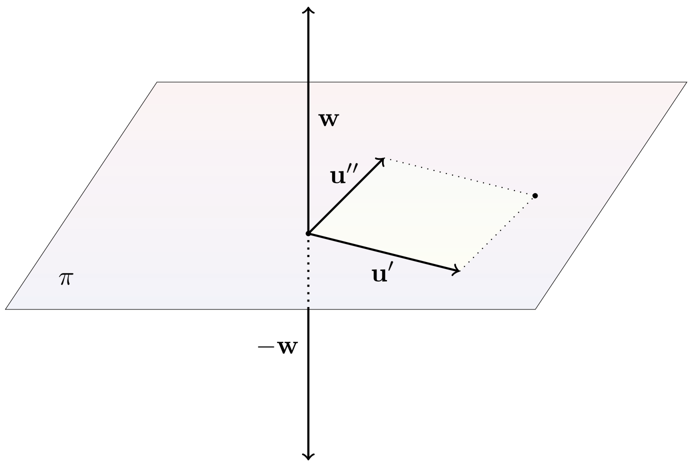
(ii) $\mathbf{u}' \wedge \mathbf{u}'' = \mathbf{0} \iff \big[\mathbf{u}', \mathbf{u}''\big]$ *è un sistema dipendente;*

(iii) $\mathbf{w} \perp \mathbf{u}'$ *e* $\mathbf{w} \perp \mathbf{u}''$.

Poiché $\mathbf{w} = l\mathbf{e}_1 + m\mathbf{e}_2 + n\mathbf{e}_3$, risulta anche

$$\mathbf{w} = \mathbf{u}' \wedge \mathbf{u}'' = \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ x' & y' & z' \\ x'' & y'' & z'' \end{vmatrix} \ .$$

Nella figura seguente sono rappresentati due vettori non paralleli $\mathbf{u}', \mathbf{u}''$ su un piano π, il parallelogramma individuato da $\mathbf{u}', \mathbf{u}''$, ed infine il vettore $\mathbf{w} = \mathbf{u}' \wedge \mathbf{u}''$ ed il suo opposto $-\mathbf{w} = \mathbf{u}'' \wedge \mathbf{u}'$ che sono entrambi ortogonali a π.



Vogliamo ora rappresentare, in forma cartesiana o parametrica, i piani in $\mathbb{E}^3$. Come osservato in generale, un piano π ha giacitura $\overrightarrow{\pi} \leq V$, un sottospazio vettoriale di dimensione 2, e per ogni $P_0 \in \pi$ si ha che $\pi = P_0 + \overrightarrow{\pi}$ ovvero

$$\pi = \left\{ P \in \mathbb{E}^3 \;\middle|\; \overrightarrow{P_0P} \in \overrightarrow{\pi} \right\} .$$

Pertanto, fissati due vettori indipendenti $\mathbf{u}', \mathbf{u}'' \in \overrightarrow{\pi}$, si ha che $\mathbf{u}', \mathbf{u}''$ generano $\overrightarrow{\pi}$ (più precisamente, il sistema $[\mathbf{u}', \mathbf{u}'']$ è una base di $\overrightarrow{\pi}$) e quindi

$$\pi = \left\{ P \in \mathbb{E}^3 \;\middle|\; \exists\, t', t'' \in \mathbb{R} \;\middle|\; \overrightarrow{P_0P} = t'\mathbf{u}' + t''\mathbf{u}'' \right\} \tag{7.10}$$

o anche

$$\pi = \left\{ P \in \mathbb{E}^3 \;\middle|\; \left[\mathbf{u}', \mathbf{u}'', \overrightarrow{P_0P}\right] \text{ è dipendente } \right\} . \tag{7.11}$$

Passando alle componenti, dalla (7.10) si deducono le equazioni parametriche di π, dalla (7.11) l'equazione cartesiana di π. Fissiamo dunque un punto $P_0 \in \mathbb{E}^3$ e, scelti due vettori indipendenti $\mathbf{u}', \mathbf{u}''$, descriviamo il piano π tale che $P_0 \in \pi$ e $\overrightarrow{\pi} = \mathcal{L}(\mathbf{u}', \mathbf{u}'')$, ovvero il piano passante per P_0 e parallelo ad $\mathbf{u}', \mathbf{u}''$. Indicato con P il generico punto dello spazio, e posto $P \equiv (x, y, z)$, $P_0 \equiv (x_0, y_0, z_0)$, $\mathbf{u}' = (l', m', n')$, $\mathbf{u}'' = (l'', m'', n'')$, abbiamo che $\overrightarrow{P_0P} = (x - x_0, y - y_0, z - z_0)$ e quindi, in base alla (7.10),

$$P \in \pi \iff \exists\, t', t'' \in \mathbb{R} \mid \overrightarrow{P_0P} = t'\mathbf{u}' + t''\mathbf{u}'' ,$$

ovvero, passando alle componenti,

$$\begin{aligned}(x - x_0, y - y_0, z - z_0) &= t'(l', m', n') + t''(l'', m'', n'') \\ &= (t'l' + t''l'', t'm' + t''m'', t'n' + t''n'') .\end{aligned}$$

Si ottiene quindi la seguente rappresentazione parametrica di π

$$\pi : \begin{cases} x - x_0 = t'l' + t''l'' \\ y - y_0 = t'm' + t''m'' \\ z - z_0 = t'n' + t''n'' \end{cases}$$

ovvero

$$\pi : \begin{cases} x = x_0 + t'l' + t''l'' \\ y = y_0 + t'm' + t''m'' \\ z = z_0 + t'n' + t''n'' \end{cases} \tag{7.12}$$

Osserviamo esplicitamente che nella rappresentazione (7.12) i coefficienti dei parametri t', t'' sono le componenti di due vettori che generano la giacitura di π. Utilizzando invece la (7.11), deve risultare dipendente il sistema

$$\big[\, (l', m', n'), (l'', m'', n''), (x - x_0, y - y_0, z - z_0) \,\big]$$

e quindi, posto

$$C = \begin{pmatrix} l' & m' & n' \\ l'' & m'' & n'' \\ x - x_0 & y - y_0 & z - z_0 \end{pmatrix} ,$$

si ha che

$$P \in \pi \iff \rho(C) = 2 , \tag{7.13}$$

poiché già sappiamo che le prime due righe di C sono indipendenti. Pertanto l'uguaglianza al II membro della (7.13) equivale al fatto che $\det C = 0$ e quindi, calcolando tale determinante con la regola di Laplace sulla terza riga, una rappresentazione cartesiana del piano π è data da

$$\pi : ax + by + cz + d = 0 \tag{7.14}$$

dove

$$a = \begin{vmatrix} m' & n' \\ m'' & n'' \end{vmatrix} \;;\; b = -\begin{vmatrix} l' & n' \\ l'' & n'' \end{vmatrix} \;;\; c = \begin{vmatrix} l' & m' \\ l'' & m'' \end{vmatrix} \;;\; d = -\begin{vmatrix} l' & m' & n' \\ l'' & m'' & n'' \\ x_0 & y_0 & z_0 \end{vmatrix}$$

(e quindi, fra l'altro, a, b, c non sono tutti nulli). Si potrebbe facilmente verificare che ogni equazione del tipo (7.14), con a, b, c non tutti nulli, rappresenta un piano, ed è poi chiaro che equazioni proporzionali rappresentano lo stesso piano.

Siamo quindi in grado di descrivere un piano nello spazio a partire da un suo punto e da due vettori indipendenti ad esso paralleli (o, equivalentemente, appartenenti alla sua giacitura). Possiamo anche partire da tre punti non allineati di un piano (ovvero non appartenenti ad una stessa retta) ed ottenere una analoga descrizione. In effetti nella geometria euclidea, che rappresenta il nostro modello, sappiamo che per tre punti non allineati passa uno ed un solo piano. Nel nostro contesto possiamo procedere come segue. Dati tre punti P_0, P_1, P_2

non allineati, possiamo porre $\mathbf{u}' = \overrightarrow{P_0P_1}$, $\mathbf{u}'' = \overrightarrow{P_0P_2}$ e descrivere il piano a partire dal punto P_0 e dai vettori $\mathbf{u}', \mathbf{u}''$.

Sia π come in (7.14) e consideriamo il vettore $\mathbf{n}_\pi = (a, b, c)$, che si dice *vettore normale* di π. Osserviamo che $\mathbf{n}_\pi \perp \pi$, ovvero $\mathbf{n}_\pi \perp \overrightarrow{\pi}$, e cioè $\mathbf{n}_\pi \perp \mathbf{w}$ per ogni $\mathbf{w} \in \overrightarrow{\pi}$. Si può provare ciò al modo seguente. Posto

$$\mathbf{u} = (-b, a, 0) \quad ; \quad \mathbf{v} = (-c, 0, a) \quad ; \quad \mathbf{w} = (0, -c, b) \tag{7.15}$$

non è difficile verificare che $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \overrightarrow{\pi}$, ovvero che $\mathbf{u}, \mathbf{v}, \mathbf{w} \parallel \pi$. Tenendo conto che a, b, c non possono annullarsi simultaneamente, vediamo che due di tali vettori costituiscono una base di $\overrightarrow{\pi}$. Ad esempio, se $a \neq 0$, $[\mathbf{u}, \mathbf{v}]$ è una base di $\overrightarrow{\pi}$. In particolare, il sistema $[\mathbf{u}, \mathbf{v}, \mathbf{w}]$ genera $\overrightarrow{\pi}$. È chiaro quindi che $\mathbf{n}_\pi \perp \pi$ se e solo se $\mathbf{n}_\pi \cdot \mathbf{u} = \mathbf{n}_\pi \cdot \mathbf{v} = \mathbf{n}_\pi \cdot \mathbf{w} = 0$, e ciò è immediato. Una rappresentazione della giacitura $\overrightarrow{\pi}$ del piano π, come sottospazio vettoriale di V nella base fissata, è data da

$$\overrightarrow{\pi} : ax + by + cz = 0 \ .$$

È infatti chiaro che i vettori $\mathbf{u}, \mathbf{v}, \mathbf{w}$ della (7.15), che generano $\overrightarrow{\pi}$, hanno componenti che verificano tale equazione.

Consideriamo ora due piani

$$\pi' : a'x + b'y + c'z + d' = 0 \quad , \quad \pi'' : a''x + b''y + c''z + d'' = 0 \ , \tag{7.16}$$

ed indichiamo con $\mathbf{n}_{\pi'} = (a', b', c')$, $\mathbf{n}_{\pi''} = (a'', b'', c'')$ i loro vettori normali. Per definizione, sappiamo che $\pi' \parallel \pi''$ se e solo se $\overrightarrow{\pi'} = \overrightarrow{\pi''}$ e cioè se e solo se le equazioni $a'x + b'y + c'z = 0$, $a''x + b''y + c''z = 0$ rappresentano lo stesso sottospazio, ovvero sono proporzionali. Quindi

$$\pi' \parallel \pi'' \iff (a', b', c') \propto (a'', b'', c'') \iff \rho \begin{pmatrix} a' & b' & c' \\ a'' & b'' & c'' \end{pmatrix} = 1 \ ,$$

o anche

$$\pi' \parallel \pi'' \iff \mathbf{n}_{\pi'} \parallel \mathbf{n}_{\pi''} \iff \big[\mathbf{n}_{\pi'}, \mathbf{n}_{\pi''}\big] \text{ è dipendente} \ .$$

Se ciò accade, ed inoltre

$$\rho \begin{pmatrix} a' & b' & c' & d' \\ a'' & b'' & c'' & d'' \end{pmatrix} = 1 \ ,$$

le (7.16) risultano proporzionali, e quindi $\pi' = \pi''$. Si dice allora che π', π'' sono *impropriamente* paralleli o coincidenti. Se invece

$$\rho \begin{pmatrix} a' & b' & c' \\ a'' & b'' & c'' \end{pmatrix} = 1 \quad , \quad \rho \begin{pmatrix} a' & b' & c' & d' \\ a'' & b'' & c'' & d'' \end{pmatrix} = 2 \ ,$$

il sistema

$$\begin{cases} a'x + b'y + c'z + d' = 0 \\ a''x + b''y + c''z + d'' = 0 \end{cases} \tag{7.17}$$

non è compatibile, si ha che $\pi' \cap \pi'' = \emptyset$ e i due piani si dicono *propriamente* paralleli. Se, infine, $\pi' \not\parallel \pi''$, deve essere

$$\rho \begin{pmatrix} a' & b' & c' \\ a'' & b'' & c'' \end{pmatrix} = \rho \begin{pmatrix} a' & b' & c' & d' \\ a'' & b'' & c'' & d'' \end{pmatrix} = 2 \ ,$$

ed il sistema (7.17) è compatibile e ridotto, ammette infinite soluzioni dipendenti da un parametro, e rappresenta, come vedremo in seguito, una retta.

Abbiamo già affrontato il problema di rappresentare, in forma cartesiana o parametrica, le rette in uno spazio affine di dimensione arbitraria e anche il caso particolare delle rette in un piano affine. Vediamo cosa succede in dettaglio in $\mathbb{E}^3$. Come osservato in generale, una retta r ha direttrice $\overrightarrow{r} \leq V$, un sottospazio vettoriale di dimensione 1, e per ogni $P_0 \in r$ si ha che $r = P_0 + \overrightarrow{r}$ ovvero

$$r = \left\{ P \in \mathbb{E}^3 \ \middle| \ \exists \ \mathbf{v} \in V \ \middle| \ \overrightarrow{P_0P} = \mathbf{v} \right\} \ .$$

Pertanto, fissato un vettore non nullo $\mathbf{u}_r \in \overrightarrow{r}$, si ha che $\mathbf{u}_r$ genera $\overrightarrow{r}$ (più precisamente, il sistema $[\mathbf{u}_r]$ è una base di $\overrightarrow{r}$) e quindi

$$r = \left\{ P \in \mathbb{E}^3 \ \middle| \ \exists \ t \in \mathbb{R} \ \middle| \ \overrightarrow{P_0P} = t\mathbf{u}_r \right\} \tag{7.18}$$

o anche

$$r = \left\{ P \in \mathbb{E}^3 \ \middle| \ \left[\mathbf{u}_r, \overrightarrow{P_0P}\right] \text{ è dipendente } \right\} \ . \tag{7.19}$$

esattamente come nel caso delle rette in un piano affine. Passando alle componenti, dalla (7.18) si deducono le equazioni parametriche di r, dalla (7.19) le equazioni cartesiane (che in questo caso saranno due). Fissato un punto $P_0 \in \mathbb{E}^3$ e scelto un vettore non nullo $\mathbf{u}$, vogliamo descrivere la retta r tale che $P_0 \in r$ ed $\overrightarrow{r} = \mathcal{L}(\mathbf{u})$. Indicato con P il generico punto del piano, e posto $P \equiv (x, y, z)$, $P_0 \equiv (x_0, y_0, z_0)$, $\mathbf{u} = (l, m, n)$, abbiamo che $\overrightarrow{P_0P} = (x - x_0, y - y_0, z - z_0)$ e quindi, in base alla (7.18),

$$P \in r \iff \exists \ t \in \mathbb{R} \mid \overrightarrow{P_0P} = t\mathbf{u} \ ,$$

ovvero, passando alle componenti,

$$(x - x_0, y - y_0, z - z_0) = t(l, m, n) = (tl, tm, tn) \ .$$

Si ottiene quindi la seguente rappresentazione parametrica di r

$$r : \begin{cases} x - x_0 = tl \\ y - y_0 = tm \\ z - z_0 = tn \end{cases}$$

ovvero

$$r : \begin{cases} x = x_0 + tl \\ y = y_0 + tm \\ z = z_0 + tn \end{cases} \ . \tag{7.20}$$

Osserviamo esplicitamente che nella rappresentazione (7.20) i coefficienti del parametro t sono le componenti di un vettore direzionale di r. Utilizzando invece la (7.19), deve risultare dipendente il sistema

$$\big[\,(l,m,n),(x-x_0,y-y_0,z-z_0)\,\big]$$

e quindi, posto

$$C=\begin{pmatrix} l & m & n \\ x-x_0 & y-y_0 & z-z_0 \end{pmatrix}\ ,$$

si ha che

$$P\in r \iff \rho(C)=1\ . \tag{7.21}$$

Poiché $\mathbf{u}\neq\mathbf{0}$, gli scalari l,m,n non sono tutti nulli. Supponiamo ad esempio che sia $l\neq 0$. Il Teorema degli orlati ci dice che l'uguaglianza al II membro della (7.21) equivale al fatto che gli orlati dell'elemento l in C siano tutti degeneri, ovvero

$$\begin{vmatrix} l & m \\ x-x_0 & y-y_0 \end{vmatrix}=0 \quad ; \quad \begin{vmatrix} l & n \\ x-x_0 & z-z_0 \end{vmatrix}=0\ ,$$

e quindi una rappresentazione cartesiana della retta r è data da

$$r:\begin{cases} -mx+ly+(mx_0-ly_0)=0 \\ -nx+lz+(nx_0-lz_0)=0 \end{cases}$$

Riassumendo, una rappresentazione cartesiana (o anche *implicita*) della retta r è del tipo

$$r:\begin{cases} a'x+b'y+c'z+d'=0 \\ a''x+b''y+c''z+d''=0 \end{cases} \tag{7.22}$$

dove, posto

$$A'=\left(\underbrace{\begin{matrix} a' & b' & c' \\ a'' & b'' & c'' \end{matrix}}_{A}\left|\begin{matrix} -d' \\ -d'' \end{matrix}\right.\right)\ ,$$

si ha che $\rho(A)=2$. In tale situazione è chiaro che le due equazioni (7.22) rappresentano due piani, diciamo π',π'', che risultano tra loro non paralleli, e si ha che

$$r=\pi'\cap\pi''\ .$$

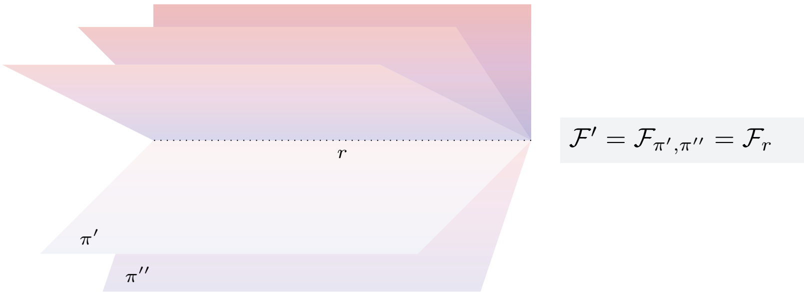
Si potrebbe facilmente verificare che ogni sistema del tipo (7.22), con $\rho(A)=2$, rappresenta una retta r e che un vettore direzionale $\mathbf{u}_r=(l,m,n)$ di r si ottiene ponendo

$$l=\begin{vmatrix} b' & c' \\ b'' & c'' \end{vmatrix} \quad ; \quad m=-\begin{vmatrix} a' & c' \\ a'' & c'' \end{vmatrix} \quad ; \quad n=\begin{vmatrix} a' & b' \\ a'' & b'' \end{vmatrix}\ , \tag{7.23}$$

(ovvero $\mathbf{u}_r=\mathbf{n}_{\pi'}\wedge\mathbf{n}_{\pi''}$, se il riferimento scelto è ortonormale).

Consideriamo ora due piani distinti π',π'', rappresentati come in (7.16). Si ha pertanto che $\rho(A')=2$ e che $\pi'\parallel\pi'' \iff \rho(A)=1$.

7.21 DEFINIZIONE. *Siano π', π'' due piani distinti. Se π', π'' non sono paralleli, e quindi la loro intersezione è una retta, posto ad esempio $r = \pi' \cap \pi''$, l'insieme $\mathcal{F}'$ dei piani che contengono r si dice fascio (proprio) di piani individuato da π', π''. Se invece $\pi' || \pi''$, e cioè $\overrightarrow{\pi}' = \overrightarrow{\pi}''$, l'insieme $\mathcal{F}''$ dei piani paralleli a π' (nonché a π''), ovvero l'insieme dei piani che ammettono $\overrightarrow{\pi}'$ come sottospazio direttore, si dice fascio improprio di iperpiani di giacitura $\overrightarrow{\pi}'$.*



La figura precedente esemplifica la nozione di fascio proprio, la successiva quella di fascio improprio.

Osserviamo che i piani che appartengono al fascio (proprio o improprio) individuato da due iperpiani distinti π', π'' sono tutti e soli quelli del tipo

$$\bar{\pi} : \lambda(a_1'x_1 + \cdots + a_n'x_n + a') + \mu(a_1''x_1 + \cdots + a_n''x_n + a'') = 0$$

dove $\lambda, \mu \in \mathbb{R}$ sono non entrambi nulli. Indichiamo con $\mathcal{F}$, o anche $\mathcal{F}_{\pi', \pi''}$, il fascio (proprio o improprio) individuato da π', π''.

