Lezioni di Geometria e Algebra

Fulvio Bisi, Francesco Bonsante, Sonia Brivio

CAPITOLO 0

Preliminari.

1. Insiemistica e logica

Il presente Capitolo introduttivo ha lo scopo di ripassare alcuni argomenti di base già visti nei corsi di studio liceali o tecnici superiori, approfondendo alcuni aspetti, allo scopo di porre le basi per un linguaggio matematico comune, corredato di simboli definiti e condivisi.

1.1. Insiemi.

La prima parte di questo Capitolo verrà dedicata a richiami riguardanti la teoria degli insiemi cosiddetta *elementare*, per distinguerla dalla teoria assiomatica degli insiemi, ben più formale, che esula dagli scopi di questi appunti e del corso stesso.

Cominciamo ricordando che assumeremo come primitivo il concetto di insieme, e che normalmente gli elementi degli insiemi che vengono presi in considerazioni formano l'insieme *ambiente* o *universo* U. Come possiamo caratterizzare un insieme A? La risposta è: *in qualunque maniera possiamo stabilire se un dato elemento appartiene* ($\in$)*o non appartiene* ($\notin$) *all'insieme* A; Questo si può ottenere, sostanzialmente, in due maniere principali.

(1) Fornendo l'elenco degli elementi dell'insieme $\{a_1, a_2, a_3\}$; ad esempio

$$A = \left\{1, 3, 7, \frac{1}{2}, -2.5, \pi\right\}$$

Notiamo che l'ordine con il quale vengono elencati gli elementi *non* è significativo. In altre parole $\{1, 2, 3\}$ e $\{3, 1, 2\}$ indicheranno lo **stesso** insieme.

In base a questo, potremo scrivere

$$\frac{1}{2} \in A, \quad 1 \in A, \quad 34 \notin A$$

(2) Caratterizzando gli elementi dell'insieme dicendo che sono quelli presi nell'universo e godono di una certa proprietà $A = \{a \in U \,|\, p(a)\}$; la barretta verticale $|$, o i due punti $:$ corrispondono, in questo contesto, alla locuzione "tale che"; quindi possiamo leggere la precedente scrittura come "A è l'insieme degli a minuscolo appartenenti ad U tali che vale la proprietà $p(a)$, dove $p(a)$ indica un qualunque enunciato che riguarda a. Ad esempio, se indichiamo con $\mathbb{N}$ l'insieme dei numeri

naturali (quelli, per capirsi, che servono al pastore per contare le sue pecore: 0 –se non ne ha nessuna– $1, 2, 3, \dots$) l'insieme

$$P = \{n \in \mathbb{N}\colon n = 2k \text{ per qualche } k \in \mathbb{N}\}$$

indica l'insieme dei numeri naturali pari. Pertanto, fra l'altro scriveremo

$$0 \in P, \quad 34 \in P, \quad 3 \notin P, \quad \pi \notin P\,.$$

La seconda modalità diventa essenziale nel momento in cui, come il caso dell'esempio, non esiste un numero finito di elementi, per cui l'elenco sarebbe impossibile. A tale proposito, ricordiamo subito che, se un insieme contiene un numero n *finito* di elementi, questo corrisponde con la *cardinalità* dell'insieme: $n = \operatorname{card}(A)$.

L'insieme che non contiene nessun elemento viene detto *insieme vuoto*, e si indica con il simbolo $\varnothing$. Osserviamo che esso è unico.

Definizione 0.1. Diremo che B è sottoinsieme di A ($B \subseteq A$) se ogni elemento di B è anche elemento di A. In simboli, introducendo il simbolo dell'implicazione (di cui parleremo estesamente più avanti):

$$b \in B \implies b \in A.$$

Se almeno un elemento di A non appartiene a B diremo che si ha un *inclusione propria*, ossia che B è un sottoinsieme proprio di A (Figura 0.1):

$$B \subset A\colon B \subseteq A, \text{ ed esiste } x \in A \,|\, x \notin B$$

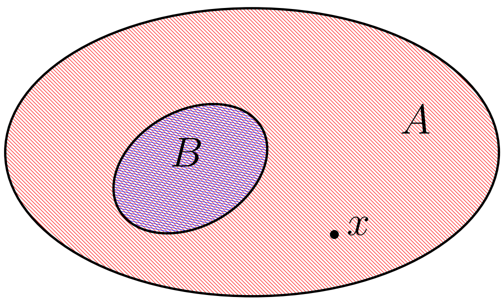

Figura 0.1. Rappresentazione schematica mediante i diagrammi di Eulero-Venn dell'inclusione propria; $B \subset A$ poiché esiste un elemento x di A che *NON* appartiene a B.