Se π', π'' non sono paralleli, e quindi i piani di $\mathcal{F}$ sono tutti e soli quelli che contengono la retta r, tale retta si dice asse del fascio. Se invece π', π'' sono paralleli, i piani di $\mathcal{F}$ sono tutti e soli quelli paralleli a π' (e anche a π''), e inoltre il generico piano $\bar{\pi}$ del fascio $\mathcal{F}$ ammette una rappresentazione cartesiana del tipo

$$\bar{\pi} : a'_1 x_1 + \cdots + a'_n x_n + k = 0$$

dove k è uno scalare.

In entrambi i casi, il piano π rappresentato dall'equazione $\pi : ax + by + cz + d = 0$ appartiene al fascio individuato da π', π'' se e solo se

$$\rho \begin{pmatrix} a & b & c & d \\ a' & b' & c' & d' \\ a'' & b'' & c'' & d'' \end{pmatrix} = 2 \ .$$

Sia ora r una retta come in (7.20) o in (7.18) e π un piano come in (7.14), ed indichiamo con $\mathbf{u}_r = (l, m, n)$ il vettore direzionale di r e con $\mathbf{n}_\pi = (a, b, c)$ il vettore normale di π. Poniamo

$$A' = \left(\begin{array}{ccc|c} \multicolumn{3}{c|}{\underbrace{\begin{matrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \end{matrix}}_{A}} & \begin{matrix} d \\ d' \\ d'' \end{matrix} \end{array} \right) .$$

Sappiamo che, per definizione, r è parallela a π se e solo se $\overrightarrow{r} \leq \overrightarrow{\pi}$, ovvero $\mathbf{u}_r \in \overrightarrow{\pi}$. Si vede che ciò equivale a dire che $\mathbf{u}_r \perp \mathbf{n}_\pi$. Pertanto

$$r \parallel \pi \iff \mathbf{u}_r \perp \mathbf{n}_\pi \iff \mathbf{u}_r \cdot \mathbf{n}_\pi = 0 \ .$$

Ricordando la descrizione (7.23) di $\mathbf{u}_r$ e calcolando $\det A$ con la regola di Laplace sulla I riga, vediamo che $\det A = \mathbf{u}_r \cdot \mathbf{n}_\pi$ e quindi

$$r \parallel \pi \iff \det A = 0$$

con $\rho(A) = 2$. Se r è parallela a π, distinguiamo due casi. Diremo che r e π sono *impropriamente paralleli* se $r \subset \pi$, e ciò equivale a dire che $\rho(A') = 2$. Altrimenti avremo che $\rho(A') = 3$, ma $\rho(A) = 2$ e quindi il sistema

$$\begin{cases} ax + by + cz + d = 0 \\ a'x + b'y + c'z + d' = 0 \\ a''x + b''y + c''z + d'' = 0 \end{cases} \tag{7.24}$$

che rappresenta $r \cap \pi$, è incompatibile. Pertanto $r \cap \pi = \emptyset$ e diremo che r e π sono *propriamente paralleli*. In sintesi

$$r \parallel \pi \iff \rho(A) = 2 \iff \det A = 0$$

e tale parallelismo è proprio, ovvero $r \cap \pi = \emptyset$, se e solo se $\rho(A') = 3$. Se invece $r \not\parallel \pi$, ovvero $\rho(A) = 3$ e $\det A \neq 0$, il sistema (7.24) è di Cramer. Se indichiamo con $(\bar{x}, \bar{y}, \bar{z})$ l'unica soluzione di (7.24), e poniamo $\bar{P} = (\bar{x}, \bar{y}, \bar{z})$, si ha che $r \cap \pi = \{\bar{P}\}$ e si dice che r e π sono incidenti. Se r è rappresentata come in (7.18), le coordinate del punto di incidenza si possono anche determinare sostituendo i secondi membri delle (7.18) al posto di x, y, z in (7.14), risolvendo la risultante equazione in t e sostituendo tale valore nelle (7.18).

7.8 Posizione reciproca tra rette

Consideriamo due rette r, s, dove r è rappresentata come in (7.22) o (7.20) ed

$$\begin{cases} a'''x + b'''y + c'''z + d''' = 0 \\ a''''x + b''''y + c''''z + d'''' = 0 \end{cases} \tag{7.25}$$

dove

$$\rho \begin{pmatrix} a''' & b''' & c''' \\ a'''' & b'''' & c'''' \end{pmatrix} = 2 \ .$$

Ricordiamo che le rette r, s sono parallele quando le loro direttrici $\overrightarrow{r}$ e $\overrightarrow{s}$ coincidono. Parleremo poi di parallelismo improprio se $r = s$ e di parallelismo proprio se $r \neq s$ (e quindi $r \cap s = \emptyset$). Diremo che r, s sono incidenti se si intersecano in un punto. Sia nel caso del parallelismo proprio che nel caso dell'incidenza è agevole verificare che esiste un unico piano π che contiene entrambe le rette. Pertanto diremo che r, s sono complanari. In caso contrario, r, s sono non complanari, o anche *sghembe*.

Vogliamo studiare le possibili posizioni reciproche di r, s. Indichiamo, come al solito, con π', π'' i piani rappresentati dalle (7.22) e con π''', π'''' quelli rappresentati dalle (7.25). Poniamo, per comodità,

$$A'_r = \left(\begin{array}{c|c} \underbrace{\begin{matrix} a' & b' & c' \\ a'' & b'' & c'' \end{matrix}}_{A_r} & \begin{matrix} -d' \\ -d'' \end{matrix} \end{array} \right) \quad ; \quad A'_s = \left(\begin{array}{c|c} \underbrace{\begin{matrix} a''' & b''' & c''' \\ a'''' & b'''' & c'''' \end{matrix}}_{A_s} & \begin{matrix} -d''' \\ -d'''' \end{matrix} \end{array} \right)$$

e

$$A' = \begin{pmatrix} A'_r \\ \hline A'_s \end{pmatrix} = \left(\begin{array}{c|c} \underbrace{\begin{matrix} a' & b' & c' \\ a'' & b'' & c'' \\ a''' & b''' & c''' \\ a'''' & b'''' & c'''' \end{matrix}}_{A} & \begin{matrix} d' \\ d'' \\ d''' \\ d'''' \end{matrix} \end{array} \right) \tag{7.26}$$

Osserviamo che

$$\rho(A_r) = \rho(A_s) = \rho(A'_r) = \rho(A'_r) = 2$$

e quindi

$$2 \leq \rho(A) \leq \rho(A') \leq 4 \ .$$

Inoltre i piani π', π'' sono incidenti e si ha che $r = \pi' \cap \pi''$ ed analogamente piani π''', π'''' sono incidenti e si ha che $s = \pi''' \cap \pi''''$. Consideriamo il seguente sistema lineare, che rappresenta l'intersezione tra le due rette r, s ed ammette le matrici A, A' descritte dalla (7.26) come matrici associate, a meno del segno degli elementi dell'ultima colonna che non incide nel calcolo dei ranghi

$$r \cap s : \begin{cases} a'x + b'y + c'z = -d' \\ a''x + b''y + c''z = -d'' \\ a'''x + b'''y + c'''z = -d''' \\ a''''x + b''''y + c''''z = -d'''' \end{cases} \tag{7.27}$$

- *Caso 1.* Supponiamo che sia $\rho(A) = \rho(A') = 2$. Allora il sistema (7.27) è compatibile, per il Teorema di Rouché-Capelli, e si riduce al sistema (7.22), o equivalentemente al sistema (7.24), che rappresentano, rispettivamente, r ed s. Pertanto $r = r \cap s = s$, ovvero r ed s coincidono. Vale anche il viceversa, nel senso che se r, s sono rappresentate come sopra ed esse coincidono, allora $\rho(A) = \rho(A') = 2$.

- *Caso 2.* Se $\rho(A) = 2$ e $\rho(A') = 3$, il sistema (7.27) è incompatibile e $r \cap s = \emptyset$. Consideriamo le rette

$$\bar{r} : \begin{cases} a'x + b'y + c'z = 0 \\ a''x + b''y + c''z = 0 \end{cases} \quad ; \quad \bar{s} : \begin{cases} a'''x + b'''y + c'''z = 0 \\ a''''x + b''''y + c''''z = 0 \end{cases}$$

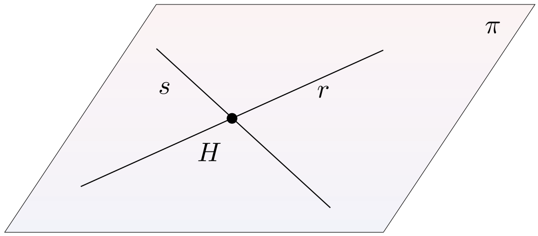
ed i piani

$$\bar{\pi}' : a'x + b'y + c'z = 0 \ ; \ \bar{\pi}'' : a''x + b''y + c''z = 0$$

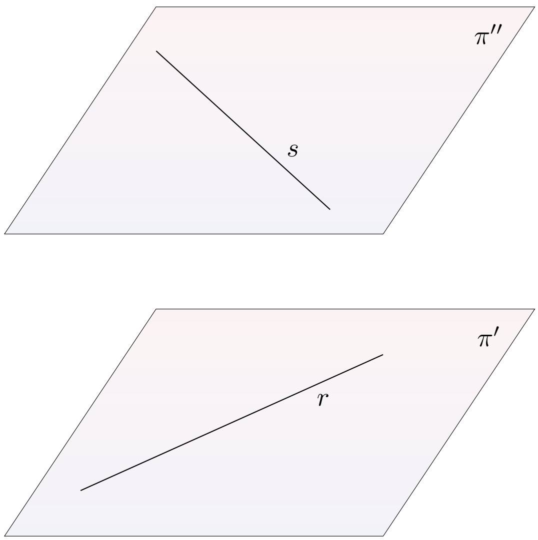
$$\bar{\pi}''' : a'''x + b'''y + c'''z = 0 \ ; \ \bar{\pi}'''' : a''''x + b''''y + c''''z = 0 \ .$$

Abbiamo che $\bar{r}, \bar{s}, \bar{\pi}', \bar{\pi}'', \bar{\pi}''', \bar{\pi}''''$ sono le rette ed i piani paralleli rispettivamente ad $r, s, \pi', \pi'', \pi''', \pi''''$, passanti per l'origine. Il fatto che $\rho(A) = 2$ ci dice che $\bar{\pi}''', \bar{\pi}''''$ appartengono al fascio individuato da $\bar{\pi}', \bar{\pi}''$, ovvero $\bar{r} = \bar{s}$, e quindi $r \parallel \bar{r} = \bar{s} \parallel s$. In definitiva, le rette r, s sono propriamente parallele. Vale anche il viceversa, nel senso che se $r \parallel s$ in senso proprio allora $\rho(A) = 2$ e $\rho(A') = 3$.

- *Caso 3.* Sia ora $\rho(A) = 3$, $\rho(A') = 3$. Il sistema (7.27) è compatibile e si riduce, cancellando opportunamente una delle quattro equazioni, ad un sistema di Cramer che ammette come unica soluzione una terna $(\bar{x}, \bar{y}, \bar{z})$. Le rette r ed s sono allora incidenti e si incontrano nel punto $\bar{P} \equiv (\bar{x}, \bar{y}, \bar{z})$. Vale anche il viceversa, nel senso che se sappiamo che r ed s si incontrano in un punto deduciamo che $\rho(A) = 3$, $\rho(A') = 3$. Il *Caso 3* è esemplificato nella seguente figura.



- *Caso 4.* Supponiamo infine che $\rho(A) = 3$, $\rho(A') = 4$, ovvero $\det(A') \neq 0$. Poiché se due rette sono complanari, ovvero esiste un piano che le contiene entrambe, esse risultano, come già noto, incidenti o parallele oppure coincidenti, e nel caso in esame nessuna di queste situazioni si verifica, le due rette r ed s saranno sghembe, ovvero non esiste alcun piano che le contiene entrambe. Anche in questo caso vale il viceversa. Osserviamo che, nel *Caso 4*, il fatto che $\rho(A') = 4$ implica che $\rho(A) = 3$, e quindi r, s sono sghembe se e solo se $\det A' \neq 0$. Il *Caso 4* è esemplificato nella seguente figura.



7.9 Questioni metriche nello spazio

7.22 Definizione. *Due rette r', r'' sono ortogonali (e si scrive $r' \perp r''$) se $\mathbf{u}_{r'} \perp \mathbf{u}_{r''}$. Quando si verifica tale situazione, può avvenire che r' ed r'' siano incidenti oppure sghembe. Due rette ortogonali ed incidenti si dicono perpendicolari.*

7.23 Definizione. *Due piani π', π'' sono ortogonali (e si scrive $\pi' \perp \pi''$) se $\mathbf{n}_{\pi'} \perp \mathbf{n}_{\pi''}$.*

7.24 Definizione. *Si dice che r e π sono ortogonali se $\mathbf{u}_r$ è ortogonale a π, ovvero a $\overrightarrow{\pi}$. Scriveremo allora $r \perp \pi$.*

Osserviamo esplicitamente che ciò equivale a dire che $\mathbf{u}_r \parallel \mathbf{n}_\pi$, o, equivalentemente, che, posto $\mathbf{u}_r = (l, m, n)$ e $\mathbf{n}_\pi = (a, b, c)$, si abbia $(l, m, n) \propto (a, b, c)$, ovvero

$$\rho \begin{pmatrix} a & b & c \\ l & m & n \end{pmatrix} = 1 \ .$$

Riassumendo, si ha che

$$r \perp \pi \iff \mathbf{u}_r \parallel \mathbf{n}_\pi \iff (l, m, n) \propto (a, b, c) \iff \rho \begin{pmatrix} l & m & n \\ a & b & c \end{pmatrix} = 1 \ .$$

Osserviamo anche che se $r \perp \pi$, allora r e π sono incidenti. È chiaro che tali definizioni non dipendono dalla scelta dei vettori direzionali delle rette r, r', r'' e dei vettori normali dei piani π, π', π''. Possiamo definire l'angolo tra le due rette r', r'' e tra i due piani π', π'' ponendo

$$\widehat{r'r''} = \min\left\{ \widehat{\mathbf{u}_{r'}\mathbf{u}_{r''}}, \pi - \widehat{\mathbf{u}_{r'}\mathbf{u}_{r''}} \right\} \quad ; \quad \widehat{\pi'\pi''} = \min\left\{ \widehat{\mathbf{n}_{\pi'}\mathbf{n}_{\pi''}}, \pi - \widehat{\mathbf{n}_{\pi'}\mathbf{n}_{\pi''}} \right\} \ .$$

La definizione di angolo tra due rette appena data ha senso anche quando le rette in questione non sono incidenti, ovvero $r' \cap r'' = \emptyset$.

Consideriamo ora due punti $P_0 \equiv (x_0, y_0 =, z_0)$, $P_1 \equiv (x_1, y_1 =, z_1)$, due rette r', r'' e due piani π', π'' (come in (7.16)). Per quanto riguarda la distanza tra due punti, osserviamo che, per definizione,

$$d(P_0, P_1) = \left|\overrightarrow{P_0P_1}\right| = \sqrt{\overrightarrow{P_0P_1} \cdot \overrightarrow{P_0P_1}}$$

e quindi, poiché $\overrightarrow{P_0P_1} = (x_1 - x_0, y_1 - y_0, z_1 - z_0)$,

$$d(P_0, P_1) = \sqrt{(x_1 - x_0)^2 + (y_1 - y_0)^2 + (z_1 - z_0)^2} \ .$$

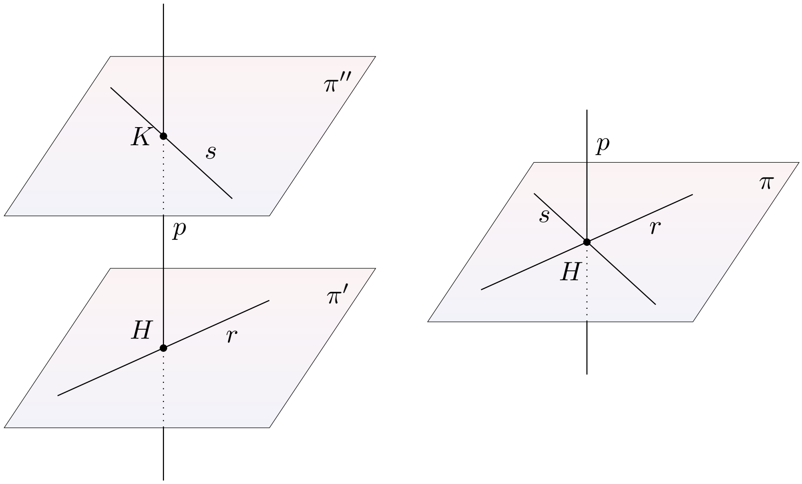
Occupiamoci ora della distanza di un punto P_0 da un piano π'. Descritta la retta r passante per P_0 e tale che $\mathbf{u}_r = \mathbf{n}_{\pi'}$, abbiamo che $r \perp \pi'$. Posto $H = r \cap \pi'$ si verifica facilmente che

$$d(P_0, H) \leq d(P_0, P) \quad \forall\ P \in \pi' \ .$$

Pertanto
$$d(P_0, \pi') = d(P_0, H)$$
e si prova che
$$d(P_0, \pi') = \frac{|a'x + b'y + c'z + d'|}{\sqrt{a'^2 + b'^2 + c'^2}} \ .$$
Studiamo la distanza tra i due piani π', π''. Se essi non sono paralleli, la loro intersezione è non vuota (è una retta) e quindi $d(\pi', \pi'') = 0$. Se invece $\pi' \parallel \pi''$, si procede al modo seguente. Se r è una qualunque retta ortogonale a entrambi i piani, detti P', P'' i punti di intersezione di r con tali piani, si ha che
$$d(\pi', \pi'') = d(P', P'') \ .$$
Consideriamo ora i sottospazi r', π' e calcoliamo $d(r', \pi')$. Se r' e π' sono incidenti è chiaro che $d(r', \pi') = 0$. Se invece $r' \parallel \pi'$, è altrettanto chiaro che
$$d(r', \pi') = d(P', \pi') \quad \forall \ P' \in r' \ .$$
La distanza tra il punto P_0 e la retta r' si calcola invece come segue. Esiste un unico piano π passante per P_0 e ortogonale ad r'. Posto allora $H = r' \cap \pi$, si ha che
$$d(P_0, r') = d(P_0, H) \ .$$
Consideriamo ora due rette r, s. Se $r \parallel s$, scelto un piano π tale che $\mathbf{n}_\pi = \mathbf{u}_r$ e posto $H = \pi \cap r$, $K = \pi \cap s$, si ha che
$$d(r, s) = d(H, K) \ .$$
Se invece r ed s sono non parallele, vale il seguente importante teorema.

7.25 Teorema della comune perpendicolare. *Siano r ed s due rette non parallele. Esiste allora un'unica retta p perpendicolare ad entrambe (che viene detta la comune perpendicolare di r ed s). Inoltre, posto $H = p \cap r$, $K = p \cap s$, si ha che $d(r, s) = d(H, K)$.*

Dimostrazione. Allo scopo di provare tale teorema, distinguiamo i due casi che possono presentarsi, ovvero che r, s siano incidenti oppure sghembe. Se r, s sono incidenti, e H è il loro punto comune, poiché un vettore ortogonale alle direzioni di r ed s è dato dal prodotto vettoriale $\mathbf{v} = \mathbf{u}_r \wedge \mathbf{u}_s$, la retta p passante per H con vettore direzionale $\mathbf{v}$ sarà certamente una perpendicolare comune ad r ed s. Essa è unica. Infatti ogni altra retta p' ortogonale sia ad r che ad s deve necessariamente essere parallela a p. Pertanto p' interseca il piano π che contiene r ed s in un unico punto, e tale punto non può che essere H, e quindi $p' = p$. È poi chiaro che in tal caso la distanza tra le due rette è nulla.



Supponiamo ora che le due rette siano sghembe e consideriamo una loro rappresentazione parametrica. Ad esempio, sia r come in (7.20) e sia

$$s : \begin{cases} x = x_1 + t'l' \\ y = y_1 + t'm' \\ z = z_1 + t'n' \end{cases}$$

I vettori direzionali delle due rette sono quindi $\mathbf{u}_r = (l, m, n)$ e $\mathbf{u}_s = (l', m', n')$ e si ha che

$$\rho \begin{pmatrix} l & m & n \\ l' & m' & n' \end{pmatrix} = 2$$

essendo $r \not\parallel s$. Indichiamo con P_t il generico punto di r ed analogamente sia $Q_{t'}$ il generico punto di s. Avremo che

$$P_t \equiv (x_0 + lt, y_0 + mt, z_0 + nt) \quad ; \quad Q_{t'} \equiv (x_1 + l't', y_1 + m't', z_1 + n't') \ .$$

In particolare $P_0 \equiv (x_0, y_0, z_0)$ e $Q_0 \equiv (x_1, y_1, z_1)$. Indichiamo con p la retta per P_t e $Q_{t'}$. Essa è, per costruzione, incidente con r ed s (nei punti P_t e $Q_{t'}$, rispettivamente) ed ha vettore direzionale $\mathbf{v} = \overrightarrow{P_tQ_{t'}}$. Pertanto

$$\mathbf{v} = \big(-lt + l't' - x_0 + x_1 \, , \, -mt + m't' - y_0 + y_1 \, , \, -nt + n't' - z_0 + z_1 \big) \ .$$

La retta p risulterà quindi una comune perpendicolare ad r ed s se e solo se $\mathbf{v} \perp \mathbf{u}_r$ e $\mathbf{v} \perp \mathbf{u}_s$, ovvero

$$\begin{cases} \mathbf{v} \cdot \mathbf{u}_r = 0 \\ \mathbf{v} \cdot \mathbf{u}_s = 0 \end{cases}$$

Tenendo conto dell'espressione di $\mathbf{v}, \mathbf{u}_r, \mathbf{u}_s$ ed effettuando i prodotti scalari su indicati, otteniamo il sistema

$$\begin{cases} l(-lt + l't' - x_0 + x_1) + m(-mt + m't' - y_0 + y_1) \\ \qquad\qquad + n(-nt + n't' - z_0 + z_1) = 0 \\ l'(-lt + l't' - x_0 + x_1) + m'(-mt + m't' - y_0 + y_1) \\ \qquad\qquad + n'(-nt + n't' - z_0 + z_1) = 0 \end{cases}$$

e cioè

$$\begin{cases} -(l^2 + m^2 + n^2)t + (ll' + mm' + nn')t' + l(x_1 - x_0) \\ \qquad\qquad + m(y_1 - y_0) + n(z_1 - z_0) = 0 \\ -(ll' + mm' + nn')t + (l'^2 + m'^2 + n'^2)t' + l'(x_1 - x_0) \\ \qquad\qquad + m'(y_1 - y_0) + n'(z_1 - z_0) = 0 \end{cases}$$

Tale sistema, di due equazioni nelle due incognite t, t', può anche scriversi nella forma più compatta

$$\begin{cases} -(\mathbf{u}_r \cdot \mathbf{u}_r)t + (\mathbf{u}_r \cdot \mathbf{u}_s)t' + \mathbf{u}_r \cdot \overrightarrow{P_0Q_0} = 0 \\ -(\mathbf{u}_r \cdot \mathbf{u}_s)t + (\mathbf{u}_s \cdot \mathbf{u}_s)t' + \mathbf{u}_s \cdot \overrightarrow{P_0Q_0} = 0 \end{cases} \tag{7.28}$$

Le matrici del sistema (7.28) sono

$$A' = \left(\underbrace{\begin{matrix} -\mathbf{u}_r \cdot \mathbf{u}_r & \mathbf{u}_r \cdot \mathbf{u}_s \\ -\mathbf{u}_r \cdot \mathbf{u}_s & \mathbf{u}_s \cdot \mathbf{u}_s \end{matrix}}_{A} \middle| \begin{matrix} \mathbf{u}_r \cdot \overrightarrow{P_0Q_0} \\ \mathbf{u}_s \cdot \overrightarrow{P_0Q_0} \end{matrix} \right)$$

Poiché $\det A = -(\mathbf{u}_r \cdot \mathbf{u}_r)(\mathbf{u}_s \cdot \mathbf{u}_s) + (\mathbf{u}_r \cdot \mathbf{u}_s)^2$, tenuto conto che i vettori $\mathbf{u}_r, \mathbf{u}_s$ sono non paralleli, e quindi indipendenti, dalla disuguaglianza di Schwarz segue che $\det A \neq 0$ e quindi il sistema (7.28) è di Cramer nelle incognite t, t'. Sia dunque (t_0, t_1') la soluzione (unica) di tale sistema. Allora il punto $H = P_{t_0} \in r$ e il punto $K = Q_{t_1'} \in s$ sono tali che la retta p per H, K è perpendicolare ad r ed s. Dall'unicità della soluzione (t_0, t_1') si deduce che tale retta p (che per costruzione è una comune incidente) è l'unica comune perpendicolare di r, s. È agevole poi la verifica del fatto che $d(r, s) = d(H, K)$. I punti $H \in r$ e $K \in s$ sono detti punti di minima distanza tra r ed s. □

7.10 Ampliamento complesso e proiettivo

Nel prossimo capitolo studieremo le coniche, ovvero le curve del piano descritte da equazioni di secondo grado. A tal fine è opportuno utilizzare un modello di piano in qualche modo più ampio di $\mathbb{E}^2$. La costruzione di tale ampliamento prevede dettagli formali che vanno ben al di là degli scopi del presente testo, e quindi sarà svolta in un modo forse formalmente insoddisfacente, ma, si spera, più intuitivo. Sappiamo che, in un riferimento fissato $\mathcal{R} = (O; \mathbf{e}_1, \mathbf{e}_2)$

ogni punto è individuato da due coordinate, ovvero due numeri reali, a, b. Immaginiamo di aggiungere altri punti, le cui coordinate siano numeri complessi, di cui almeno uno non reale. Ad esempio sarà possibile considerare il punto $Q \equiv (1, i)$. L'insieme così ottenuto sarà indicato con il simbolo $\mathbb{E}^2_{\mathbb{C}}$ e sarà detto *ampliamento complesso* del piano reale affine (ed euclideo) $\mathbb{E}^2$. I punti di $\mathbb{E}^2$ saranno detti punti reali, gli altri saranno detti punti immaginari. Osserviamo che se applichiamo le usuali formule di cambiamento di coordinate da $\mathcal{R}$ ad un altro riferimento $\mathcal{R}' = (O'; \mathbf{e}'_1, \mathbf{e}'_2)$, un punto di coordinate reali (a, b) in $\mathcal{R}$ avrà coordinate (a', b') in $\mathcal{R}'$ che risulteranno ancora reali. Analogamente, se le coordinate del punto in questione sono non entrambe reali in $\mathcal{R}$, tali saranno anche in $\mathcal{R}'$. In altre parole un punto sarà reale, oppure immaginario, indipendentemente dal riferimento di $\mathbb{E}^2$ scelto, in quanto sono reali le matrici coinvolte nelle formule di cambiamento delle coordinate. Per essere più precisi, applichiamo la Proposizione 7.14 al nostro contesto. Supponiamo che

$$P \equiv_{\mathcal{R}} (x, y) \ ; \quad P \equiv_{\mathcal{R}'} (x', y') \ . \tag{7.29}$$

Supponiamo inoltre che

$$\mathbf{e}_1 = a_{1,1}\mathbf{e}'_1 + a_{1,2}\mathbf{e}'_2 \ ; \quad \mathbf{e}_2 = a_{2,1}\mathbf{e}'_1 + a_{2,2}\mathbf{e}'_2 \ ; \quad O \equiv_{\mathcal{R}'} (a_{1,3}, a_{2,3}) \ . \tag{7.30}$$

Poniamo

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix} \ ; \quad B = \begin{pmatrix} a_{1,3} \\ a_{2,3} \end{pmatrix} \ ; \quad A' = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \end{pmatrix} \ . \tag{7.31}$$

Allora la Proposizione 7.14 ci assicura che

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = A \cdot \begin{pmatrix} x \\ y \end{pmatrix} + B = \begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} a_{1,3} \\ a_{2,3} \end{pmatrix} \tag{7.32}$$

ovvero

$$\begin{cases} x' = a_{1,1}x + a_{1,2}y + a_{1,3} \\ y' = a_{2,1}x + a_{2,2}y + a_{2,3} \end{cases} \tag{7.33}$$

Conviene ricordare che se, come spesso accadrà, i riferimenti sono ortonormali, la matrice A risulta ortogonale.

Consideriamo un punto $Q \equiv (z, w)$. Le coordinate z, w sono numeri complessi. Possiamo quindi considerare i loro coniugati $\bar{z}$, e $\bar{w}$. Il punto $\bar{Q} = (\bar{z}, \bar{w})$ si dice complesso coniugato di Q. Osserviamo che Q è un punto reale se e solo se $\bar{Q} = Q$. L'applicazione

$$\chi : \mathbb{E}^2_{\mathbb{C}} \longrightarrow \mathbb{E}^2_{\mathbb{C}}$$

che associa ad ogni punto Q il suo coniugato $\bar{Q}$ si dice *coniugio*. Essa fissa i punti reali di $\mathbb{E}^2_{\mathbb{C}}$ e si ha che $\chi\chi = id_{\mathbb{E}^2_{\mathbb{C}}}$.

Consideriamo ora una retta r di $\mathbb{E}^2$, di equazione

$$r : ax + by + c = 0 \tag{7.34}$$

ed il luogo geometrico r' dei punti di $\mathbb{E}^2_{\mathbb{C}}$ le cui coordinate (eventualmente anche immaginarie) soddisfano la (7.34). Il luogo r' è costituito dai punti di r e da altri eventuali punti immaginari. r' si dice ampliamento complesso della retta reale r. Ad esempio, se

$$r : x - y + 1 = 0 \tag{7.35}$$

abbiamo che $P' \equiv (0, 1)$, $P'' \equiv (-1, 0)$, $P'' \equiv (2, 3)$ sono punti di r, ovvero sono punti reali di r', mentre $Q' \equiv (i, i + 1)$ è un punto immaginario di r'. La distinzione tra r e r', laddove non indispensabile, sarà omessa. In $\mathbb{E}^2_{\mathbb{C}}$ possiamo considerare luoghi geometrici descritti da equazioni di primo grado in x, y, ma con coefficienti non necessariamente reali. Chiameremo tali luoghi ancora rette. Ad esempio, sia

$$s : x + iy = 0\ . \tag{7.36}$$

Osserviamo che il punto $O \equiv (0, 0)$ è reale ed appartiene ad s, ma è l'unico punto reale di s. Altri punti, necessariamente immaginari, di s sono $R' \equiv (i, -1)$, $R'' \equiv (1, i)$, $R''' \equiv (3 + i, 3i - 1)$. Una retta descritta mediante un'equazione a coefficienti reali è detta *retta reale*. Può accadere che una retta sia descritta da un'equazione a coefficienti non reali, ma sia comunque una retta reale. Ad esempio, l'equazione

$$ix - iy + i = 0$$

descrive ancora la retta r della (7.35). È infatti vero, in generale, che equazioni proporzionali descrivono lo stesso luogo geometrico (ovvero hanno le stesse soluzioni). Diremo quindi che una retta è *immaginaria* se non può essere descritta da un'equazione a coefficienti reali. Non è difficile verificare che una retta immaginaria possiede al più un punto reale, mentre una retta reale possiede infiniti punti reali. Quindi una retta che possiede almeno due punti reali è reale (e possiede infiniti punti reali).

Sia r una retta di equazione (7.34) (reale o immaginaria). I coefficienti a, b, c dell'equazione sono quindi numeri complessi. Indicheremo con $\bar{r}$ la retta

$$\bar{r} : \bar{a}x + \bar{b}y + \bar{c} = 0\ ,$$

dove $\bar{a}, \bar{b}, \bar{c}$ sono i complessi coniugati di a, b, c. Diremo che $\bar{r}$ è la retta complessa coniugata di r. È interessante notare che una retta è reale se e solo se coincide con la sua coniugata.

Osserviamo che anche il concetto di vettore libero, o applicato, va ampliato. Avremo quindi vettori di componenti non necessariamente reali. Ad esempio, il vettore $\overrightarrow{R'R''}$ avrà componenti $(2+i, -1+2i)$, e sarà un vettore direzionale di s. Va sottolineato che una retta immaginaria può anche avere vettore direzionale reale. È il caso, ad esempio, della retta

$$t : x + i = 0$$

che è immaginaria ed ha vettore direzionale $\mathbf{u}_t = (0, -1)$, o anche $(0, 1)$. Il concetto di parallelismo si estende all'insieme $\mathcal{L}$ di tutte le rette (reali o immaginarie) di $\mathbb{E}^2_{\mathbb{C}}$. Ad esempio vediamo che la retta immaginaria t è parallela all'asse delle y, descritto dall'equazione $x = 0$, che è una retta reale. Sottolineiamo il fatto che i vettori e i punti che compaiono nei riferimenti che noi selezioniamo sono tutti reali. Parleremo pertanto di *riferimenti reali.*

Consideriamo ora l'insieme $\mathcal{L}$ delle rette di $\mathbb{E}^2_{\mathbb{C}}$. La relazione $\parallel$ di parallelismo è un'equivalenza in $\mathcal{L}$. Ricordiamo che se $r', r'' \in \mathcal{L}$ e $\mathbf{u}_{r'} = (l', m')$, $\mathbf{u}_{r''} = (l'', m'')$ sono vettori direzionali di tali rette, abbiamo che

$$r' \parallel r'' \iff (l', m') \propto (l'', m'') \ .$$

Un elemento dell'insieme quoziente $\mathcal{L}/_{\parallel}$, ovvero una classe di rette parallele, è detto anche talvolta, *direzione*. Definiamo un nuovo insieme

$$\hat{\mathbb{E}}^2 = \mathbb{E}^2_{\mathbb{C}} \bigcup \mathcal{L}/_{\parallel} \ .$$

Un elemento di $\hat{\mathbb{E}}^2$ sarà detto punto. Se esso appartiene ad $\mathbb{E}^2_{\mathbb{C}}$, sarà un *punto proprio*, se invece appariene a $\mathcal{L}/_{\parallel}$, ovvero è una direzione, sarà detto anche *punto improprio.*

Sia ora $\mathbb{K} = \mathbb{R}$ oppure $\mathbb{K} = \mathbb{C}$. Dalla potenza cartesiana $\mathbb{K}^3$ escludiamo l'origine, e consideriamo quindi l'insieme

$$\mathbb{X} = \mathbb{K}^3 - \{\mathbf{0}\} \ .$$

In $\mathbb{X}$ consideriamo la relazione $\propto$ di proporzionalità: se $(a, b, c), (a', b', c') \in \mathbb{X}$, poniamo

$$(a, b, c) \propto (a', b', c') \iff \exists\, \lambda \in \mathbb{K} - \{0\} \,\big|\, (a', b', c') = \lambda(a, b, c) \ .$$

Osserviamo che $\propto$ è una relazione d'equivalenza in $\mathbb{X}$. Indicheremo con $\mathbb{P}^2_{\mathbb{K}}$ l'insieme quoziente $\mathbb{X}/\propto$, che è noto come *piano proiettivo numerico* (reale o complesso). Se $(a, b, c) \in \mathbb{X}$, la sua classe si indicherà con il simbolo $[a, b, c]$. Si potrebbe verificare che tale classe, ovvero l'insieme delle terne non nulle proporzionali ad (a, b, c), è la retta per l'origine e per (a, b, c), privata dell'origine, o anche il sottospazio vettoriale di $\mathbb{K}^3$ generato da (a, b, c), privato del vettore nullo. Pertanto talvolta si dice che il piano proiettivo può vedersi, geometricamente, come l'insieme delle rette dello spazio passanti per l'origine, o anche come l'insieme delle direzioni nello spazio.

Consideriamo un punto $P \in \mathbb{E}^2_{\mathbb{C}}$, ovvero, con la terminologia appena introdotta, un punto proprio di $\hat{\mathbb{E}}^2$, di coordinate, che d'ora in avanti chiameremo *affini*, (a, b). Diremo che P ha come coordinate *omogenee* la terna $(a, b, 1)$, o una qualunque altra terna ad essa proporzionale (che risulterà necessariamente non nulla). Quindi la coppia di coordinate affini di un punto è univocamente determinata, mentre la terna di coordinate omogenee dello stesso punto è definita solo a meno di proporzionalità. Comunque l'ultima coordinata omogenea non sarà mai nulla. Possiamo anche dire che le coordinate omogenee di P sono date dalla classe $[a, b, 1] \in \mathbb{P}^2_{\mathbb{C}}$. Scriveremo $P \equiv [a, b, 1]$. Ad esempio l'origine O ha coordinate affini $(0, 0)$ e coordinate omogenee date da una qualunque terna del tipo $(0, 0, k)$ dove k è uno scalare non nullo, o più semplicemente O ha coordinate omogenee $[0, 0, 1]$ (o anche $[0, 0, 37]$!!) Se il punto $P' \equiv (3, 5)$, ovvero ha coordinate affini $(3, 5)$, avremo che $P' \equiv [3, 5, 1]$, o anche $P' \equiv [6, 10, 2]$. Consideriamo ora un punto $P \in \mathcal{L}/_{\parallel}$, ovvero, con la terminologia appena introdotta, un punto improprio di $\hat{\mathbb{E}}^2$. In altri termini $P = [r]$ è una classe di rette parallele, rappresentata, ad esempio, dalla retta r. Sia $\mathbf{u}_r = (l, m)$ un vettore direzionale di r. Diremo che P ha coordinate omogenee $[l, m, 0]$, ovvero che una sua terna di coordinate omogenee è data da una qualunque terna $(\alpha, \beta, 0)$, dove α, β sono numeri direzionali di r (che come è noto sono definiti solo a meno di proporzionalità). Se r è la retta descritta dalla (7.35), abbiamo che $\mathbf{u}_r = (1, 1)$, e se $P = [r] \in \mathcal{L}/_{\parallel} \subset \hat{\mathbb{E}}^2$, le coordinate omogenee di P sono $[1, 1, 0]$. Abbiamo in tal modo costruito un'applicazione

$$\hat{c} : \hat{\mathbb{E}}^2 \longrightarrow \mathbb{P}^2_{\mathbb{C}}$$

che associa ad ogni punto di $\hat{\mathbb{E}}^2$ un elemento di $\mathbb{P}^2_{\mathbb{C}}$, che è biettiva.

Se $P' \in \mathbb{E}^2_{\mathbb{C}}$ ha coordinate affini (x', y'), ogni terna (x'_1, x'_2, x'_3) di coordinate omogenee di P' sarà tale che $x'_3 \neq 0$ e $(x', y', 1) \propto (x'_1, x'_2, x'_3)$, e ciò equivale a dire che

$$x' = \frac{x'_1}{x'_3} \; ; \; y' = \frac{x'_2}{x'_3} \; ; \; x'_3 \neq 0 \; . \tag{7.37}$$

Le (7.37) ci consentono agevolmente di passare dalle coordinate affini a quelle omogenee e viceversa. Osserviamo che se, invece, $P' \in \mathcal{L}/_{\parallel}$, e cioè P' è una classe di rette parallele, se $P' \equiv [l', m', 0]$ vuol dire che (l', m') è un vettore direzionale di una qualunque delle rette che rappresentano P'.

Consideriamo la generica retta r di equazione (7.34) e l'equazione omogenea di primo grado in x_1, x_2, x_3

$$ax_1 + bx_2 + cx_3 = 0 \; . \tag{7.38}$$

Osserviamo che il punto $P' \in \mathbb{E}^2_{\mathbb{C}}$ di coordinate affini (x', y') e di coordinate omogenee (x'_1, x'_2, x'_3), con $x'_3 \neq 0$, appartiene ad r se e solo se la coppia (x', y') è una soluzione della (7.34), ma anche se e solo se la terna (x'_1, x'_2, x'_3) è una soluzione della (7.38). Però la (7.38) è soddisfatta anche dalla terna $(-b, a, 0)$ (ed ogni altra ad essa proporzionale) che rappresenta la direzione di r, che è un punto improprio di $\hat{\mathbb{E}}^2_{\mathbb{C}}$ che possiamo indicare, ad esempio, con il simbolo P_r. Pertanto la (7.38) rappresenta in $\hat{\mathbb{E}}^2_{\mathbb{C}}$ il luogo geometrico $r \cup \{P_r\}$, che indicheremo con il simbolo $\hat{r}$ e chiameremo *retta ampliata.* Scriveremo quindi

$$\hat{r} : ax_1 + bx_2 + cx_3 = 0 \; .$$

Osserviamo esplicitamente che anche la terna nulla $(0, 0, 0)$ è soluzione della (7.38), ma non va considerata, in quanto, nel nostro contesto, essa non ha significato geometrico, nel senso che non rappresenta in coordinate omogenee alcun punto di $\hat{\mathbb{E}}^2_{\mathbb{C}}$.