Due insiemi sono *uguali* (o coincidono) se contengono gli stessi elementi, ossia se $A \subseteq B$ e $B \subseteq A$. Ad esempio, con i simboli di prima, l'insieme $A = \{n \in P \,|\, n \le 10\}$ e l'insieme $B = \{0, 2, 4, 6, 8, 10\}$ contengono esattamente gli stessi elementi, dunque sono uguali: $A = B$.

Ogni studente nel proprio corso di studi ha incontrato insiemi di numeri ed ha imparato a svolgere alcune operazioni in questi insiemi. Il primo di questi è l'insieme $\mathbb{N}$ dei **numeri naturali**:

$$\mathbb{N} = \{0, 1, 2, 3, \dots\}$$

Il secondo insieme è l'insieme $\mathbb{Z}$ dei **numeri interi o relativi**:

$$\mathbb{Z} = \{0, \pm 1, \pm 2, \pm 3, \dots\}$$

ricordiamo che il sottoinsieme di $\mathbb{Z}$ degli interi positivi viene identificato con l'insieme dei numeri naturali. Il terzo insieme è l'insieme $\mathbb{Q}$ dei **numeri razionali**, ossia delle frazioni:

$$\mathbb{Q} = \left\{ \pm\frac{1}{2}, \pm\frac{1}{3}, \pm\frac{2}{3}, \pm\frac{4}{1}, \dots \right\}$$

ricordiamo che il sottoinsieme di $\mathbb{Q}$ delle frazioni apparenti (ossia, quelle con denominatore unitario, o equivalenti a queste ultime) viene identificato con $\mathbb{Z}$. Infine l'ultimo insieme numerico conosciuto è l'insieme $\mathbb{R}$ dei **numeri reali**, di cui fanno parte i **numeri razionali** ed i **numeri irrazionali**, cioè i numeri decimali che non sono periodici. Sono numeri irrazionali ad esempio:

$$\sqrt{2},\ \pi,\ \mathrm{e}.$$

Abbiamo quindi l'inclusione di insiemi:

$$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}.$$

Conveniamo di indicare con un asterisco $()^*$ l'insieme numerico corrispondente privato dello zero:

$$\mathbb{N}^* := \mathbb{N} - \{0\},$$

eccetera.

Riprenderemo in seguito nella sezione 2 in modo più formale questi insiemi numerici, introducendo anche l'ultimo insieme che useremo a volte nel corso: quello dei numeri complessi $\mathbb{C}$.

Vi sono due operazioni fondamentali che possiamo eseguire con gli insiemi (Figura 0.2):

Definizione 0.2. Siano A e B due insiemi.

- **L'unione** dei due insiemi, indicata con $A \cup B$, è l'insieme che contiene gli elementi di A e B, in comune e non.
- **L'intersezione** dei due insiemi, indicata con $A \cap B$, è l'insieme che contiene gli elementi che A e B hanno in comune.
- L'insieme **differenza** $A - B$ è formato privando A degli (eventuali) elementi che sono anche in B (ossia, togliendo ad A gli elementi di $A \cap B$).

Come caratterizziamo tutti i possibili sottoinsiemi di un dato insieme?

Definizione 0.3. Dato un insieme A, l'insieme formato da tutti i sottoinsiemi di A viene detto insieme delle parti di A:

$$\mathcal{P}(A) := \{B \mid B \subseteq A\}$$

Osservazione 0.1. Con il simbolo $:=$ indicheremo le uguaglianze per definizione.

Esempio 0.2. Sia $A = \{1, 5, 7\}$. L'insieme delle parti di A è:

$$\mathcal{P}(A) = \{\varnothing, A, \{1\}, \{5\}, \{7\}\{1, 5\}, \{5, 7\}, \{1, 7\}\}.$$