È interessante sottolineare che se r, s sono due rette distinte in $\mathbb{E}^2_{\mathbb{C}}$ e $\hat{r}, \hat{s}$ sono le corrispondenti rette ampliate in $\hat{\mathbb{E}}^2_{\mathbb{C}}$, accade che r, s possono essere incidenti, ed in tal caso ammettono un punto di intersezione, oppure parallele (propriamente), ed in tal caso non si intersecano. Invece in ogni caso $\hat{r}, \hat{s}$ si intersecano in un punto, che sarà un punto proprio quando r, s sono incidenti, e tale punto coincide con il punto di incidenza di r, s, e sarà invece un punto improprio nel caso in cui r, s siano parallele, ed allora il punto di intersezione sarà proprio la direzione comune delle due rette. Osserviamo infine che, in coordinate omogenee, affinché la (7.38) sia effettivamente una equazione di primo grado, occorre e basta che gli scalari a, b, c non si annullino simultaneamente. Ha pertanto senso l'equazione

$$x_3 = 0$$

che rappresenta il luogo di tutti i punti impropri di $\hat{\mathbb{E}}^2_{\mathbb{C}}$. Tale luogo si indica con il simbolo r_∞ ed è noto come la *retta impropria* del piano ampliato. Quando non sarà strettamente necessario, indicheremo con lo stesso simbolo una retta r e la corrispondente retta ampliata $\hat{r}$.

Concludiamo il capitolo con alcune osservazioni sulle formule di cambiamento delle coordinate non affini ma omogenee. Consideriamo la situazione descritta nelle formule (7.29) e successive, e definiamo una nuova matrice, che chiamiamo *matrice completa del cambiamento*

$$\tilde{A} = \left(\begin{array}{c|c} A & B \\ \hline 0\ 0 & 1 \end{array}\right) . \tag{7.39}$$

Si verifica facilmente che

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \tilde{A} \cdot \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} .$$

Pertanto, se, in termini di coordinate omogenee, abbiamo che $P \equiv_{\mathcal{R}} [x_1, x_2, x_3]$ e $P \equiv_{\mathcal{R}'} [x'_1, x'_2, x'_3]$, la terna di coordinate omogenee $\begin{pmatrix} x'_1 \\ x'_2 \\ x'_3 \end{pmatrix}$ è data, a meno di proporzionalità, dal prodotto $\tilde{A} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$, ovvero esiste uno scalare non nullo k tale che

$$k \begin{pmatrix} x'_1 \\ x'_2 \\ x'_3 \end{pmatrix} = \tilde{A} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} .$$

Capitolo 8

Coniche

8.1 Circonferenza, ellisse, iperbole, parabola

Fissato un punto $C \equiv (x_0, y_0)$ ed un numero reale positivo h, si definisce *circonferenza* di centro C e raggio h il luogo γ dei punti del piano che distano h da C, ovvero

$$\gamma = \left\{ R \mid d(R, C) = h \right\} .$$

Poiché, posto $R \equiv (x_1, y_1)$, si ha che

$$d(R, C) = \left|\overrightarrow{RC}\right| = \sqrt{(x_1 - x_0)^2 + (y_1 - y_0)^2} ,$$

il luogo γ è rappresentato dall'equazione

$$\sqrt{(x_1 - x_0)^2 + (y_1 - y_0)^2} = h$$

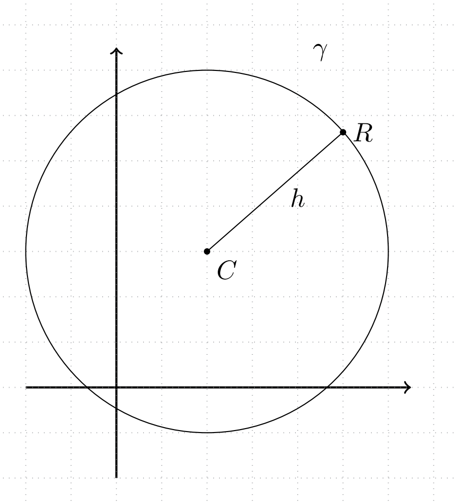
o anche

$$(x_1 - x_0)^2 + (y_1 - y_0)^2 = h^2$$

che è del tipo

$$x^2 + y^2 + 2ax + 2by + c = 0 , \ (c < a^2 + b^2) ,$$

dove $a = -x_0$, $b = -y_0$, $h = \sqrt{a^2 + b^2 - c}$. Si verifica che ogni equazione di questo tipo rappresenta una circonferenza, il cui centro ha coordinate $(-a, -b)$, di raggio $h = \sqrt{a^2 + b^2 - c}$.



$$C \equiv (2,3),\ \ h = 4$$
$$\gamma : x^2 + y^2 - 2x - 3y - 3 = 0$$

Osserviamo che se $h = 0$ otteniamo il luogo costituito dal solo punto reale C, che possiamo considerare una circonferenza degenere di raggio nullo. Se invece $h < 0$, si tratta di una circonferenza immaginaria, nel senso che il luogo γ è privo di punti reali, ma dotato di punti immaginari. Ad esempio, al luogo

$$\gamma : x^2 + y^2 + 1 = 0$$

appartiene il punto $A \equiv (i, 0)$.

Descriviamo ora la costruzione dell'ellisse. Consideriamo due punti distinti F', F'' e fissiamo, per comodità, il riferimento che ha come origine O il punto medio di $F'F''$, asse delle ascisse passante per F', F'', con F'' nel semiasse positivo, ed asse delle ordinate passante per O, ortogonale all'asse delle ascisse, con l'ovvio orientamento. Sia, ad esempio, $d(F', F'') = 2c$. Ciò vuol dire che $F' \equiv (-c, 0)$ e $F'' \equiv (c, 0)$. Sia $a > c$. Consideriamo il luogo γ dei punti $R \equiv (x_1, y_1)$ tali che la somma delle distanze di R da F' e F'' sia $2a$, ovvero

$$\gamma = \left\{ R \;\middle|\; d(R, F') + d(R, F'') = 2a \right\} .$$

Abbiamo che

$$R \in \gamma \iff \sqrt{(x_1 + c)^2 + y_1^2} + \sqrt{(x_1 - c)^2 + y_1^2} = 2a .$$

Con una quadratura si ottiene

$$(x_1 + c)^2 + y_1^2 + (x_1 - c)^2 + y_1^2 + 2\sqrt{(x_1 + c)^2 + y_1^2}\sqrt{(x_1 - c)^2 + y_1^2} = 4a^2$$

ovvero

$$(x_1 + c)^2 + y_1^2 + (x_1 - c)^2 + y_1^2 - 4a^2 = -2\sqrt{(x_1 + c)^2 + y_1^2}\sqrt{(x_1 - c)^2 + y_1^2} .$$

Con una ulteriore quadratura si ottiene

$$\big((x_1+c)^2+y_1^2+(x_1-c)^2+y_1^2-4a^2\big)^2=4\big((x_1+c)^2+y_1^2\big)\big((x_1-c)^2+y_1^2\big)\ ,$$

ovvero

$$(2x_1^2+2y_1^2+2c^2-4a^2)^2=4(x_1^2+y_1^2+c^2+2cx_1)(x_1^2+y_1^2+c^2-2cx_1)$$

o ancora

$$(x_1^2+y_1^2+c^2-2a^2)^2=(x_1^2+y_1^2+c^2)^2-4c^2x_1^2$$

o anche

$$(x_1^2+y_1^2+c^2)^2+4a^4-4a^2(x_1^2+y_1^2+c^2)=(x_1^2+y_1^2+c^2)^2-4c^2x_1^2$$

e quindi

$$4a^4-4a^2(x_1^2+y_1^2+c^2)=-4c^2x_1^2\ .$$

Riordinando e dividendo per 4, abbiamo

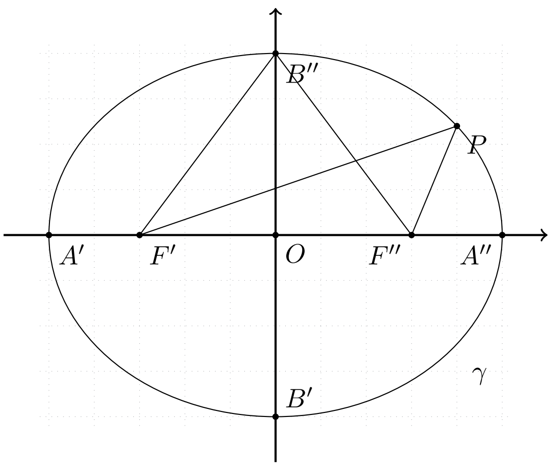
$$(a^2-c^2)x_1^2+a^2y_1^2=a^2(a^2-c^2)\ .$$

Poniamo allora $b=\sqrt{a^2-c^2}$, e cioè $b^2=a^2-c^2$, dividiamo per a^2b^2 e otteniamo

$$\frac{x_1^2}{a^2}+\frac{y_1^2}{b^2}=1\ .$$

Pertanto una rappresentazione di γ è la seguente:

$$\gamma:\frac{x^2}{a^2}+\frac{y^2}{b^2}=1\ . \tag{8.1}$$



$a=5,\ b=4,\ c=3$

$$\gamma:\frac{x^2}{25}+\frac{y^2}{16}=1$$

Il luogo γ si dice *ellisse* di centro O e semiassi a,b, e la (8.1) è la rappresentazione dell'ellisse in forma canonica. I punti $A'\equiv(-a,0)$, $A''\equiv(a,0)$, $B'\equiv(0,-b)$,

$B'' \equiv (0, b)$ di intersezione di γ con gli assi cartesiani si dicono *vertici*, mentre i punti F', F'' sono i *fuochi* dell'ellisse.

Studiamo ora la costruzione dell'iperbole. Utilizziamo gli stessi punti F', F'' e lo stesso riferimento di prima e consideriamo un numero reale positivo $a < c$. Consideriamo il luogo γ dei punti $R \equiv (x_1, y_1)$ tali che la differenza delle distanze tra R ed F', F'' (in valore assoluto) sia $2a$:

$$\gamma = \left\{ R \mid |d(R, F') - d(R, F'')| = 2a \right\} .$$

Si ha quindi che

$$R \in \gamma \iff \left| \sqrt{(x_1 + c)^2 + y_1^2} - \sqrt{(x_1 - c)^2 + y_1^2} \right| = 2a .$$

Con una doppia quadratura, in analogia con il caso dell'ellisse, abbiamo che

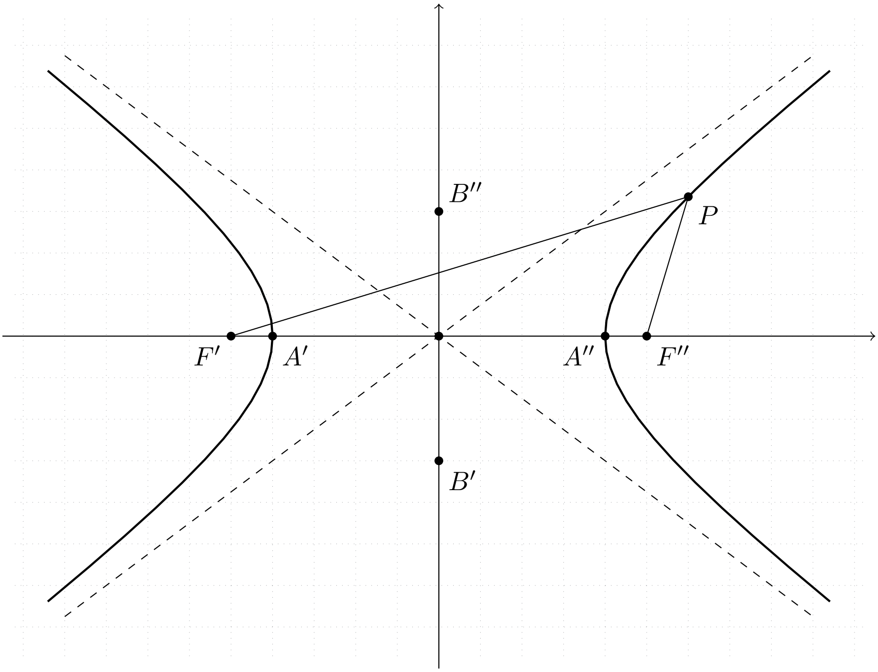
$$(c^2 - a^2)x_1^2 - a^2 y_1^2 = a^2(c^2 - a^2) .$$

Poniamo allora $b = \sqrt{c^2 - a^2}$, e cioè $b^2 = c^2 - a^2$, dividiamo per $a^2 b^2$ e otteniamo

$$\frac{x_1^2}{a^2} - \frac{y_1^2}{b^2} = 1 .$$

Il luogo γ così descritto si dice *iperbole* di centro O, semiasse maggiore a e semiasse minore b, ed è rappresentato dall'equazione

$$\gamma : \frac{x^2}{a^2} - \frac{y^2}{b^2} = 1 . \tag{8.2}$$



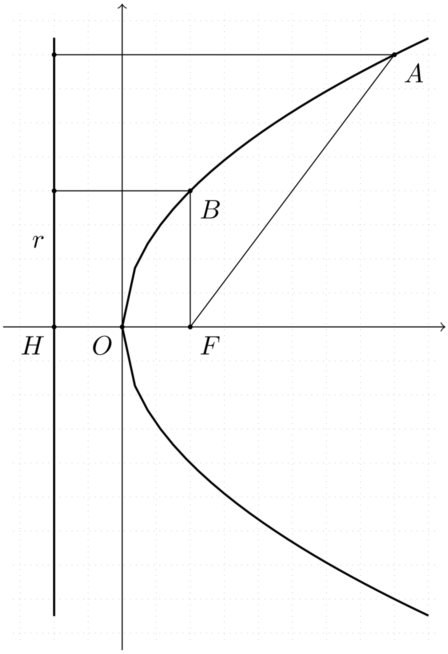
$$a = 4,\ b = 3,\ c = 5\ ,\ \gamma : \frac{x^2}{16} - \frac{y^2}{9} = 1$$

La (8.2) è la rappresentazione dell'iperbole in forma canonica. I punti F', F'' sono i *fuochi* dell'iperbole, i punti $A' \equiv (-a, 0)$, $A'' \equiv (a, 0)$ di intersezione di γ con l'asse delle ascisse sono i *vertici* reali dell'iperbole, ed infine le rette

$$s' : bx - ay = 0\ ,\ s'' : bx + ay = 0$$

sono gli *asintoti* di γ.

Consideriamo ora una retta r ed un punto $F \notin r$. Sia ad esempio $d(F, r) = p > 0$. La retta s per F perpendicolare ad r interseca r in un punto H. Sia O il punto medio del segmento di estremi H ed F, s l'asse delle ascisse, e la perpendicolare ad s passante per O l'asse delle ordinate, opportunamente orientati come in figura.



$$\gamma : y^2 = 8x$$
$$p = 4,\ F \equiv (2,0)$$
$$A \equiv (8,8),\ B \equiv (2,4)$$
$$d(A,F) = d(A,r) = 8$$
$$d(B,F) = d(B,r) = 4$$

Nel riferimento così ottenuto avremo che $F \equiv (\frac{p}{2}, 0)$. Consideriamo il luogo γ dei punti $R \equiv (x_1, y_1)$ equidistanti da F ed r. Avremo cioè

$$\gamma = \left\{ R \;\middle|\; d(R,F) = d(R,r) \right\} .$$

Poiché

$$d(R,F) = \sqrt{(x_1 - \frac{p}{2})^2 + y_1^2}\ ,\ d(R,r) = x_1 + \frac{p}{2}$$

avremo che

$$R \in \gamma \quad \Longleftrightarrow \quad \sqrt{(x_1 - \frac{p}{2})^2 + y_1^2} = x_1 + \frac{p}{2}\ ,$$

ovvero, con una quadratura,

$$(x_1 - \frac{p}{2})^2 + y_1^2 = (x_1 + \frac{p}{2})^2$$

e cioè

$$y_1^2 = 2px_1\ .$$

Il luogo così descritto si dice *parabola*, l'origine O appartiene ad essa e si dice *vertice*, F è il *fuoco*, r la *direttrice*. La parabola ammette quindi la seguente rappresentazione:

$$\gamma : y^2 = 2px\ . \tag{8.3}$$

La (8.3) è la rappresentazione della parabola in forma canonica.

8.2 Generalità sulle coniche

In questo capitolo useremo sempre le seguenti notazioni. Indicheremo con X, Y, W, Z i vettori colonna

$$X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \ ; \ Y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \ ; \ W = \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} \ ; \ Z = \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} \ .$$

Useremo le notazioni X^t, Y^t, W^t, Z^t per gli stessi vettori in notazione di riga. X rappresenterà la terna delle incognite, Y, W, Z le terne delle coordinate omogenee di punti del piano ampliato $\widehat{\mathbb{E}}^2_{\mathbb{C}}$, spesso indicati con i simboli R, S, T. In $\widehat{\mathbb{E}}^2_{\mathbb{C}}$ si consideri un riferimento ortonormale $\mathcal{R} = (O; \mathbf{e}_1, \mathbf{e}_2)$ reale. Abbiamo quindi una biezione

$$\widehat{c} : \widehat{\mathbb{E}}^2_{\mathbb{C}} \longrightarrow \mathbb{P}^2(\mathbb{C}) \ ,$$

come descritto alla fine del capitolo precedente. Se $R \in \widehat{\mathbb{E}}^2_{\mathbb{C}}$ e $\widehat{c}(R) = [y_1, y_2, y_3]$, scriveremo anche, per brevità, $R \equiv Y^t$.

Consideriamo lo spazio vettoriale numerico complesso di dimensione 3, $\mathbb{C}^3$, con una base ordinata fissata; per semplicità possiamo ad esempio scegliere quella canonica. Sia

$$f : \mathbb{C}^3 \times \mathbb{C}^3 \longrightarrow \mathbb{C}$$

una forma bilineare simmetrica (non nulla). Esisterà dunque una matrice simmetrica 3×3

$$B = \begin{pmatrix} b_{1,1} & b_{1,2} & b_{1,3} \\ b_{2,1} & b_{2,2} & b_{2,3} \\ b_{3,1} & b_{3,2} & b_{3,3} \end{pmatrix}$$

tale che se $Y, W \in \mathbb{C}^3$ si abbia

$$f(Y, W) = f(y_1, y_2, y_3; w_1, w_2, w_3) = Y^t \cdot B \cdot W = \sum_{i,j=1}^{3} b_{i,j} y_i w_j \ . \qquad (8.4)$$

8.1 DEFINIZIONE. *Si dice* conica *associata alla forma f il luogo geometrico γ_f, o anche γ, descritto dall'equazione*

$$\gamma : \sum_{i,j=1}^{3} b_{i,j} x_i x_j = 0 \qquad (8.5)$$

che è omogenea di secondo grado nelle incognite x_1, x_2, x_3.

8.2 DEFINIZIONE. *Se la matrice (non nulla) B è reale, o proporzionale ad una matrice reale, la conica γ si dice a sua volta reale.*

Osserviamo che, in virtù della simmetria di f, o di B, la (8.5) si riduce a

$$\gamma : b_{1,1}x_1^2 + 2b_{1,2}x_1x_2 + b_{2,2}x_2^2 + 2b_{1,3}x_1x_3 + 2b_{2,3}x_2x_3 + b_{3,3}x_3^2 = 0 \ . \qquad (8.6)$$

Poiché si è supposto che f non è nulla, non tutti i coefficienti della matrice B possono annullarsi simultaneamente. Quindi il rango di B è almeno 1 e la (8.5), o anche (8.6), è effettivamente un'equazione omogenea di secondo grado nelle incognite x_1, x_2, x_3. Osserviamo esplicitamente che la conica γ, pur essendo una conica *reale*, potrebbe non possedere alcun punto reale, oppure possedere un unico punto reale, o infiniti punti reali, mentre possiede sempre infiniti punti immaginari. Un esempio molto semplice di conica reale priva di punti reali è fornito dalla conica

$$\gamma : x_1^2 + x_2^2 + x_3^2 = 0 \ .$$

Infatti l'unica soluzione reale di tale equazione è data dalla terna nulla $(0,0,0)$, che non rappresenta, in coordinate omogenee, alcun punto di $\widehat{\mathbb{E}}^2_{\mathbb{C}}$.

Se γ è una conica reale definita dalla (8.6) e P' è un punto proprio di coordinate affini (x', y') e omogenee $[x_1', x_2', x_3']$ (con $x_3' \neq 0$), è chiaro che $P' \in \gamma$ se e solo se la terna (x_1', x_2', x_3') è soluzione della (8.6), ma è agevole verificare che $P' \in \gamma$ se e solo se la coppia (x', y') è soluzione della seguente equazione di secondo grado non omogenea

$$b_{1,1}x^2 + 2b_{1,2}xy + b_{2,2}y^2 + 2b_{1,3}x + 2b_{2,3}y + b_{3,3} = 0 \qquad (8.7)$$

in x, y, che chiameremo *equazione non omogenea associata a* γ. In altri termini, come già osservato nel caso delle rette nel piano affine e nel piano ampliato, abbiamo che il luogo descritto (in coordinate affini) dalla (8.7) coincide con la parte propria del luogo descritto (in coordinate omogenee) dalla (8.6). Però γ possiede anche alcuni punti impropri, che si possono studiare solo in coordinate omogenee mediante la (8.6). Tali punti avranno la terza coordinata omogenea nulla, e quindi costituiscono l'intersezione $r_\infty \cap \gamma$ della conica γ con la retta impropria r_∞. Osserviamo che se r è una qualunque retta reale nel piano ampliato $\widehat{\mathbb{E}}^2_{\mathbb{C}}$ e γ è una conica reale vale il seguente risultato, di cui omettiamo la dimostrazione.

8.3 Teorema. *L'intersezione $r \cap \gamma$ è di uno dei seguenti tipi:*

(a) $r \cap \gamma$ è un punto (reale);

(b) $r \cap \gamma$ è costituita da due punti reali;

(c) $r \cap \gamma$ è costituita da due punti immaginari;

(d) $r \cap \gamma$ coincide con r, ovvero $r \subseteq \gamma$.

Nel caso (a) diremo che r è tangente, nel caso (b) secante, nel caso (c) esterna alla conica. Nel caso (d) può accadere che $\gamma = r$ oppure che γ sia l'unione di

r e di un'altra retta. Il teorema precedente, applicato al caso $r = r_\infty$, ci dice che se r_∞ non è contenuta nella conica, γ possiede un unico punto improprio, oppure due punti impropri reali o infine due punti impropri immaginari.

La conica γ induce una relazione tra i punti del piano ampliato.

8.4 Definizione. *I punti $R \equiv Y^t$, $S \equiv W^t$ si dicono coniugati rispetto a γ, o anche rispetto ad f, se*

$$f(Y, W) = 0 \ .$$

Scriveremo allora $R \sim_f S$.

Alla luce di tale definizione, la conica γ può vedersi come il luogo dei punti autoconiugati, ovvero

$$\gamma = \left\{ R \in \widehat{\mathbb{E}}^2_{\mathbb{C}} \mid R \sim_f R \right\} \ .$$

Si verifica agevolmente che si si cambia riferimento ortonormale reale l'equazione di γ nel nuovo riferimento è ancora omogenea di secondo grado in tre incognite. Si verifica anche, ma non così agevolmente, che se g è un'altra forma simmetrica f e g rappresentano la stessa conica se e solo se sono proporzionali ordinatamente i loro coefficienti:

$$\gamma_f = \gamma_g \iff f \propto g \ .$$

8.5 Definizione. *Un punto $V \in \widehat{\mathbb{E}}^2_{\mathbb{C}}$ si dice punto doppio di γ se è coniugato ad ogni altro punto di $\widehat{\mathbb{E}}^2_{\mathbb{C}}$, ovvero*

$$V \sim_f R \, , \ \forall R \in \widehat{\mathbb{E}}^2_{\mathbb{C}} \ ,$$

o ancora, posto $V \equiv Z^t$,

$$f(Z, Y) = 0 \, , \ \forall Y \in \mathbb{C}^3 \ .$$

Indicheremo con $\mathcal{D}$ l'insieme dei punti doppi di γ.

Osserviamo che gli eventuali punti doppi di γ sono necessariamente reali, e appartengono alla conica.

8.6 Definizione. *Per ogni punto $R \in \widehat{\mathbb{E}}^2_{\mathbb{C}}$ definiamo un sottoinsieme $\omega(R) \subseteq \widehat{\mathbb{E}}^2_{\mathbb{C}}$ al modo seguente:*

$$\omega(R) = \left\{ S \in \widehat{\mathbb{E}}^2_{\mathbb{C}} \mid R \cong R \right\} \ .$$

Il sottoinsieme $\omega(R)$ prende il nome di luogo polare di R rispetto a γ.

Nel caso in cui $\omega(R)$ sia una retta, parleremo di *retta polare* o semplicemente di *polare* di R.

8.7 Proposizione. *Per ogni punto R di $\widehat{\mathbb{E}}^2_{\mathbb{C}}$ abbiamo che:*

(i) se $R \in \mathcal{D}$, $\omega(R) = \widehat{\mathbb{E}}^2_{\mathbb{C}}$;

(ii) se $R \notin \mathcal{D}$, $\omega(R)$ *è una retta.*

Pertanto, se γ è non degenere, non esistono punti doppi e ogni punto possiede retta polare, e quindi ω può riguardarsi come applicazione da $\widehat{\mathbb{E}}^2_{\mathbb{C}}$ all'insieme $\widehat{\mathbb{L}}^2_{\mathbb{C}}$ delle rette di $\widehat{\mathbb{E}}^2_{\mathbb{C}}$. Dalla simmetria della forma f discende facilmente il seguente risultato.

8.8 Teorema di reciprocità. *Sia γ una conica e siano R, S due punti di $\widehat{\mathbb{E}}^2_{\mathbb{C}}$. Allora*

$$R \in \omega(S) \iff S \in \omega(R) .$$

In particolare un punto doppio appartiene al luogo polare di ogni altro punto.

8.9 Teorema. *Sia γ una conica reale, descritta dalla* (8.6)*. L'insieme $\mathcal{D}$ dei suoi punti doppi è di uno dei seguenti tipi:*

(a) $\mathcal{D} = \emptyset$ *se e solo se* $\rho(B) = 3$;

(b) $\mathcal{D} = \{V\}$, *un singleton, se e solo se* $\rho(B) = 2$;

(c) $\mathcal{D} = r$, *una retta, se e solo se* $\rho(B) = 1$.

8.10 Definizione. *La conica γ si dirà non degenere se $\mathcal{D} = \emptyset$, degenere (o riducibile) se $\mathcal{D} \neq \emptyset$. In particolare, se $\mathcal{D} = \{V\}$ e $\rho(B) = 2$ γ è semplicemente degenere, se $\mathcal{D}$ è una retta e $\rho(B) = 1$ γ è doppiamente degenere.*

Sulle coniche doppiamente degeneri non c'è molto da dire. Si verifica che $\rho(B) = 1$ se e solo se esiste un polinomio omogeneo di primo grado

$$p(x_1, x_2, x_3) = a_1 x_1 + a_2 x_2 + a_3 x_3$$

(a coefficienti non tutti nulli) tale che

$$f(x_1, x_2, x_3; x_1, x_2, x_3) = (p(x_1, x_2, x_3))^2 .$$

Se indichiamo con r la retta rappresentata dal polinomio p, abbiamo che $\gamma = r = \mathcal{D}$, e diciamo che i punti di r sono contati due volte, in quanto sono doppi. La retta r risulta necessariamente reale.

8.11 Esempio. *Sia* $\gamma : x_1^2 - 2x_1x_2 + x_2^2 + 2x_1x_3 - 2x_2x_3 + x_3^2 = 0$. *La matrice associata è*

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

che ha chiaramente rango 1. *Si vede subito che*

$$x_1^2 - 2x_1x_2 + x_2^2 + 2x_1x_3 - 2x_2x_3 + x_3^2 = (x_1 - x_2 + x_3)^2 \ .$$

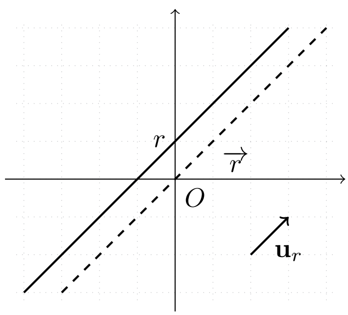
Quindi γ *non è altro che la retta*

$$r : x_1 - x_2 + x_3 = 0$$

contata due volte. La parte propria di r *è rappresentata, in coordinate affini, dall'equazione*

$$x - y + 1 = 0 \ .$$

Il punto improprio di r*, che è anche l'unico punto improprio di* γ*, è il punto* $P_r \equiv [1, 1, 0]$.



$$r : x_1 - x_2 + x_3 = 0$$

$$\gamma : (x_1 - x_2 + x_3)^2 = 0$$

$$\gamma = \mathcal{D} = r$$

Il caso delle coniche semplicemente degeneri è un po' più articolato. In analogia con il caso precedente, accade che $\rho(B) = 2$ se e solo se esistono due polinomi

$$p'(x_1, x_2, x_3) = a_1'x_1 + a_2'x_2 + a_3'x_3 \ ; \ p''(x_1, x_2, x_3) = a_1''x_1 + a_2''x_2 + a_3''x_3$$

(non identicamente nulli), tali che

$$f(x_1, x_2, x_3; x_1, x_2, x_3) = p'(x_1, x_2, x_3) \cdot p''(x_1, x_2, x_3) \ .$$

Pertanto, dette r', r'' le rette rappresentate dalle equazioni

$$p'(x_1, x_2, x_3) = 0 \ ; \ p''(x_1, x_2, x_3) = 0 \ ,$$

si ha che $\gamma = r' \cup r''$ e $\mathcal{D} = \{V\}$, dove V è il punto di intersezione delle due rette, che è sicuramente reale, e può essere proprio (nel caso di rette incidenti) o improprio (quando $r' \parallel r''$).

8.12 Esempio. *(Coppia di rette reali parallele). Sia $\gamma : x_1^2 - 2x_1x_2 + x_2^2 - x_3^2 = 0$. La matrice associata è*

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

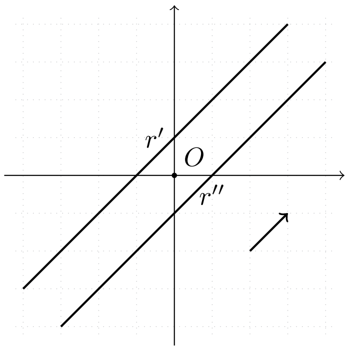
che ha chiaramente rango 2. Si vede subito che

$$x_1^2 - 2x_1x_2 + x_2^2 - x_3^2 = (x_1 - x_2 + x_3) \cdot (x_1 - x_2 - x_3) \ .$$

Pertanto, posto

$$r' : x_1 - x_2 + x_3 = 0 \ ; \ r'' : x_1 - x_2 - x_3 = 0 \ ,$$

vediamo che $\gamma = r' \cup r''$ e $\mathcal{D} = \{V\}$, dove $V \equiv [1, 1, 0]$. Le rette r', r'' sono parallele e $V = P_{r'} = P_{r''}$ è il punto improprio che rappresenta la direzione comune delle due rette.



$$r' : x_1 - x_2 + x_3 = 0 \ ; \ r'' : x_1 - x_2 - x_3 = 0$$

$$\gamma : (x_1 - x_2 + x_3)(x_1 - x_2 - x_3) = 0$$

$$\gamma = r' \cup r'' \ ; \ \mathcal{D} = \{V\} \ ; \ V \equiv [1, 1, 0]$$

8.13 Esempio. *(Coppia di rette reali incidenti). Sia $\gamma : x_1^2 - x_2^2 - 2x_2x_3 - x_3^2 = 0$. La matrice associata è*

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

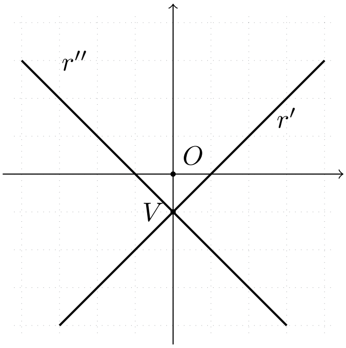
che ha chiaramente rango 2. Si vede subito che

$$x_1^2 - x_2^2 - 2x_2x_3 - x_3^2 = (x_1 - x_2 - x_3) \cdot (x_1 + x_2 + x_3) \ .$$

Pertanto, posto

$$r' : x_1 - x_2 - x_3 = 0 \ ; \ r'' : x_1 + x_2 + x_3 = 0 \ ,$$

vediamo che $\gamma = r' \cup r''$ e $\mathcal{D} = \{V\}$, dove $V \equiv [0, -1, 1]$. Le rette r', r'' sono incidenti e V è il punto reale proprio che le due rette hanno in comune.



$$r' : x_1 - x_2 - x_3 = 0 \ ; \ r'' : x_1 + x_2 + x_3 = 0$$

$$\gamma : (x_1 - x_2 - x_3)(x_1 + x_2 + x_3) = 0$$

$$\gamma = r' \cup r'' \ ; \ \mathcal{D} = \{V\} \ ; \ V \equiv [0, -1, 1]$$

8.14 Esempio. *(Coppia di rette immaginarie parallele, con direzione reale). Consideriamo la conica* $\gamma : x_1^2 + 2x_1x_2 + x_2^2 + x_3^2 = 0$. *La matrice associata è*

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

che ha chiaramente rango 2. *Si vede subito che*

$$x_1^2 + 2x_1x_2 + x_2^2 + x_3^2 = (x_1 + x_2 + ix_3) \cdot (x_1 + x_2 - ix_3) \ .$$

Pertanto, posto

$$r' : x_1 + x_2 + ix_3 = 0 \ ; \ r'' : x_1 + x_2 - ix_3 = 0 \ ,$$

vediamo che $\gamma = r' \cup r''$ *e* $\mathcal{D} = \{V\}$, *dove* $V \equiv [-1, 1, 0]$. *Le rette* r', r'' *sono immaginarie parallele e* V *è il punto reale improprio, ovvero la direzione, che le due rette hanno in comune.*

8.15 Esempio. *(Coppia di rette immaginarie incidenti in un punto reale). Consideriamo la conica* $\gamma : x_1^2 + x_2^2 + 2x_1x_3 + x_3^2 = 0$. *La matrice associata è*

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

che ha chiaramente rango 2. *Si vede subito che*

$$x_1^2 + x_2^2 + 2x_1x_3 + x_3^2 = (x_1 + x_3)^2 + x_2^2 = (x_1 + ix_2 + x_3) \cdot (x_1 - ix_2 + x_3) \ .$$

Pertanto, posto

$$r' : x_1 + ix_2 + x_3 = 0 \ ; \ r'' : x_1 - ix_2 + x_3 = 0 \ ,$$

vediamo che $\gamma = r' \cup r''$ *e* $\mathcal{D} = \{V\}$, *dove* $V \equiv [-1, 0, 1]$. *Le rette* r', r'' *sono immaginarie incidenti e* V *è il punto reale proprio che le due rette hanno in comune.*

8.16 Esempio. *(Coppia di rette reali di cui una è la retta impropria). Consideriamo la conica* $\gamma : x_1x_3 + x_2x_3 + x_3^2 = 0$. *La matrice associata è proporzionale a*

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

che ha chiaramente rango 2. Si vede subito che

$$x_1x_3 + x_2x_3 + x_3^2 = x_3 \cdot (x_1 + x_2 + x_3) \ .$$

Pertanto, posto

$$r' = r_\infty : x_3 = 0 \ ; \ r'' : x_1 + x_2 + x_3 = 0 \ ,$$

vediamo che $\gamma = r_\infty \cup r''$ *e* $\mathcal{D} = \{V\}$*, dove* $V \equiv [-1, 1, 0]$.

Supponiamo ora che γ sia non degenere, e cioè si abbia $\det B \neq 0$, ovvero $\rho(B) = 3$. Deduciamo che γ non contiene rette e che ogni retta interseca γ in uno o due punti. Consideriamo allora l'insieme $r_\infty \cap \gamma$ dei punti impropri della conica.

8.17 Definizione. *La conica non degenere* γ *si dice ellisse se possiede due punti impropri immaginari, iperbole se possiede due punti impropri reali, parabola se possiede un solo punto improprio (reale).*

Il seguente risultato fornisce uno strumento per classificare una conica non degenere.

8.18 Teorema. *Sia* γ *una conica reale rappresentata dalla* (8.6) *e supponiamo che* γ *non contenga la retta impropria* r_∞*. Allora si ha che*

$$r_\infty \cap \gamma = \begin{cases} 2 \textit{ punti immaginari} & \iff \det B_{(3,3)} > 0 \\ 1 \textit{ punto reale} & \iff \det B_{(3,3)} = 0 \\ 2 \textit{ punti reali} & \iff \det B_{(3,3)} < 0 \end{cases}$$

8.19 Corollario. *Sia* γ *una conica reale non degenere rappresentata dalla* (8.6)*. Allora si ha che*

$$\gamma \ \ \text{è} \begin{cases} \textit{un'ellisse} & \iff \det B_{(3,3)} > 0 \\ \textit{una parabola} & \iff \det B_{(3,3)} = 0 \\ \textit{un'iperbole} & \iff \det B_{(3,3)} < 0 \end{cases}$$

Nel caso in cui γ sia una conica semplicemente degenere, il teorema precedente ci assicura che la conica è unione di due rette immaginarie incidenti se $\det B_{(3,3)} > 0$, di due rette parallele (reali o immaginarie, ma di direzione reale) se $\det B_{(3,3)} = 0$, di due rette reali incidenti se $\det B_{(3,3)} < 0$.

Come già osservato, il fatto che γ sia una conica reale, ovvero una conica rappresentata da un'equazione a coefficienti reali, non garantisce l'esistenza in γ di punti reali. E' pertanto utile il seguente risultato, di non facile dimostrazione.

8.20 Teorema. *La conica reale* γ *rappresentata dalla* (8.6) *possiede almeno un punto reale se e solo se*

$$\det B_{(3,3)} \leq 0 \quad \vee \quad b_{1,1} \cdot \det B \leq 0 \ .$$

Pertanto una conica che non possiede punti reali è caratterizzata come segue:

$$\det B_{(3,3)} > 0 \quad \wedge \quad b_{1,1} \cdot \det B > 0 \ . \tag{8.8}$$

La seconda delle (8.8) implica che $\det B \neq 0$, e quindi γ è non degenere, e la prima delle (8.8) ci dice che si tratta di un'ellisse. Un'ellisse priva di punti reali è detta immaginaria, ed è caratterizzata dalla seconda delle (8.8). Un'ellisse dotata di punti reali si dice invece ordinaria.

Ci occuperemo ora di alcune applicazioni del concetto di retta polare nel caso in cui la conica γ sia non degenere. Abbiamo già osservato che in tale situazione ogni punto possiede una retta polare. Più precisamente, vale il seguente risultato.

8.21 Proposizione. *L'applicazione* $\omega : \widehat{\mathbb{E}}^2_{\mathbb{C}} \to \widehat{\mathbb{L}}^2_{\mathbb{C}}$ *che ad ogni punto associa la sua (retta) polare è biettiva.*

Di conseguenza, per ogni retta r esisterà esattamente un punto R tale che $\omega(R) = r$ e diremo che R è il *polo* di r. Inoltre

8.22 Corollario. *Per ogni punto* R *abbiamo che*

$$R \in \omega(R) \iff R \in \gamma \ .$$

In tal caso la polare di R coincide con la tangente di γ in R. Dalla definizione di retta polare deduciamo che se $R \in \widehat{\mathbb{E}}^2_{\mathbb{C}}$ e $R \equiv Y^t$, o, in altri termini, R ha coordinate omogenee y_1, y_2, y_3, la retta polare di R si può rappresentare, in modo compatto, come segue:

$$\omega(R) : Y^t B X = 0 \ ,$$

ovvero, in modo più esplicito,

$$\omega(R) : a x_1 + b x_2 + c x_3 = 0 \ ,$$

dove

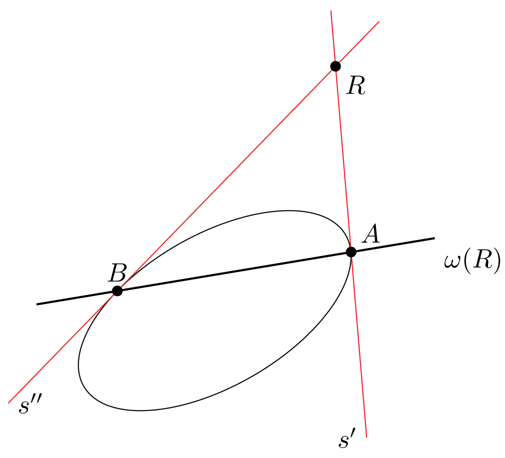
$$a = Y^t \cdot B^1 = y_1 b_{1,1} + y_2 b_{2,1} + y_3 b_{3,1} \ ;$$
$$b = Y^t \cdot B^2 = y_1 b_{1,2} + y_2 b_{2,2} + y_3 b_{3,2} \ ;$$

$$c = Y^t \cdot B^3 = y_1 b_{1,3} + y_2 b_{2,3} + y_3 b_{3,3} \ .$$

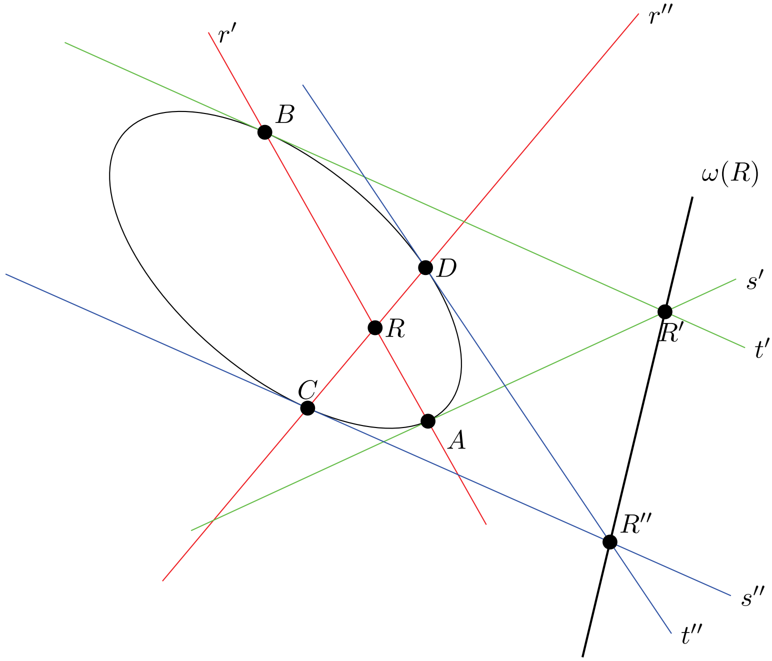
Ricordiamo che $R \in \gamma$ se e solo se

$$f(y_1, y_2, y_3; y_1, y_2, y_3) = f(Y, Y) = Y^t B Y = 0 \ .$$

Se invece accade che $f(Y, Y) > 0$, diremo che R è esterno a γ, mentre parleremo di punto interno quando $f(Y, Y) < 0$. Graficamente nel caso in cui R sia esterno, la polare si costruisce nel modo indicato nella seguente figura.