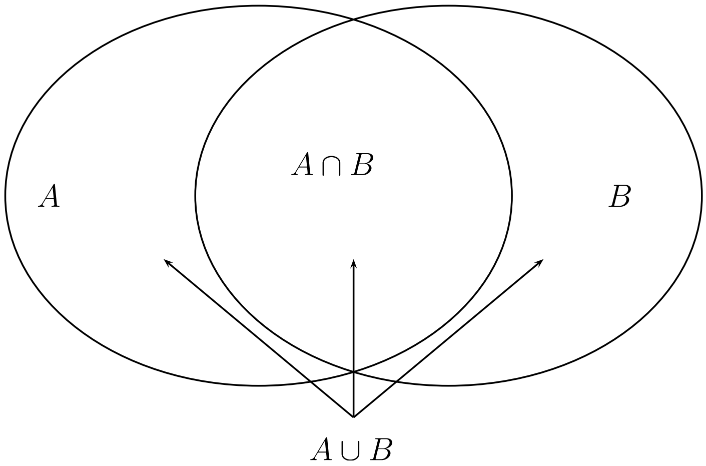

FIGURA 0.2. Rappresentazione schematica mediante i diagrammi di Eulero-Venn degli insiemi A, B, della loro intersezione $A \cap B$ e dell'unione $A \cup B$ (le tre aree insieme).

Se il numero di elementi di un insieme è finito, ha senso chiedersi quanti insiemi si troveranno in $\mathcal{P}(A)$. La risposta viene da un teorema che dimostreremo basandoci su un noto principio valido per i numeri naturali:

PROPOSIZIONE 0.1 (Principio di induzione). *Sia A un sottoinsieme di $\mathbb{N}$ caratterizzato dalle seguenti proprietà:*

(1) $0 \in A$;
(2) *se* $n \in A$, *allora* $n + 1 \in A$.

L'insieme A coincide con $\mathbb{N}$.

Una dimostrazione che una certa proprietà vale per tutti i numeri $n \in \mathbb{N}$ può essere basata sul Principio 0.1 mostrando che la proprietà vale per 0 (o per 1, se si parte da questo numero) e che, se vale per n *(ipotesi induttiva)*, allora vale per $n + 1$.

Siamo in grado ora di enunciare il nostro semplice

TEOREMA 0.2. *Sia A un insieme di cardinalità* $\text{card}(A) = n \in \mathbb{N}$. *La cardinalità di $\mathcal{P}(A)$ è allora 2^n.*

DIMOSTRAZIONE. Forniamo la dimostrazione come esempio di applicazione del principio di induzione.

- Il primo passo è verificare che l'affermazione contenuta nella tesi sia vera per $n = 1$ (in realtà, vale anche per 0, visto che la cardinalità dell'insieme vuoto è nulla, per definizione: per maggiore chiarezza, partiamo con il caso più concreto)

  Se A ha cardinalità $n = 1$, contiene solo 1 elemento; i suoi possibili sottoinsiemi sono l'insieme vuoto $\varnothing$ e l'insieme stesso A: l'insieme delle parti contiene 2 elementi, quindi la cardinalità $\text{card}(A) = 2 = 2^1$.
- Supponiamo ora (*ipotesi induttiva*) che la tesi sia vera per un certo numero n:

$$\text{card}(A) = n \implies \text{card}(\mathcal{P}(A)) = 2^n.$$

Mostriamo che allora è vera per $n+1$; costruiamo un insieme B unendo ad A un insieme B che contiene un solo elemento, diverso da tutti quelli di A. Ad esempio

$$\begin{aligned} A &= \{\diamondsuit, \clubsuit, \spadesuit, \heartsuit\} \\ B &= \{\bullet\} \\ A \cup B &= \{\diamondsuit, \clubsuit, \spadesuit, \heartsuit, \bullet\}\,. \end{aligned}$$

È chiaro che $A \cup B$ contiene $n+1$ elementi. Quanti sottoinsiemi possiamo fare per esso?

Tutti quelli che sono sottoinsiemi di A, che sono 2^n, e poi tutti quelli che otteniamo unendo a ciascuno di essi l'elemento $\bullet$: questi sono tutti diversi da quelli di A soltanto (perché contengono un elemento che *non* appartiene ad A), e non ve ne sono altri. Quindi, in tutto abbiamo $2^n + 2^n$ sottoinsiemi, cioè $2(2^n) = 2^{n+1}$, come volevamo dimostrare.

Quindi, siccome la tesi vale per $n = 1$, vale anche per $n = 2 = 1 + 1$; allora vale per $n = 3 = 2 + 1$; allora vale per $n = 4 = 3 + 1$, ecc.