Si verifica infatti che esistono esattamente due rette reali s', s'' per R tangenti a γ, e, detti A, B i punti di tangenza, il Teorema di Reciprocità ci assicura che la retta r per A, B coincide con la polare di R. Se invece R è interno, si verifica che ogni retta reale per R è secante. Scelte quindi due rette r', r'' per R ed indicati con A, B i punti in cui r' interseca γ ed C, D i punti in cui r'' interseca γ, consideriamo le tangenti a γ nei punti A, B, C, D. Indichiamo tali rette con i simboli s', t', s'', t'' e consideriamo il punto R' in cui s' interseca t' e il punto R'' in cui s'' interseca t''. Il Teorema di Reciprocità ci assicura, anche in questo caso, che la retta r per R', R'' coincide con la polare di R.



Nella teoria delle coniche reali non degeneri riveste una notevole importanza lo studio delle polari dei punti impropri.

8.23 Definizione. *Sia γ una conica reale non degenere e sia $R \in r_\infty$.*

(i) *Se la polare $\omega(R)$ del punto improprio R è una retta propria, essa si dice* diametro *di γ, associato alla direzione rappresentata da R.*

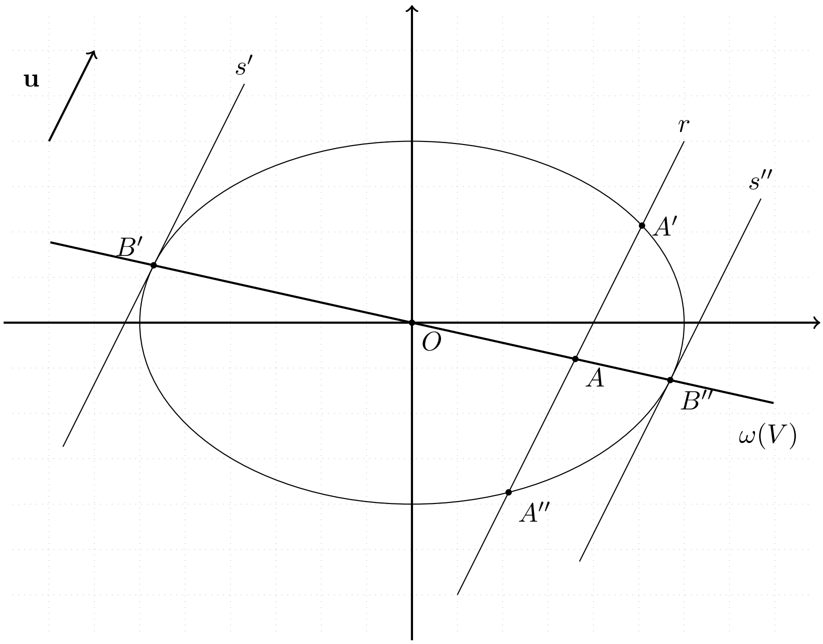
(ii) *Se $\omega(R)$ è un diametro e $R \in \gamma \cap r_\infty$, $\omega(R)$ si dice* asintoto *di γ.*

(iii) *Se $\omega(R)$ è un diametro e risulta ortogonale alla direzione rappresentata da R, $\omega(R)$ si dice* asse *di γ.*

8.24 Teorema. *I diametri di una conica reale non degenere formano un fascio, che sarà proprio nel caso dell'ellisse e dell'iperbole, improprio nel caso della parabola.*

Nel caso dell'ellisse e dell'iperbole, il punto C comune a tutti i diametri si dice *centro* della conica. Per tale motivo l'ellisse e l'iperbole sono dette *coniche a centro*.

8.25 Osservazione. *Per ogni punto improprio R non appartenente a γ, la polare $\omega(R)$, se è una retta propria (cioè è un diametro) ed è secante, è un asse di simmetria di γ rispetto* *alla direzione rappresentata da R, nel senso che se $R \equiv [\alpha, \beta, 0]$ ed r è una retta di direzione $\mathbf{u}_r = (\alpha, \beta)$, detti A', A'' i punti di intersezione di r con γ ed A il punto di intersezione di r con $\omega(R)$, si ha che A è il punto medio di A', A''.*



$$\mathbf{u} = (1, 2),\ V \equiv [1, 2, 0]$$

8.26 Osservazione. *Le coniche a centro hanno due assi, che sono ortogonali tra loro. Se $R \in r_\infty$ ed $\omega(R)$ è un asse, detto S il punto improprio di $\omega(R)$, $\omega(S)$ è l'altro asse.*

8.27 Osservazione. *La parabola ha un unico asse, il cui punto improprio è l'unico punto improprio della conica.*

8.28 Definizione. *Un punto proprio si dice vertice se appartiene all'intersezione di un asse con la conica.*

8.29 Osservazione. *(i) Un'ellisse ordinaria che non sia una circonferenza ha quattro vertici reali.*

(ii) Un'iperbole ha due vertici reali e due vertici immaginari (che sono tra loro complessi coniugati).

(iii) Una parabola possiede un unico vertice reale.

Abbiamo finora ritrovato tutti gli elementi caratterizzanti di una conica, così come appaiono dalla iniziale descrizione elementare, tranne i fuochi. Essi necessitano di un approccio un po' più sofisticato.

8.30 Definizione. *Si dicono* punti ciclici *del piano i punti (impropri immaginari)*

$$C_1 \equiv [1, i, 0] \ , \ C_2 \equiv [1, -i, 0] \ .$$

Ogni retta propria passante per un punto ciclico si dice isotropa*.*

Per ogni punto proprio R passano due rette isotrope: la retta r' per R e C_1 e la retta r'' per R e C_2.

8.31 Definizione. *Un* fuoco *di una conica reale non degenere γ è un punto proprio $R \notin \gamma$ tale che le rette isotrope per R siano tangenti a γ.*

Osserviamo che una conica a centro dotata di punti reale (iperbole o ellisse ordinaria) ha due fuochi reali e due immaginari coniugati, mentre una parabola ha un solo fuoco, che risulta reale.

8.3 Riduzione in forma canonica

Consideriamo una conica γ rappresentata come in (8.6), e studiamo, in coordinate affini, la sua parte propria, rappresentata come in (8.7). In tale equazione abbiamo una parte quadratica

$$b_{1,1}x^2 + 2b_{1,2}xy + b_{2,2}y^2 = (x, y) \cdot B_{(3,3)} \cdot \begin{pmatrix} x \\ y \end{pmatrix}$$

ed una parte lineare

$$2b_{1,3}x + 2b_{2,3}y = 2\,(b_{1,3}, b_{2,3}) \cdot \begin{pmatrix} x \\ y \end{pmatrix} \ ,$$

per cui la (8.7) si scrive anche

$$\gamma : (x, y) \cdot B_{(3,3)} \cdot \begin{pmatrix} x \\ y \end{pmatrix} + 2\,(b_{1,3}, b_{2,3}) \cdot \begin{pmatrix} x \\ y \end{pmatrix} + b_{3,3} = 0 \tag{8.9}$$

oppure

$$\gamma : (x, y, 1) \cdot B \cdot \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = 0 \ . \tag{8.10}$$

Poiché $B_{(3,3)}$, al pari di B, è simmetrica, esiste una matrice ortogonale

$$P = \begin{pmatrix} p_{1,1} & p_{1,2} \\ p_{2,1} & p_{2,2} \end{pmatrix} \in O(2)$$

tale che

$$P^t \cdot B \cdot P = \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix} \ .$$

Consideriamo allora il riferimento ortonormale $\mathcal{R}' = (O, \mathbf{e}'_1, \mathbf{e}'_2)$ (con lo stesso origine O di $\mathcal{R}$) tale che P sia la matrice di passaggio, ovvero $M^{\mathcal{R}}_{\mathcal{R}'} = P$. Ciò vuol dire che se, detto R il generico punto, si ha che

$$R \equiv_{\mathcal{R}} (x, y) \ , \ R \equiv_{\mathcal{R}'} (x', y') \ ,$$

si avrà anche che

$$\begin{pmatrix} x \\ y \end{pmatrix} = P \cdot \begin{pmatrix} x' \\ y' \end{pmatrix} \ , \ (x, y) = (x', y') \cdot P^t$$

e quindi γ sarà rappresentata, in $\mathcal{R}'$ dall'equazione

$$\gamma : (x', y') \cdot P^t \cdot B_{(3,3)} \cdot P \cdot \begin{pmatrix} x' \\ y' \end{pmatrix} + 2\,(b_{1,3}, b_{2,3}) \cdot P \cdot \begin{pmatrix} x' \\ y' \end{pmatrix} + b_{3,3} = 0 \tag{8.11}$$

ovvero

$$\lambda x'^2 + \mu y'^2 + 2b'_{1,3}x' + 2b'_{2,3}y' + b_{3,3} = 0 \tag{8.12}$$

dove

$$b'_{1,3} = b_{1,3}p_{1,1} + b_{2,3}p_{2,1} \ , \ b'_{2,3} = b_{1,3}p_{1,2} + b_{2,3}p_{2,2} \ .$$

Il cambiamento di riferimento appena descritto è una rotazione, eventualmente seguita da un cambio di orientamento degli assi cartesiani. A seguito di tale cambiamento di riferimento, i nuovi assi cartesiani risultano paralleli agli assi della conica, nel caso delle coniche non degeneri a centro (ellisse e iperbole). Nel caso della parabola avremo invece un asse cartesiano parallelo all'asse della parabola.

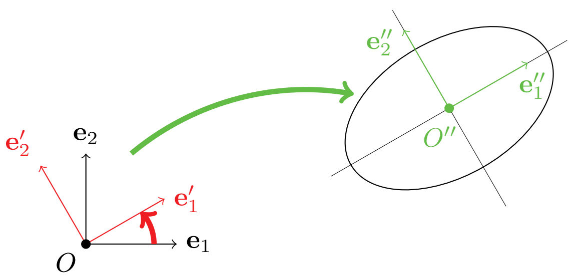
Ora dobbiamo distinguere due casi. Per quanto riguarda il primo caso, supponiamo che gli autovalori λ, μ siano entrambi non nulli. La (8.12) può scriversi allora

$$\lambda\big(x'^2 + 2\frac{b'_{1,3}}{\lambda}\big) + \mu\big(y'^2 + 2\frac{b'_{2,3}}{\mu}\big) + b_{3,3} = 0$$

o anche

$$\lambda\big(x'^2 + 2\frac{b'_{1,3}}{\lambda}x' + \frac{b'^2_{1,3}}{\lambda^2}\big) + \mu\big(y'^2 + 2\frac{b'_{2,3}}{\mu}y' + \frac{b'^2_{2,3}}{\mu^2}\big) + b_{3,3} - \frac{b'^2_{1,3}}{\lambda} - \frac{b'^2_{2,3}}{\mu} = 0$$

ovvero
$$\lambda\big(x' + \frac{b'_{1,3}}{\lambda}\big)^2 + \mu\big(y' + \frac{b'_{2,3}}{\mu}\big)^2 + b_{3,3} - \frac{b'^2_{1,3}}{\lambda} - \frac{b'^2_{2,3}}{\mu} = 0 \ .$$
Consideriamo allora un ulteriore riferimento $\mathcal{R}'' = (O'', \mathbf{e}'_1, \mathbf{e}'_2)$ (questa volta si tratta di una traslazione, ovvero cambia l'origine ma non la base di vettori) tale che, posto $R \equiv_{\mathcal{R}''} (x'', y'')$, si abbia
$$\begin{cases} x'' = x' + \dfrac{b'_{1,3}}{\lambda} \\ y'' = y' + \dfrac{b'_{2,3}}{\mu} \end{cases}$$
In altri termini, la matrice completa del cambiamento da $\mathcal{R}'$ a $\mathcal{R}''$ (vedi (7.39)) è
$$\tilde{A} = \left(\begin{array}{c|c} I_2 & \begin{smallmatrix} b'_{1,3} \\ b'_{1,3} \end{smallmatrix} \\ \hline 0\ 0 & 1 \end{array}\right) \ . \tag{8.13}$$
In $\mathcal{R}''$ la conica è quindi rappresentata da
$$\gamma : \lambda x''^2 + \mu y''^2 + b''_{3,3} = 0 \tag{8.14}$$
dove
$$b''_{3,3} = b_{3,3} - \frac{b'^2_{1,3}}{\lambda} - \frac{b'^2_{2,3}}{\mu} \ .$$
Se $b''_{3,3} \neq 0$, con semplici passaggi si perviene ad una rappresentazione del tipo (8.1) quando λ, μ sono concordi e $b''_{3,3}$ è discorde da λ, μ, e si tratterà di una ellisse ordinaria, si perviene ad una rappresentazione del tipo (8.2) quando λ, μ sono discordi, e si tratterà allora di una iperbole. Se invece $\lambda, \mu, b''_{3,3}$ sono concordi, avremo una ellisse immaginaria.



Se $b''_{3,3} = 0$ la (8.14) si riduce a
$$(\sqrt{|\lambda|}x'' + \sqrt{|\mu|}y'')(\sqrt{|\lambda|}x'' - \sqrt{|\mu|}y'') = 0$$

e γ è l'unione di due rette reali, quando λ, μ sono concordi. Se invece λ, μ sono discordi, ad esempio $\lambda > 0$ e $\mu < 0$, otterremo l'equazione

$$(\sqrt{\lambda}x'' + i\sqrt{-\mu}y'')(\sqrt{\lambda}x'' - i\sqrt{-\mu}y'') = 0$$

e γ sarà l'unione di due rette immaginarie incidenti in un punto reale.

Consideriamo ora il secondo caso, ovvero supponiamo che uno degli autovalori, ad esempio μ, si annulli. Allora la (8.12) si riduce a

$$\lambda x'^2 + 2b'_{1,3}x' + 2b'_{2,3}y' + b_{3,3} = 0$$

ovvero

$$\lambda\big(x'^2 + 2\frac{b'_{1,3}}{\lambda}x' + \frac{b'^2_{1,3}}{\lambda^2}\big) + 2b'_{2,3}y' + +b_{3,3} - \frac{b'^2_{1,3}}{\lambda} = 0$$

ovvero

$$\lambda\big(x' + \frac{b'_{1,3}}{\lambda}\big)^2 + \mu\big(y' + \frac{b'_{2,3}}{\mu}\big)^2 + b''_{3,3} = 0 \ ,$$

dove

$$b''_{3,3} = b_{3,3} - \frac{b'^2_{1,3}}{\lambda} \ .$$

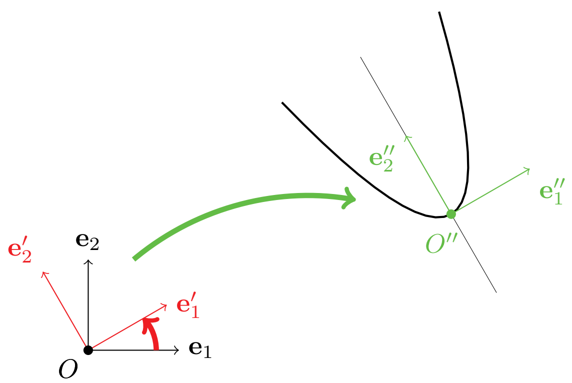
Se $b'_{2,3} \neq 0$, utilizziamo la traslazione

$$\begin{cases} x'' = x' + \dfrac{b'_{1,3}}{\lambda} \\ y'' = y' + \dfrac{\lambda b_{3,3} - b'^2_{1,3}}{2b'_{2,3}\lambda} \end{cases}$$

ed otteniamo la rappresentezione

$$\gamma : \lambda x''^2 + 2b'_{2,3}y'' = 0 \ ,$$

che equivale ad un'equazione del tipo (8.3), e γ è una parabola.



Se invece $b'_{2,3} = 0$, la conica è degenere. Utilizziamo la traslazione

$$\begin{cases} x'' = x' + \dfrac{b'_{1,3}}{\lambda} \\ y'' = y' \end{cases}$$

ed otteniamo la rappresentezione

$$\gamma : \lambda x''^2 + b''_{3,3} = 0 \ .$$

In questo caso, se $b''_{3,3} \neq 0$, γ è l'unione di due rette parallele, che saranno reali se $\lambda, b''_{3,3}$ sono discordi, immaginarie (ma con direzione reale) se $\lambda, b''_{3,3}$ sono concordi. Se infine $b''_{3,3} = 0$, γ è doppiamente degenere.

In definitiva ogni conica può rappresentarsi, in un opportuno riferimento, in forma canonica.

Capitolo 9

Esercizi

Esercizi sul primo capitolo

1. *Relazioni in un insieme.* Si considerino le seguenti relazioni h, h' in $\mathbb{Z}$:

$$a\,h\,b \iff b^2 - a^2 \text{ è pari } ; \; a\,h'\,b \iff b^2 - a^2 \text{ è multiplo di } 4 \; .$$

(i) Si verifichi che h, h' sono relazioni d'equivalenza.

(ii) Si verifichi che h, h' coincidono, entrambe, con la relazione dell'Esempio 1.12.

2. *Relazioni in un insieme.* Si considerino le seguenti relazioni h, h' in $\mathbb{N}$:

$$a\,h\,b \iff a \text{ è un divisore di } b$$

$$a\,h'\,b \iff a, b \text{ hanno un divisore comune (non invertibile) } .$$

(i) Si verifichi che h è una relazione d'ordine.

(ii) Si verifichi che h' non è una relazione d'ordine, né d'equivalenza, ma verifica la proprietà simmetrica.

3. *Applicazioni.* Si consideri l'applicazione $f : \mathbb{R}^+ \longrightarrow \mathbb{R}^+$ definita ponendo

$$f(\alpha) = \frac{\alpha}{1+\alpha} \; .$$

(i) Si verifichi che f è iniettiva.

(ii) Si verifichi che $\operatorname{im} f =]0, 1[$.

4. *Applicazioni.* Si consideri l'applicazione $g : \mathbb{R}^+ \longrightarrow]0, 1[$ definita ponendo

$$g(\alpha) = \frac{\alpha}{1+\alpha} \; .$$

(i) Si verifichi che g è biettiva.

(ii) Si determini l'inversa di g.

5. *Operazioni insiemistiche.* Sia X un insieme non vuoto e indichiamo con $P(X)$ l'insieme delle parti di X. Consideriamo le classiche operazioni insiemistiche $\cap, \cup, -$, ovvero l'intersezione, l'unione e la differenza tra sottoinsiemi.

(i) Si provi che $\big(P(X), \cap)\big)$ e $\big(P(X), \cup)\big)$ sono monoidi commutativi, specificando qual è l'elemento neutro.

(ii) Si provi che in tali strutture gli unici elementi simmetrizzabili sono quelli neutri.

(iii) Si provi che l'operazione di differenza non è associativa.

(iv) Considerata una parte propria Y di X ed indicato con $P_Y(X)$ l'insieme delle parti di X contenenti Y, si provi che $\big(P_Y(X), \cap\big)$ è una parte stabile di $\big(P(X), \cap\big)$, ed è a sua volta un monoide, ovvero, come si suol dire, è un sottomonoide, mentre $\big(P_Y(X), \cup\big)$ è una parte stabile di $\big(P(X), \cup)\big)$, ed è a sua volta un semigruppo, ma non un monoide.

6. *Moltiplicazione su una circonferenza.* Sia S^1 la circonferenza unitaria del piano euclideo, ovvero l'insieme dei punti P le cui coordinate x, y in un fissato riferimento monometrico ortogonale soddisfino la relazione $x^2 + y^2 = 1$. Per ogni $\theta \in \mathbb{R}$, indichiamo con P_θ il punto di coordinate $(\cos\theta, \sin\theta)$. È facile verificare che tutti e soli i punti di S^1 sono di questo tipo. Definiamo una operazione $\bullet$ in S^1 ponendo

$$P_\theta \bullet P_{\theta'} = P_{\theta+\theta'} \ .$$

(i) Si provi che $\bullet$ è una operazione interna in S^1 e $(S^1, \bullet)$ è un gruppo abeliano con elemento neutro P_0, ovvero il punto di coordinate $(1, 0)$.