Vale, insomma, per ogni $n \in \mathbb{N}$ (per $n = 0$, $2^0 = 1$, e l'unico sottoinsieme dell'insieme vuoto è l'insieme vuoto stesso, quindi anche in questo caso l'affermazione è corretta). $\square$

1.2. Logica matematica elementare.

Nella logica matematica, si considerano *proposizioni* che hanno la possibilità di essere valutate come vere (V) o false (F) senza ambiguità (ossia, per esempio, senza che questo riguardi un'opinione o simili). Le proposizioni vengono di solito indicati con le lettere minuscole: $p, q, r, \ldots$; può accadere che una data proposione contenga una o più variabili, che vengono messe tra parentesi: $p(x)$, $q(x, y), \ldots$. Ad esempio

$$p = \text{“7 è un numero pari”},$$

$$q = \text{“36 è divisibile per 9”};$$

$$r(x) = \text{“}x\text{ è un numero primo”}.$$

Evidentemente, p è falsa, q è vera, $r(x)$ è vera per $x = 31$, e falsa per $x = 25$, ad esempio.

La negazione di una proposizione è simboleggiata da una barra sopra la lettera che indica la proposizione. Per esempio, se

$$p = \text{“42 è divisibile per 7”}$$

la sua negazione è

$$\overline{p} = \text{“42 NON è divisibile per 7”}.$$

Per formulare enunciati, risulta utile l'utilizzo di due simboli, detti *quantificatori*:

(1) il quantificatore universale $\forall$ ("per ogni"): indica che ciò che segue si assume per tutti gli esemplari indicati, senza esclusione. Per esempio, se A indica l'insieme dei numeri primi, possiamo scrivere

$$\forall n \in A,\ n \text{ ha come divisori solo se stesso e } 1.$$

(2) il quantificatore esistenziale $\exists$ ("esiste"): indica l'esistenza di almeno un elemento, per cui vale ciò che segue, ma non necessariamente questo deve valere per tutti, o esiste un solo elemento che si comporta in tale maniera (se si vuole indicare anche l'unicità, si fa seguire al simbolo un punto esclamativo: $\exists!$, che si legge "esiste un unico"). Per esempio, se con A indichiamo i numeri primi, e con B i numeri naturali maggiori di 1, possiamo dire che l'enunciato

$$\forall n \in B\ \exists p \in A \,|\, p \text{ è divisore di } n$$

è vero (convincetevi di questo ragionando sul suo significato).

Occorre fare attenzione alla negazione di enunciati che contengono i quantificatori: se voglio negare *tutti i bottoni di questo cassetto sono neri* (una quantificazione "per ogni") basta trovare *almeno UN bottone NON nero* (una quantificazione "di esistenza del contrario"). In simboli, possiamo scrivere che l'opposto di

$$\forall x, \quad p(x)$$

sarà

$$\exists x \,|\, \overline{p(x)}.$$

Vale anche il viceversa: la negazione di

$$\exists y \,|\, q(y)$$

è

$$\forall y, \quad \overline{q(y)}.$$

I due connettivi logici fondamentali sono la congiunzione ("e"), che rappresentiamo con il simbolo $\wedge$ e la disgiunzione ("o, oppure"), che rappresentiamo con il simbolo $\vee$. Osserviamo che la proposizione

$$p \wedge q$$

è vera solo se p e q sono vere, altrimenti è falsa; invece, la proposizione

$$p \vee q$$

è vera se *almeno* una delle due fra p e q è vera, e falsa solo se sono entrambe false; in altre parole, non si intende un "o" esclusivo (per chi ha dimestichezza con il latino, corrisponde al connettivo *vel*, non al connettivo *aut*).

Basandoci sui connettivi, possiamo definire formalmente le operazioni di base sugli insiemi (Figura 0.2): *l'unione fra due insiemi* A e B sarà quindi

$$A \cup B := \{x \in U \,|\, x \in A \vee x \in B\}$$

e *l'intersezione*

$$A \cap B := \{x \in U \,|\, x \in A \wedge x \in B\}$$

Una struttura importante della logica, e della matematica, è quello di *implicazione*; i teoremi sono normalmente enunciati in forma di implicazione del tipo "se queste condizioni si verificano (ipotesi), allora queste conseguenze si devono verificare (tesi)". In simboli descriveremo la struttura come

$$p \implies q \tag{0.1}$$

che possiamo leggere

- p implica q;
- se p allora q;
- p è condizione sufficiente (c.s.) per q;
- q è condizione necessaria (c.n.) per p.

La doppia implicazione, ossia l'implicazione reciproca di p e q corrisponde all'equivalenza; pertanto in simboli

$$p \iff q$$

si leggerà

- p se e solo se (sse) q;
- p equivale a q;
- p è condizione necessaria e sufficiente per q (c.n.&s.)