7. *Strutture su reali positivi.* Si considerino i seguenti sottoinsiemi di $\mathbb{R}$:

$$A =]0,1[\ ; \ B =]0,1] \ ; \ C =]0,+\infty[= \mathbb{R}^+ \ .$$

(i) Si verifichi se le operazioni di addizione e moltiplicazione sono stabili in questi sottoinsiemi.

(ii) Si verifichi se $(A; \cdot)$, $(B; \cdot)$, $(C; \cdot)$ sono gruppi.

8. *Coppie di interi.* Si consideri l'insieme $\mathbb{Z} \times \mathbb{Z}$ delle coppie ordinate di numeri interi con le operazioni $+, \cdot$ definite componente per componente.

(i) Si verifichi che $(\mathbb{Z} \times \mathbb{Z}; +, \cdot)$ è un anello commutativo unitario.

(ii) Si verifichi che $(\mathbb{Z} \times \mathbb{Z}; +, \cdot)$ non è un dominio d'integrità, ovvero si verifichi che esistono due elementi non nulli con prodotto nullo.

9. *Applicazioni a valori reali.* Sia X un insieme non vuoto e si consideri l'insieme $C(X)$ delle applicazioni di X in $\mathbb{R}$. Si definiscano in $C(X)$ due operazioni $\oplus, \odot$ ponendo

$$(f \oplus g)(x) = f(x) + g(x) \ ; \ (f \odot g)(x) = f(x) \cdot g(x)$$

per ogni $f, g \in C(X)$ e per ogni $x \in X$.

(i) Si verifichi che $(C(X); \oplus, \odot)$ è un anello commutativo unitario.

(ii) Si verifichi che $(C(X); \oplus, \odot)$ non è un dominio d'integrità (a patto che l'insieme X abbia almeno due elementi).

10. *Polinomi - MCD.* Si determini il massimo comun divisore monico tra le seguenti coppie di polinomi.

(i) $f = x^3 - x^2 + x - 1$; $g = x^4 - x^3 - x^2 - x - 2$.

(ii) $f = x^4 - 1$; $g = x^3 - 6x^2 + 11x - 6$.

(iii) $f = x^3 - x^2 + x - 1$; $g = x^3 - 6x^2 + 11x - 6$.

11. *Polinomi.* Sia $f = a_0 + a_1x + \cdots + a_nx^n \in \mathbb{Q}[x]$.

(i) Si provi che f è proporzionale ad un polinomio a coefficienti interi.

(ii) Supponiamo che f sia un polinomio a coefficienti interi. Si provi che se $\frac{u}{v}$ è una radice di f, con u, v interi coprimi, ovvero privi di fattori comuni non invertibili, allora u divide a_0 e v divide a_n.

(iii) Supponiamo che f sia un polinomio a coefficienti interi. Si provi che ogni sua radice intera divide a_0.

(iv) Supponiamo che f sia un polinomio monico a coefficienti interi. Si provi che ogni sua radice è intera.

12. *Polinomi.* Si determinino due polinomi in $\mathbb{Z}[x]$ che ammettono le stesse radici, ma non sono proporzionali.

Esercizi sul secondo capitolo

1. *Dipendenza ed indipendenza lineare.* In uno spazio vettoriale V sul campo reale, sia $\mathcal{S} = [\mathbf{u}, \mathbf{v}, \mathbf{w}]$ un sistema di vettori di V.

(i) Si verifichi che il sistema $\mathcal{S}' = [\mathbf{u}+\mathbf{v}+\mathbf{w}, -\mathbf{u}+\mathbf{v}, 2\mathbf{u}+\mathbf{w}]$ è dipendente.

(ii) Si verifichi che, se $\mathcal{S}$ è indipendente, allora anche il sistema $\mathcal{S}'' = [\mathbf{u} + \mathbf{v} + \mathbf{w}, \mathbf{u} + \mathbf{v}, \mathbf{u}]$ è indipendente.

2. *Dipendenza ed indipendenza lineare nello spazio dei polinomi.* Sia $V = \mathbb{K}[x]$ lo spazio vettoriale dei polinomi sul campo $\mathbb{K}$ e si consideri il sistema $\mathcal{S} = [1, x, x^2, x^3]$.

(i) Si provi che ogni polinomio di grado al più 3 dipende da $\mathcal{S}$.

(ii) Si provi che ogni polinomio di grado maggiore di 3 non dipende da $\mathcal{S}$.

3. *Generatori in* $\mathbb{R}^3$. Sia $V = \mathbb{R}^3$, e si considerino i vettori

$$\mathbf{v}_1 = (1,1,1) \; ; \; \mathbf{v}_2 = (1,1,2) \; ; \; \mathbf{v}_3 = (1,1,3) \; ; \; \mathbf{e}_1 = (1,0,0) \; .$$

(i) Si verifichi il sistema $\mathcal{S} = [\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{e}_1]$ e il sistema $[\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3]$ sono dipendenti.

(ii) Si verifichi il sistema $[\mathbf{v}_1, \mathbf{v}_2]$ e il sistema $[\mathbf{v}_1, \mathbf{v}_2, \mathbf{e}_1]$ sono indipendenti.
(iii) Si determinino tutti i sottosistemi indipendenti massimali di $\mathcal{S}$.
(iv) Si verifichi $[\mathbf{v}_1, \mathbf{v}_2, \mathbf{e}_1]$ e $\mathcal{S}$ sono sistemi di generatori (e quindi $[\mathbf{v}_1, \mathbf{v}_2, \mathbf{e}_1]$ è una base).

4. *Completamento a base.* Si considerino i vettori $\mathbf{u} = (1, 1, 1)$ e $\mathbf{v} = (1, 1, 0)$ di $\mathbb{R}^3$.

(i) Si verifichi che $[\mathbf{u}, \mathbf{v}]$ è indipendente.
(ii) Si completi $[\mathbf{u}, \mathbf{v}]$ a base di $\mathbb{R}^3$, utilizzando uno dei vettori $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ della base standard. Si evidenzi anche quale di tali vettori non può essere utilizzato.

5. *Sottospazi in* $\mathbb{R}^3$. Si considerino i seguenti sottoinsiemi di $\mathbb{R}^3$:

$$A = \left\{ (a, b, c) \in \mathbb{R}^3 \mid a + b + c = 0 \right\} ;$$

$$B = \left\{ (a, b, c) \in \mathbb{R}^3 \mid a + b + c = 1 \right\} .$$

(i) Si provi che $A \leq \mathbb{R}^3$ e che $B \not\leq \mathbb{R}^3$.
(ii) Si determinino due elementi non proporzionali di A e si verifichi che ne costituiscono una base.
(iii) Detto $\mathcal{S}$ il sistema determinato al punto precedente, si completi $\mathcal{S}$ a base, utilizzando un vettore della base standard.

6. *Somme di sottospazi.* Si consideri lo spazio vettoriale $V = \mathbb{R}[x]$.

(i) Si determinino due sottospazi $U, W \leq \mathbb{R}[x]$ tali che risulti $\mathbb{R}[x] = U \oplus W$.
(ii) Si determinino due sottospazi $U, W \leq \mathbb{R}[x]$ tali che risulti $\mathbb{R}[x] = U + W$, somma NON diretta.

7. *Bandiere di sottospazi.* Sia V uno spazio vettoriale di dimensione n e si consideri una n-pla $(U_1, \ldots, U_n)$ di sottospazi di V tale che si abbia

$$U_1 < U_2 < \cdots < U_n = V \ .$$

Una tale sequenza si dice *bandiera completa di sottospazi.*

(i) Si dimostri che $\dim U_j = j$, per ogni j.
(ii) Si provi che esiste una base ordinata $\mathcal{B} = (\mathbf{u}_1, \ldots, \mathbf{u}_n)$ di V tale che $U_j = \mathcal{L}(\mathbf{u}_1, \ldots, \mathbf{u}_j)$, per ogni j.

8. *Immagini di sottospazi.* Sia $f : V \to V'$ un'applicazione lineare.

(i) Sia $U \leq V$. Si provi che $f(U) \leq V'$ e che, se U è finitamente generato, $\dim(f(U)) \leq \dim U$.
(ii) Sia $U' \leq V'$. Si provi che $f^{-1}(U') \leq V$ e che, se $f^{-1}(U')$ è finitamente generato, $\dim U' \leq \dim f^{-1}(U')$.

9. *Omotetie.* Sia $f : V \to V$ un'omotetia di ragione $k \neq 0$.

(i) Si verifichi che f è un isomorfismo, la cui inversa è ancora un'omotetia, di ragione k^{-1}.

(ii) Si provi che per ogni sottospazio U di V si ha che $f(U) = U$.

10. *Applicazioni lineari definite su una base ed estese per linearità.* Si determini l'applicazione lineare $f; \mathbb{R}^2 \to \mathbb{R}^3$ definita ponendo

$$f(1,0) = (1,1,1) \ ; \ f(0,1) = (1,2,3)$$

ed estendendo poi per linearità.

$\left[f(a,b)=(a+b,a+2b,a+3b) \right]$

11. *Applicazioni lineari definite su una base ed estese per linearità.* Si determini l'applicazione lineare $f; \mathbb{R}^2 \to \mathbb{R}^3$ definita ponendo

$$f(1,1) = (1,0,-1) \ ; \ f(1,-1) = (-3,0,3)$$

ed estendendo poi per linearità. $\left[f(a,b)=(-a+2b,0,a-2b) \right]$

12. *Nucleo e immagine.* Sia $f : V \to V$ un endomorfismo, e sia $\dim V = n$.

(i) Si stabilisca se è possibile che $\ker f = \operatorname{im} f$ quando n è pari e quando n è dispari.

(ii) Nel caso in cui ciò è possibile, si costruisca un esempio.

13. *Nucleo, immagine e somme di sottospazi.* Sia V uno spazio vettoriale di dimensione 3.

(i) Si costruisca un endomorfismo $f : V \to V$ tale che $V = \ker f \oplus \operatorname{im} f$.

(ii) Si costruisca un endomorfismo $g : V \to V$ tale che $\ker g \leq \operatorname{im} g$.

(iii) Si costruisca un endomorfismo $h : V \to V$ tale che $\operatorname{im} h \leq \ker h$.

14. *Applicazioni lineari composte.* Siano V, V', V'' spazi vettoriali su un campo $\mathbb{K}$ e siano

$$f : V \longrightarrow V' \quad ; \quad g : V' \longrightarrow V''$$

due applicazioni lineari.

(i) Si verifichi che $\operatorname{im}(g \circ f) \leq \operatorname{im} g \leq V''$.

(ii) Si verifichi che $\ker f \leq \ker(g \circ f) \leq V$.

(iii) Supposto che gli spazi siano tutti finitamente generati, si verifichi che $\dim(\operatorname{im}(g \circ f)) \leq \dim(\operatorname{im} f)$.

15. *Endomorfismi ed Automorfismi.* Sia V uno spazio vettoriale su un campo $\mathbb{K}$. Un endomorfismo $f : V \to V$ che sia anche un isomorfismo si dice anche *automorfismo*. Indichiamo con $\operatorname{End}(V)$ l'insieme degli endomorfismi di V, e con $\operatorname{Aut}(V)$ l'insieme degli automorfismi di V.

(a) Si verifichi l'operazione di composizione tra applicazioni rende $\mathrm{End}(V)$ un monoide (non commutativo).

(b) Si verifichi l'operazione di composizione tra applicazioni rende $\mathrm{Aut}(V)$ un gruppo (non abeliano).

Esercizi sul terzo capitolo

1. *Prodotto di matrici triangolari alte.* Siano $A, B \in \mathcal{M}_{2,2}(\mathbb{R})$. Si verifichi che sono veri i seguenti enunciati.

(i) Se A, B sono triangolari alte, tale è anche $A \cdot B$.

(ii) Se A, B sono triangolari alte, e gli elementi della diagonale principale di A e di B sono tutti uguali ad 1, lo stesso vale anche per $A \cdot B$.

2. *Prodotto di matrici e commutatività.* Siano $A, B \in \mathcal{M}_{n,n}(\mathbb{R})$ e sia

$$A = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

una matrice diagonale in cui gli elementi della diagonale sono a due a due distinti. Si provi che $A \cdot B = B \cdot A$ se e solo se B è diagonale.

3. *Prodotto di matrici righe per righe.* Si consideri l'insieme $\mathcal{M}_{2,2}(\mathbb{R})$ delle matrici quadrate di ordine 2.

(i) In analogia con il prodotto righe per colonne, si definisca formalmente un prodotto *righe per righe* in $\mathcal{M}_{2,2}(\mathbb{R})$.

(ii) Si considerino le matrici

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

e si verifichi, utilizzando tale prodotto, che $(AB)C \neq A(BC)$, e quindi non vale la proprietà associativa. Si verifichi, altresì, che $(BA)C = B(AC)$.

4. *Rango e determinante.* Si consideri la seguente matrice 4×4:

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

(i) Si calcoli il rango di A, con i vari metodi disponibili.

(ii) Si calcoli il determinante di A, con i vari metodi disponibili.

(iii) Si spieghi se è vero che il risultato di (i) implica quello di (ii) e/o viceversa.

5. *Matrice inversa.* Posto $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ e osservato che A è invertibile, si spieghi perché SICURAMENTE la sua inversa possiede elementi non interi.

6. *Inverse destre e sinistre.* Sia

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

(i) Si provi che A non ammette inverse a destra.

(ii) Si costruiscano alcuni esempi di inversa sinistra di A.

7. *Un sottomonoide.* Si consideri il monoide $(\mathcal{M}_{2,2}; \cdot)$ delle matrici quadrate di ordine 2 sul campo $\mathbb{R}$ con l'operazione di prodotto righe per colonne. Sia

$$S = \left\{ \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \mid a, b \in \mathbb{R} \right\} .$$

(i) Si provi che S è stabile rispetto al prodotto righe per colonne ed è quindi un sottomonoide di $\mathcal{M}_{2,2}$, essendo anche $I_2 \in S$.

(ii) Si provi che una matrice $A \in S$ è invertibile se e solo se $a \neq 0$ o $b \neq 0$.

(iii) Si provi che l'inversa di una matrice invertibile di S è ancora un elemento di S.

8. *Invertibilità.* Sia A una matrice 2×3 e B una matrice 3×2 (su $\mathbb{R}$).

(i) Si verifichi che $\det(B \cdot A) = 0$.

(ii) Posto

$$A = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 1 \end{pmatrix} \; ; \; B = \begin{pmatrix} 1 & 1 \\ 2 & 0 \\ t & t \end{pmatrix} ,$$

si stabilisca per quali valori del parametro t la matrice $A \cdot B$ risulta invertibile.

Esercizi sul quarto capitolo

1. *Cramer* 2×2. Si studi la compatibilità e si determinino le eventuali soluzioni dei seguenti sistemi lineari.

(i)

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

(ii)

$$\begin{cases} x+y=2 \\ x-y=2 \end{cases}$$

(iii)

$$\begin{cases} x+y=2 \\ x-y=t \end{cases} \qquad \text{al variare del parametro } t$$

2. *Cramer* 3×3. Si studi la compatibilità e si determinino le eventuali soluzioni dei seguenti sistemi lineari.

(i)

$$\begin{cases} x+y+z=3 \\ x-y=0 \\ x+3z=4 \end{cases}$$

(ii)

$$\begin{cases} x+y+z=3 \\ x-y=t-2 \\ x+2z=t+2 \end{cases} \qquad \text{al variare del parametro } t$$

3. *Cramer* 4×4. Si studi la compatibilità e si determinino le eventuali soluzioni del seguente sistema lineare. Si utilizzino entrambi i metodi.

$$\begin{cases} x_1+x_2+x_3+x_4=4 \\ x_1-x_2+x_3-x_4=0 \\ x_1+x_2+x_3-x_4=2 \\ x_1+2x_2+3x_3=6 \end{cases}$$

4. *Cramer* 4×4 *con parametro.* Si studi la compatibilità e si determinino le eventuali soluzioni del seguente sistema lineare al variare del parametro t. Si utilizzino entrambi i metodi.

$$\begin{cases} x_1+x_2+x_3+x_4=4 \\ x_1-x_2+x_3-x_4=t-1 \\ x_1+x_2+x_3-x_4=t+1 \\ x_1+2x_2+3x_3=6 \end{cases}$$

5. 3×3 *con parametro.* Si studi la compatibilità e si determinino le eventuali soluzioni del seguente sistema lineare al variare del parametro t. Si utilizzino entrambi i metodi.

$$\begin{cases} x_1+x_2+x_3=2 \\ x_1+x_2+tx_3=2 \\ tx_2+2x_3=1 \end{cases}$$

6. 3×4 *da ridurre.* Si studi la compatibilità e si determinino le eventuali soluzioni del seguente sistema lineare. Si utilizzino entrambi i metodi.

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

7. 3×4 *con parametro.* Si studi la compatibilità e si determinino le eventuali soluzioni del seguente sistema lineare. Si utilizzino entrambi i metodi.

$$\begin{cases} x_1 + 2x_2 + 3x_3 = 1 \\ 2x_1 + 3x_2 + 4x_3 + x_4 = 2 \\ x_1 + x_2 + x_3 + tx_4 = 1 \end{cases}$$

8. 2×2 *con due parametri.* Si studi la compatibilità e si determinino le eventuali soluzioni del seguente sistema lineare al variare dei parametri h, k. Si utilizzino entrambi i metodi.

$$\begin{cases} hx - ky = 2 \\ hx + ky = 2 \end{cases}$$

9. *Rappresentazioni di sottospazi.* Sia V uno spazio vettoriale reale di dimensione 3, e sia $\mathcal{B} = (\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ una sua base ordinata. Si determini una base del sottospazio W di V che ammette la seguente rappresentazione cartesiana:

$$W : \begin{cases} x + y - z = 0 \\ 2x + z = 0 \\ x - y + 2z = 0 \end{cases}$$

10. *Rappresentazioni di sottospazi.* Sia V uno spazio vettoriale reale di dimensione 3, e sia $\mathcal{B} = (\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ una sua base ordinata. Si considerino i vettori $\mathbf{u} = \mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3$, $\mathbf{w} = 2\mathbf{e}_1 - \mathbf{e}_3$ e il sottospazio $W = \mathcal{L}(\mathbf{u}, \mathbf{w})$. Si costruisca una rappresentazione cartesiana di W (rispetto alla base $\mathcal{B}$).

11. *Rappresentazioni di sottospazi.* Sia V uno spazio vettoriale reale di dimensione 4, e sia $\mathcal{B} = (\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4)$ una sua base ordinata. Si considerino i vettori $\mathbf{u} = \mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3 + \mathbf{e}_4$, $\mathbf{w} = 2\mathbf{e}_1 - \mathbf{e}_2 + \mathbf{e}_3$ e il sottospazio $W = \mathcal{L}(\mathbf{u}, \mathbf{w})$. Si costruisca una rappresentazione cartesiana di W (rispetto alla base $\mathcal{B}$).

Esercizi sul quinto capitolo

1. *Matrice associata ad f in varie basi, caso numerico.* Sia $f : \mathbb{R}^2 \to \mathbb{R}^3$ l'applicazione lineare definita ponendo

$$f(a, b) = (a + b, 2a, a - 2b) \quad \forall\ (a, b) \in \mathbb{R}^2\ .$$

Consideriamo le basi standard $\mathcal{B}$ e $\mathcal{B}'$ di $\mathbb{R}^2$ ed $\mathbb{R}^3$ rispettivamente, ed inoltre le basi $\mathcal{B}'' = \big((1,1),(1,-1)\big)$ di $\mathbb{R}^2$ e $\mathcal{B}''' = \big((1,0,1),(1,1,1),(1,1,-1)\big)$ di $\mathbb{R}^3$.

(i) Si determini la matrice $M_{\mathcal{B}}^{\mathcal{B}'}(f)$.

(ii) Si determini la matrice $M_{\mathcal{B}''}^{\mathcal{B}'}(f)$.

(iii) Si determini la matrice $M_{\mathcal{B}}^{\mathcal{B}'''}(f)$.

(iv) Si determini la matrice $M_{\mathcal{B}''}^{\mathcal{B}'''}(f)$.