L'enunciato simbolico (0.1) equivale a dire che *se non vale* q*, allora non può valere* p, che chiameremo l'enunciato *contronominale* del primo: in simboli, scriveremo che

$$(p \implies q) \quad \iff \quad (\overline{q} \implies \overline{p}) \tag{0.2}$$

Questo viene utile, perché a volte l'enunciato contronominale è di più agevole approccio; ad esempio, invece di "se n è dispari allora non è divisibile per 10" risulta più chiaro o più facilmente dimostrabile la forma contronominale: "se n è divisibile per 10 allora è pari (non è dispari)".

Analogamente, a volte per dimostrare un enunciato $p \implies q$, si può procedere *per assurdo*, ossia, si suppone che l'ipotesi p sia vera, e si mostra che, se la tesi q fosse falsa, si arriva ad una contraddizione (per esempio, che p è falsa, oppure che una proposizione r è contemporaneamente vera E falsa).

1.3. Ancora sugli insiemi.

Vediamo ora alcune operazioni e alcuni concetti ulteriori sugli insiemi che possono essere utili per descrivere in modo sintetico vari argomenti che ci troveremo ad affrontare.

Definizione 0.4. Dati due insiemi A e B (non necessariamente distinti), chiameremo *prodotto cartesiano* $A \times B$ l'insieme delle coppie ordinate (a, b) ottenibili prendendo come primo elemento della coppia un elemento di A e come secondo elemento un elemento di B:

$$A \times B := \{(a, b) \mid a \in A,\, b \in B\}$$

(Diamo per assodato che sia chiaro il concetto di *coppia ordinata* (a, b). È importante ricordare che la coppia (a, b) e la coppia (b, a) **NON** sono coincidenti (ossia, importa anche l'ordine con cui diamo gli elementi a e b).

Definizione 0.5. Se gli insiemi che consideriamo sono i sottoinsiemi di un certo ambiente o universo U, dato un insieme A diremo che il suo *complementare* in U è l'insieme degli elementi di U che non stanno in A:

$$\mathcal{C}_U(A) := \{x \in U \mid x \notin A\}$$

(quando sarà chiaro in quale ambiente ci si trova, ometteremo l'indice U nel simbolo, e scriveremo semplicemente $\mathcal{C}(A)$).

A volte occorre considerare quali elementi di un insieme A sono *correlati* agli elementi di un altro insieme B (ossia, si vuole identificare un particolare sottoinsieme di $A \times B$); scriveremo $a \,\mathcal{R}\, b$ per esprimere che esiste questa relazione fra $a \in A$ e $b \in B$ (per esempio, la relazione "è figlio di").

Osservazione 0.3. Osserviamo che la conoscenza della relazione $\mathcal{R}$ fra gli insiemi A e B equivale a selezionare un particolare sottoinsieme del prodotto cartesiano $A \times B$; in altre parole, possiamo indentificare $\mathcal{R}$ con un elemento delle parti di $A \times B$:

$$\mathcal{R} \in \mathcal{P}(A \times B).$$

Di particolare interesse risulta il caso delle relazioni in cui A e B coincidono, e, ulteriormente, vi sono fra esse alcune che godono di proprietà peculiari: sono le relazioni di equivalenza.

Definizione 0.6. Una relazione fra gli elementi di un insieme A viene detta **di equivalenza** se valgono per essa le seguenti proprietà, valide per ogni scelta di elementi in A:

(1) proprietà riflessiva: $a \,\mathcal{R}\, a$;
(2) proprietà simmetrica: $x \,\mathcal{R}\, y \implies y \,\mathcal{R}\, x$;
(3) proprietà transitiva: $x \,\mathcal{R}\, y,\ y \,\mathcal{R}\, z \implies x \,\mathcal{R}\, z$.

In altre parole le proprietà dicono che ogni elemento è in relazione con se stesso; che quando x è in relazione con y, y è nella **stessa** relazione con x; che se possiamo dire che x è in relazione con y e che y è in relazione con z, allora x è in quella relazione con z.

L'uguaglianza fra numeri è un esempio banale di relazione di equivalenza. In geometria euclidea, la relazione di equivalenza per antonomasia fra figure piane è "ha la stessa area di": infatti, due figure che hanno la stessa area si dicono *equivalenti*. Per indicare che due elementi $a, a' \in A$ sono equivalenti, scriveremo

$$a \simeq a'\,.$$

Una volta che in un insieme viene introdotta una relazione di equivalenza $\simeq$, possiamo radunare in un sottoinsieme di A tutti gli elementi che sono equivalenti ad uno scelto:

Definizione 0.7. dato A ed una relazione di equivalenza $\mathcal{R}(\simeq)$ in esso, scelto un elemento $a \in A$, l'insieme di tutti gli elementi di A che sono equivalenti ad esso viene detto la sua classe di equivalenza $[a]$, modulo $\mathcal{R}$:

$$[a] := \{x \in A \mid x \simeq a\}\,.$$

Osservazione 0.4. Lasciamo allo studente il compito di convincersi delle seguenti proposizioni:

(1) per ogni $a \in A$, la classe $[a]$ è non vuota: infatti, contiene almeno a, visto che $a \simeq a$;
(2) Dati $a, b \in A$ le due classi $[a]$ e $[b]$ hanno intersezione non vuota se e solo se $a \simeq b$ e questo avviene se solo se $[a] = [b]$;
(3) l'unione di tutte le possibili classi di equivalenza coincide con l'insieme di partenza:
$$\bigcup_{a \in A} ([a]) = A.$$
In altre parole, tutto questo si esprime sinteticamente dicendo che le classi di equivalenza formano una *partizione* di A.

Esempio 0.5. Consideriamo l'insieme dei numeri naturali $\mathbb{N}$ e la relazione di equivalenza
$$\mathcal{R} = \text{“ha lo stesso resto nella divisione per 5 di”}.$$
Per esempio, con questa relazione $7 \simeq 12$, perché dividendo sia 7 che 12 per 5 si ottiene come resto 2 (formalmente, si direbbe che $7 \simeq 12 \pmod 5$, letto “7 è equivalente a 12 modulo 5”).

Cominciamo a chiederci quali numeri sono nella classe di equivalenza di 0: essi sono tutti i multipli di 5, che danno 0, appunto, come resto della divisione per 5:
$$[0] = \{0, 5, 10, 15, 20, 25, \dots\};$$
continuando, ci rendiamo conto che nella classe di equivalenza di 1 stanno tutti i primi successivi dei multipli di 5:
$$[1] = \{1, 6, 11, 16, 21, 26, \dots\};$$
e così via con le classi di equivalenza di $2, 3, 4$:
$$[2] = \{2, 7, 12, 17, 22, 27, \dots\};$$
$$[3] = \{3, 8, 13, 18, 23, 28, \dots\};$$
$$[4] = \{4, 9, 14, 19, 24, 29, \dots\};$$
e abbiamo finito: poiché $0 \simeq 5$ la classe $[5]$ coincide con $[0]$, la classe $[6]$ coincide con $[1]$, ecc.

Osserviamo che ogni classe è disgiunta dalle altre, e che riunendo le classi riotteniamo $\mathbb{N}$: abbiamo partizionato $\mathbb{N}$ in 5 sottoinsiemi. Naturalmente, mediante altre classi di equivalenza, avrei potuto partizionare l'insieme diversamente (sia per numero di sottoinsiemi, che per gli elementi che si trovano in ciascun sottoinsieme).

1.4. Funzioni.

Il concetto di *funzione* o *applicazione* fra due insiemi è tra quelli chiave nella matematica moderna; vale quindi la pena di dedicare un paragrafo a ripassare le definizioni più comuni in questo campo. Sinteticamente, parleremo di funzione quando esistono due insiemi fatti in modo che esista una "macchinetta" di qualunque tipo (una legge matematica, una tabella ecc.) che consente di prendere *un qualunque* elemento del primo insieme e trovare, grazie alla "macchinetta", *un solo* elemento nel secondo insieme che gli corrisponde. Più formalmente abbiamo la seguente

Definizione 0.8. Dati un insieme A, detto *dominio* ed un insieme B, detto *codominio*, viene detta funzione (o applicazione) f fra i due insiemi una legge che permette di associare ad ogni elemento a di A *uno ed un solo* elemento b di B. In simboli, scriveremo

$$f\colon A \to B$$

che leggeremo "f che va da A in B", o anche f che porta A in B". Fissato a, l'elemento b che gli corrisponde verrà detto *immagine di a secondo (o attraverso, o mediante) la f*, e per esprimere questo scriveremo $b = f(a)$, o anche $a \mapsto b$ o $a \overset{f}{\mapsto} b$ (Figura 0.3).