2. *Matrice associata ad f in varie basi, caso astratto.* Siano V, V' due spazi vettoriali sul campo reale, e siano

$$\mathcal{B} = \big(\,\mathbf{e}_1, \mathbf{e}_2\,\big) \ ; \ \mathcal{B}' = \big(\,\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3\,\big) \ .$$

Sia $f : V \to V'$ l'applicazione lineare definita ponendo

$$f(\mathbf{e}_1) = \mathbf{e}'_1 + \mathbf{e}'_2 \quad ; \quad f(\mathbf{e}_2) = \mathbf{e}'_1 - \mathbf{e}'_2 + \mathbf{e}'_3$$

ed estendendo poi per linearità.

(i) Si verifichi che $\mathcal{B}'' = (\mathbf{e}_1 + \mathbf{e}_2, \mathbf{e}_1 - \mathbf{e}_2)$ è una base di V e che $\mathcal{B}''' = (\mathbf{e}'_1 + \mathbf{e}'_2, \mathbf{e}'_2 + \mathbf{e}'_3, \mathbf{e}'_1 + \mathbf{e}'_2 + \mathbf{e}'_3)$ è una base di V'.

(ii) Si determini la matrice $M_{\mathcal{B}}^{\mathcal{B}'}(f)$.

(iii) Si determini la matrice $M_{\mathcal{B}''}^{\mathcal{B}'''}(f)$.

3. *Rango del prodotto di matrici.* Siano $A \in \mathcal{M}_{m,n}$, $B \in \mathcal{M}_{n,p}$ due matrici compatibili.

(i) Come conseguenza dell'Esercizio 14 del secondo capitolo, parte (i) e parte (ii), si verifichi che $\rho(B \cdot A) \leq \rho(A)$ e $\rho(B \cdot A) \leq \rho(B)$.

(ii) Si verifichi che due matrici coniugate hanno lo stesso rango.

4. *Cambi di base.* Sia V uno spazio vettoriale su un campo $\mathbb{K}$ e sia $\mathcal{B} = (\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ una sua base ordinata. Si consideri la matrice

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

(i) Si determini una base $\mathcal{B}'$ tale che $M_{\mathcal{B}'}^{\mathcal{B}} = A$.

$\big[\,\mathcal{B}'=(\mathbf{e}_1+\mathbf{e}_3, \mathbf{e}_1+2\mathbf{e}_3, \mathbf{e}_2)\,\big]$

(ii) Si determini una base $\mathcal{B}''$ tale che $M_{\mathcal{B}}^{\mathcal{B}''} = A$.

[si calcoli A^{-1} e si proceda come in (i)]

5. *Derivate ed autovalori.* Si consideri lo spazio vettoriale $V = \mathbb{R}[x]$ dei polinomi, e sia $D : \mathbb{R}[x] \to \mathbb{R}[x]$ l'operatore di derivazione, definito ponendo

$$D(a_0 + a_1 x + \cdots + a_n x^n) = a_1 + 2a_2 x + \cdots + n a_n x^{n-1} \ .$$

(i) Si verifichi che 0 è un autovalore per D e si determini l'autospazio corrispondente.

(ii) Si verifichi che non possono esistere autovalori non nulli.

6. *Autovalori ed automorfismi.* Sia V uno spazio vettoriale su un campo $\mathbb{K}$ e sia $f \in \mathrm{Aut}(V)$ (ovvero sia f un isomorfismo di V in V).

(i) Si verifichi che λ è un autovalore per f se e solo se λ^{-1} è un autovalore per f^{-1}.

(ii) Sia λ uno scalare. Si verifichi che se λ è un autovalore per f gli autospazi $V_\lambda(f)$ e $V_{\lambda^{-1}}(f^{-1})$ coincidono.

(iii) Si verifichi che f è diagonalizzabile se e solo se f^{-1} è diagonalizzabile.

7. *Diagonalizzabilità su vari campi.* Sia V uno spazio vettoriale su un campo $\mathbb{K}$, sia $\mathcal{B} = (\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4)$ una sua base ordinata, e sia $f \in \mathrm{End}(V)$ definito ponendo

$$f(\mathbf{e}_1) = -\mathbf{e}_2 \; ; f(\mathbf{e}_2) = \mathbf{e}_1 \; ; f(\mathbf{e}_3) = \mathbf{e}_1 + 2\mathbf{e}_4 \; ; f(\mathbf{e}_4) = 2\mathbf{e}_2 + \mathbf{e}_3 \; .$$

(i) Si studi la diagonalizzabilità di f nel caso in cui $\mathbb{K} = \mathbb{Q}$.

(ii) Si studi la diagonalizzabilità di f nel caso in cui $\mathbb{K} = \mathbb{R}$.

(iii) Si studi la diagonalizzabilità di f nel caso in cui $\mathbb{K} = \mathbb{C}$.

8. *Diagonalizzabilità con due parametri.* Sia V uno spazio vettoriale sul campo $\mathbb{R}$, sia $\mathcal{B} = (\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ una sua base ordinata, e sia $f \in \mathrm{End}(V)$ definito ponendo

$$f(\mathbf{e}_1) = \mathbf{e}_1 + b\mathbf{e}_3 \; ; f(\mathbf{e}_2) = 2\mathbf{e}_2 \; ; f(\mathbf{e}_3) = a\mathbf{e}_1 + \mathbf{e}_3 \; ,$$

dove a, b sono parametri reali.

(i) Si stabilisca per quali valori dei parametri a, b l'endomorfismo f ammette autovalori multipli.

(ii) Si stabilisca per quali valori dei parametri a, b l'endomorfismo f non soddisfa alla prima parte della quarta condizione del Teorema Spettrale.

(iii) Si studi la diagonalizzabilità di f.

9. *Diagonalizzabilità con un parametro.* Sia V uno spazio vettoriale sul campo $\mathbb{R}$, sia $\mathcal{B} = (\mathbf{e}_1, \mathbf{e}_2)$ una sua base ordinata, e sia $f \in \mathrm{End}(V)$ definito ponendo

$$f(\mathbf{e}_1) = t^2\mathbf{e}_1 \; ; f(\mathbf{e}_2) = (t-1)\mathbf{e}_1 + (2-t^2)\mathbf{e}_2 \; ,$$

dove t è un parametro reale. Si discuta la diagonalizzabilità di f al variare del parametro t.

Esercizi sul sesto capitolo

1. *Vettori isotropi.* Si consideri la forma bilineare dell'Esempio 6.10.

(i) Si determinino tutti i vettori isotropi.

(ii) Si stabilisca se essi formano un sottospazio.

2. *Complemento ortogonale.* Sia V uno spazio vettoriale euclideo di dimensione 4, e sia $\mathcal{B} = (\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4)$ una sua fissata base ordinata ortonormale. Sia $W \leq V$.

(i) Si provi che $V = W \oplus W^{\perp}$.

(ii) Si considerino i vettori $\mathbf{u} = (1, 0, 1, 0)$ e $\mathbf{v} = (1, 1, 2, 2)$. Posto $W = \mathcal{L}(\mathbf{u}, \mathbf{v})$, si descriva $W^{\perp}$.

3. *Gruppo ortogonale.* Sia O_2 il gruppo delle matrici ortogonali 2×2 su $\mathbb{R}$.

(i) Si verifichi che per ogni $\theta \in \mathbb{R}$ la matrice

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

appartiene ad O_2.

(ii) Si verifichi che l'insieme $SO_2 = \{A_\theta \mid \theta \in \mathbb{R}\}$ è un sottogruppo (abeliano) del gruppo (non abeliano) O_2.

(iii) Si verifichi che la matrice $\left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$ appartiene ad O_2 ma non ad SO_2.

4. *Gram-Schmidt.* Si consideri lo spazio vettoriale euclideo $\mathbb{R}^3$ ed il suo sistema di vettori

$$\mathcal{S} = \big((1, 1, 1), (1, 2, 2), (1, 2, 1) \big) \ .$$

(i) Si verifichi che $\mathcal{S}$ è una base, ma non ortonormale.

(ii) Si ortonormalizzi $\mathcal{S}$.

5. *Diagonalizzazione ortogonale.* Si consideri uno spazio vettoriale euclideo V ed una sua base ortonormale $\mathcal{B} = (\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$. Sia $f : V \to V$ l'endomorfismo definito ponendo

$$f(\mathbf{e}_1) = 2\mathbf{e}_3 \ ; \ f(\mathbf{e}_2) = 2\mathbf{e}_2 \ ; \ f(\mathbf{e}_3) = 2\mathbf{e}_1$$

ed estendendo poi per linearità.

(i) Si verifichi che f è ortogonalmente diagonalizzabile.

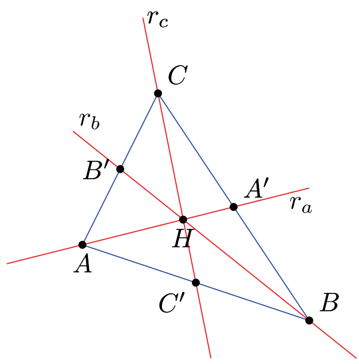
(ii) Si determini una base ortonormale di autovettori.

Esercizi sul settimo capitolo

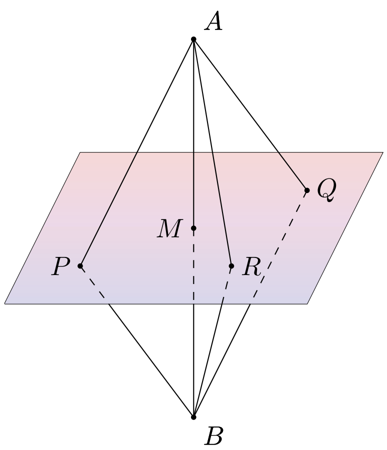
1. *Parallelismo.* Sia $\mathbb{A}$ uno spazio affine reale di dimensione n. Si provi la non transitività della relazione di parallelismo tra sottospazi.

2. *Punto medio.* Sia $\mathbb{A}$ uno spazio affine reale di dimensione n. Si provi l'esistenza ed unicità del punto medio tra due punti.

3. *Baricentro.* Sia $\mathbb{A}$ un piano affine reale e siano A, B, C tre punti non allineati. Indichiamo com A', B', C' i punti medi delle coppie di punti (B, C), (A, C), (A, B) rispettivamente. Si provi che le rette r_a, r_b, r_c individuate dalle coppie (A, A'), (B, B'), (C, C') si incontrano in un punto H, che è noto come baricentro della terna (A, B, C), o anche del triangolo di vertici A, B, C.



4. *Asse di due punti.* Sia $\mathbb{A}$ uno spazio euclideo reale di dimensione n. Si provi che l'insieme dei punti equidistanti da due punti dati è un iperpiano.



5. *Distanze in* $\mathbb{E}^2$. Si consideri un punto $P_0 \equiv (x_0, y_0) \in \mathbb{E}^2$ ed una retta $r : ax + by + c = 0$. Si verifichi che

$$d(P_0, r) = \frac{|ax_0 + by_0 + c|}{\sqrt{a^2 + b^2}} .$$

6. *Prodotto misto.* Nello spazio (di dimensione 3) dei vettori liberi ordinari associato allo spazio euclideo tridimensionale, considerati tre vettori $\mathbf{u}, \mathbf{v}, \mathbf{w}$, lo scalare $\mathbf{u} \wedge \mathbf{v} \cdot \mathbf{w}$ si dice prodotto misto dei tre vettori. In un riferimento ortonormale, poniamo $\mathbf{u} = (\ell, m, n), \mathbf{v} = (\ell', m', n'), \mathbf{w} = (\ell'', m'', n'')$ e

$$A = \begin{pmatrix} \ell & m & n \\ \ell' & m' & n' \\ \ell'' & m'' & n'' \end{pmatrix} .$$

(i) Si verifichi che $\mathbf{u} \wedge \mathbf{v} \cdot \mathbf{w} = \mathbf{v} \wedge \mathbf{w} \cdot \mathbf{u} = \mathbf{w} \wedge \mathbf{u} \cdot \mathbf{v}$.

(ii) Si verifichi che $|\mathbf{u} \wedge \mathbf{v} \cdot \mathbf{w}| = \det A$.

(iii) Siano $A, B, C, D \in \mathbb{E}^3$ e sia

$$\overrightarrow{AB} = \mathbf{u} \ ; \ \overrightarrow{AC} = \mathbf{v} \ ; \ \overrightarrow{AD} = \mathbf{w} .$$

Si verifichi che $|\mathbf{u} \wedge \mathbf{v} \cdot \mathbf{w}|$ è il volume del parallelepipedo che ammette $\bar{A}B, \bar{A}C, \bar{A}D$ come spigoli.

7. *Piani in* $\mathbb{E}^4$. Si costruiscano due piani π, π' in uno spazio affine di dimensione 4, in modo che l'intersezione $\pi \cap \pi'$ sia un punto.

8. *Fasci di piani.* Sia $\mathcal{F}$ un fascio di piani in $\mathbb{E}^3$.

(i) Nel caso in cui $\mathcal{F}$ è improprio, si provi che $\forall\ P \in \mathbb{E}^3\ \exists!\ \pi \in \mathcal{F}$ tale che $P \in \pi$.

(ii) Nel caso in cui $\mathcal{F}$ è proprio, ad esempio di asse r, si provi che

$$\forall\ P \in \mathbb{E}^3 - r\ \exists!\ \pi \in \mathcal{F} \ \text{ tale che } P \in \pi .$$

9. *Rette e fasci di piani.* Siano r ed s due rette distinte in $\mathbb{E}^3$ e si consideri il fascio proprio di piani $\mathcal{F}_r$ di asse r. Si provi che esiste un (unico) piano $\pi \in \mathcal{F}_r$ che contiene s se e solo se r ed s sono incidenti o parallele.

10. *Rappresentazione di sottospazi.* Si considerino quattro punti A, B, C, D in $\mathbb{E}^4$. In un fissato riferimento sia

$$A \equiv (1, 1, 1, 1) \ , \ B \equiv (1, 2, 1, 2) \ , \ C \equiv (2, 1, 2, 1) \ , \ D \equiv (2, 2, 2, 2) .$$

(i) Si determinino i vettori $\overrightarrow{AB}, \overrightarrow{AC}, \overrightarrow{AD}$.

(ii) Si determini il più piccolo sottospazio affine contenente i punti A, B (e si verifichi che è una retta).

(iii) Si determini il più piccolo sottospazio affine contenente i punti A, B, C (e si verifichi che è un piano).

(iv) Si determini il più piccolo sottospazio affine contenente i punti A, B, C, D e si verifichi che esso coincide con il sottospazio del punto precedente.

Esercizi sull'ottavo capitolo

1. *Circonferenze.* Sia

$$\gamma : b_{1,1}x_1^2 + 2b_{1,2}x_1x_2 + b_{2,2}x_2^2 + 2b_{1,3}x_1x_3 + 2b_{2,3}x_2x_3 + b_{3,3}x_3^2 = 0$$

una conica. Si verifichi che γ è una circonferenza se e solo se

$$b_{1,1} = b_{2,2},\ b_{1,2} = 0\ , b_{1,1}b_{3,3} \leq b_{1,3}^2 + b_{2,3}^2\ .$$

2. *Coniche prive di punti reali.* Si spieghi perché una conica priva di punti reali è necessariamente un'ellisse.

3. *Ellissi ed angoli.* Sia γ un'ellisse, siano F', F'' i suoi fuochi, e sia $P \in \gamma$. Ad esempio si ponga

$$\gamma : x^2 + 2xy + 2x - 3 = 0\ ;\ P \equiv (1, -1)\ .$$

(i) Si determini la tangente $r = \omega(P)$.

(ii) Si calcolino la coordinate dei fuochi.

(iii) Siano r', r'' le rette passanti per P e per F', F'' rispettivamente. Si verifichi che gli angoli α e β che r forma con r' ed r'' coincidono.

(iv) Si verifichi tale proprietà in generale (difficile!!).

4. *Coniche a centro.* Si considerino le seguenti coniche in coordinate affini

$$\gamma_\pm : 2x^2 + 2xy \pm y^2 + 2x = 0\ .$$

(i) Si scrivano le corrispondenti equazioni in coordinate omogenee.

(ii) Si provi che γ_+ è un'ellisse ordinaria e γ_- è una iperbole.

(iii) Si calcolino gli asintoti di γ_+ e γ_-.

(iv) Si calcolino gli assi di γ_+ e γ_-.

(v) Si calcoli il centro di γ_+ e γ_-.

(vi) Si calcolino i vertici di γ_+ e γ_-.

5. *Polarità.* Si considerino i punti

$$O \equiv (0,0),\ A \equiv (-1,0),\ B \equiv (-1,2),\ C \equiv (1,1),\ D \equiv (-1,1)$$

e l'ellisse ordinaria

$$\gamma_+ : 2x^2 + 2xy + y^2 + 2x = 0$$

(i) Si provi che $O, A, B \in \gamma_+$ e si determini la tangente alla conica in tali punti.

(ii) Si provi che $C, D \notin \gamma_+$ e si determini la polare in tali punti.

6. *Diametri.* Si consideri l'iperbole $\gamma_- : 2x^2 + 2xy - y^2 + 2x = 0$.

(i) Si determini un diametro $r = \omega(P_\infty)$ di γ_-.

(ii) Si determinino i punti di $r \cap \gamma_-$.

(iii) Si determinino le tangenti s' ed s'' in tali punti.

(iv) Si verifichi che $s' \parallel s''$.

7. *Fuochi.* Si determinino i fuochi delle coniche γ_+ e γ_- degli esercizi precedenti.

8. *Polarità definita da una conica semplicemente degenere.* Siano r', r'' due rette distinte, e sia $\gamma = r' \cup r''$ (conica semplicemente degenere). Indichiamo con V il punto in cui r' ed r'' si intersecano. Sia infine s' una retta per V.

(i) Si provi che se $P, Q \in s'$ $(P, Q \neq V)$ si ha $\omega(P) = \omega(Q)$.

(ii) Si indichi con s'' la polare comune a tutti i punti di s' distinti da V. Si provi che per ogni $R \in s''$ si ha $\omega(R) = s'$.

(iii) Nel fascio di rette $\mathcal{F}$ di centro V si definisca un'applicazione $\ell : \mathcal{F} \to \mathcal{F}$ al modo seguente. Per ogni $s' \in \mathcal{F}$ si ponga $\ell(s') = s''$, la polare comune di tutti i punti di s' distinti da V. Si provi che ℓ ha carattere involutorio, ovvero $\ell \circ \ell = id_{\mathcal{F}}$, e che le rette r', r'' sono (tutti e soli i) punti fissi di tale involuzione, ovvero $\ell(r') = r'$ ed $\ell(r'') = r''$.

9. *Famiglie di coniche dipendenti da un parametro.* Si consideri, al variare del parametro reale t, la conica $\gamma_t : tx^2 + 2xy + 2y^2 + 2tx + 1 = 0$.

(i) Si determini, per ogni valore di t, il luogo $\gamma_t \cap r_\infty$ dei punti impropri.

(ii) Si classifichi, per ogni valore di t, la conica γ_t.

(iii) Si scelga il valore di t per cui γ_t sia una parabola, e si determini l'asse.

(iv) Si ponga $t = 1$ e si calcoli il centro.

(v) Si ponga $t = 0$ e si determinino gli asintoti.

10. *Famiglie di coniche dipendenti da un parametro.* Si consideri, al variare del parametro reale t, la conica $\gamma_t : tx^2 + 2xy + y^2 + 2tx + 2y + 1 = 0$.

(i) Si determini, per ogni valore di t, il luogo $\gamma_t \cap r_\infty$ dei punti impropri.

(ii) Si classifichi, per ogni valore di t, la conica γ_t.

(iii) Si scelga il valore di t per cui la conica γ_t sia doppiamente degenere, e si determini la retta (tutta costituita di punti doppi) in cui si spezza la conica.

(iv) Si ponga $t = 2$ e si calcolino i vertici della conica.

(v) Si ponga $t = 0$ e si determinino gli asintoti.

AREE SCIENTIFICO–DISCIPLINARI

AREA 01 – **Scienze matematiche e informatiche**

AREA 02 – Scienze fisiche

AREA 03 – Scienze chimiche

AREA 04 – Scienze della terra

AREA 05 – Scienze biologiche

AREA 06 – Scienze mediche

AREA 07 – Scienze agrarie e veterinarie

AREA 08 – Ingegneria civile e architettura

AREA 09 – Ingegneria industriale e dell'informazione

AREA 10 – Scienze dell'antichità, filologico–letterarie e storico–artistiche

AREA 11 – Scienze storiche, filosofiche, pedagogiche e psicologiche

AREA 12 – Scienze giuridiche

AREA 13 – Scienze economiche e statistiche

AREA 14 – Scienze politiche e sociali

Il catalogo delle pubblicazioni di Aracne editrice è su

www.aracneeditrice.it

Compilato il 2 agosto 2013, ore 15:23
con il sistema tipografico $\LaTeX\,2_\epsilon$

Finito di stampare nel mese di agosto del 2013
dalla «ERMES. Servizi Editoriali Integrati S.r.l.»
00040 Ariccia (RM) – via Quarto Negroni, 15
per conto della «Aracne editrice S.r.l.» di Roma