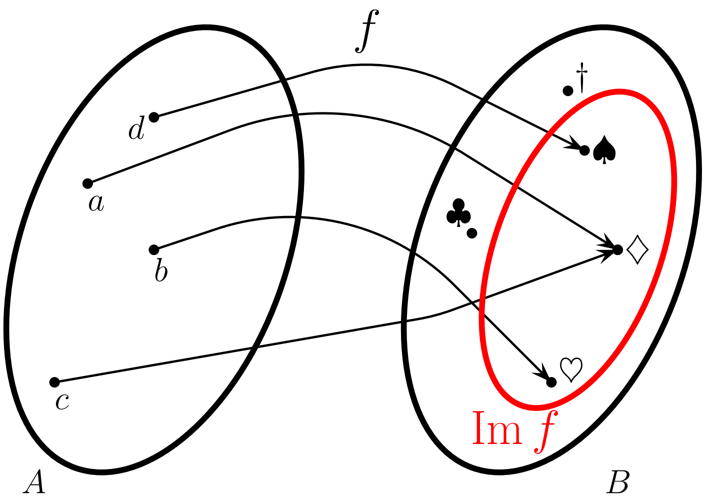

Figura 0.3. Rappresentazione schematica mediante i diagrammi di Eulero-Venn di una funzione $f\colon A \to B$. Si noti che, fissato $x \in A$, l'immagine $f(x)$ è unica: non possono corrispondere ad x due o più punti nel codominio B; nulla vieta, però, che $y \in B$ sia immagine di diversi elementi in A. L'ovale rosso racchiude gli elementi di $\operatorname{Im} f$, ossia l'insieme dei "punti di atterraggio" delle frecce.

Esempio 0.6. Vediamo alcuni esempi di funzione:

(1) Sia A l'insieme dei cittadini italiani iscritti al primo anno di Ingegneria di una certa sede universitaria (per esempio, Pavia o Mantova); sia B l'insieme di tutti i cognomi dei cittadini italiani. Consideriamo la funzione f che associa ad ogni elemento di A (ossia, ad ogni studente del primo anno cittadino italiano) il suo cognome.
(2) Sia A l'insieme dei comuni italiani, sia B l'insieme delle regioni italiane; consideriamo la f che associa ad ogni comune la regione di appartenenza.
(3) Sia A l'insieme dei numeri naturali, e sia B l'insieme dei numeri naturali. Consideriamo la funzione f che associa ad ogni numero naturale n il suo quadrato (ossia, $f(n) = n^2$).
(4) Sia A l'insieme dei numeri reali, e sia B sempre l'insieme dei numeri reali. Consideriamo la funzione f che associa ad ogni numero reale x il suo quadrato (ossia, $f(x) = x^2$). Anche se la legge che permette di trovare l'immagine è la stessa di prima, le due funzioni sono differenti, perché hanno *domini differenti*, e non basterebbe cambiare solo il codominio.
(5) Sia A l'insieme delle coppie (p, q) di numeri razionali, sia B l'insieme dei numeri reali. Consideriamo la f che associa ad ogni coppia (p, q) la loro somma: $f(p, q) = p + q$.

OSSERVAZIONE 0.7. Possiamo interpretare la conoscenza della funzione

$$f \colon A \to B$$

come la selezione di un particolare elemento di $\mathcal{P}(A \times B)$.

Risulta interessante qualificare tutti gli elementi del codominio che sono "raggiunti da almeno una freccia".

DEFINIZIONE 0.9. Sia $f \colon A \to B$ una funzione. L'insieme di tutte le possibili immagini ottenibili tramite la f viene chiamata **immagine** $\operatorname{Im} f$ della funzione:

$$\operatorname{Im} f := \{y \in B \mid y = f(x), \text{ al variare di } x \in A\} = \{f(x), x \in A\}\,.$$

In altre parole, possiamo dire che l'immagine di f è l'insieme di tutti gli elementi del codominio **per i quali esiste almeno un elemento nel dominio che viene trasformato in essi**:

$$\operatorname{Im} f := \{y \in B \mid \exists\, x \in A \colon y = f(x)\}\,.$$

Inoltre, preso un sottoinsieme D del dominio ($D \subset A$), chiamiamo **immagine di D secondo la f** l'insieme di tutte le immagini degli elementi di D, e la indicheremo con $f(D)$:

$$f(D) := \{y \in B \mid y = f(x), \text{ al variare di } x \in D\}\,. \tag{0.3}$$

Risulta evidente che $\operatorname{Im} f = f(A)$. Inoltre, per convenzione $f(\varnothing) = \varnothing$.

OSSERVAZIONE 0.8. In altre parole, per verificare che un certo elemento y del codominio appartenga all'immagine di f, occorre (e basta) trovare un elemento x del dominio che ha come immagine y: $y = f(x)$.

Ha anche interesse individuare, fissato un elemento del codominio, quali elemento del dominio "portano" in esso:

Definizione 0.10. Sia $f\colon A \to B$ una funzione. Fissato $b \in B$, l'insieme di tutti gli elementi di A che hanno per immagine b viene viene chiamato **controimmagine** $f^{-1}(b)$:

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

Inoltre, preso un sottoinsieme C del codominio ($C \subset B$), chiamiamo **controimmagine di C secondo la f** l'insieme di tutte le controimmagini degli elementi di C, e la indicheremo con $f^{-1}(C)$:

$$f^{-1}(C) := \{x \in A \mid f(x) = y, \text{ al variare di } y \in C\}\,.$$

Inoltre, per convenzione $f^{-1}(\varnothing) = \varnothing$.

Esempio 0.9. Considerando la f rappresentata schematicamente nella Figura 0.3, abbiamo:

$$\begin{aligned} f^{-1}(\{\heartsuit\}) &= \{b\}\,, \\ f^{-1}(\{\diamondsuit\}) &= \{a, c\}\,, \\ f^{-1}(\{\clubsuit\}) &= \varnothing. \end{aligned}$$

Osservazione 0.10. Data $f\colon A \to B$, la controimmagine del codominio è tutto il dominio:

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

Lo studente avrà sicuramente maggiore dimestichezza con le funzioni reali di una variabile reale, per le quali spesso viene usata una rappresentazione grafica nel piano cartesiano (formalizzeremo meglio il concetto di riferimento cartesiano ortogonale nel Capitolo 1, ma consideriamo che la sua conoscenza a livello di base faccia parte del bagaglio culturale di ogni studente proveniente dalle scuole superiori). In effetti, il concetto di "grafico di una funzione" vale in generale:

Definizione 0.11. Sia $f\colon A \to B$ una funzione; il **grafico** $\mathcal{G}_f$ della funzione è il sottoinsieme del prodotto cartesiano $A \times B$ di tutte le coppie (a, b) formate al variare di a in A, con b immagine di a secondo f:

$$\mathcal{G}_f := \{(a, b) \in A \times B \mid b = f(a), \quad \forall\, a \in A\}$$

Se ritorniamo all'Osservazione 0.7, possiamo dire che, in ultima analisi *la funzione è il grafico.*

Osservazione 0.11. In base alla definizione di funzione $f\colon A \to B$, vi sono alcune condizioni che si devono sempre verificare: cerchiamo di vederle esplicitamente, dando anche l'interpretazione di queste in termini di diagrammi schematici di Eulero-Venn

(1) Devo sempre potere prendere un elemento nel dominio e trovare un'immagine nel codominio. In altre parole, da ogni elemento del dominio parte **sempre** una freccia.
(2) Fissato un elemento del dominio, a questo corrisponde una sola immagine. Quindi da ogni elemento del dominio parte **una sola** freccia.

(3) Non è detto che ad ogni immagine corrisponda solo un elemento del dominio: quindi per un elemento del dominio possono esservi **una, nessuna o più di una** freccia che vi arriva.

Quest'ultimo punto ci fa capire che diventa utile classificare ulteriormente le funzioni in base al loro comportamento; diamo quindi le seguenti definizioni.

DEFINIZIONE 0.12. Sia $f\colon A \to B$ una funzione:

(1) f viene detta **surgettiva (o suriettiva)** se la sua immagine coincide con il codominio, ossia: $B = \operatorname{Im} f$. In altre parole, per una funzione surgettiva, preso un qualunque elemento y di B, questo è immagine di *almeno* un elemento x di A:
$$\forall\, y \in B \quad \exists\, x \in A \,|\, y = f(x)\,,$$
dunque, in ogni elemento di B arriva **almeno una freccia** (Figura 0.4);

(2) f viene detta **iniettiva** se porta *elementi distinti in elementi distinti*, ossia
$$\forall\, x_1,\, x_2 \in A \quad x_1 \neq x_2 \implies f(x_1) \neq f(x_2)\,,$$
o anche, nella forma contronominale, se le immagini di due elementi coincidono, i due elementi devono coincidere
$$\forall\, x_1,\, x_2 \in A \quad f(x_1) = f(x_2) \implies x_1 = x_2\,.$$
Schematicamente, per una funzione iniettiva in ogni elemento del codominio può arrivare **solo una freccia, o nessuna** (Figura 0.5);

(3) f viene detta **bigettiva** (o biiettiva), o anche f **è una corrispondenza biunivoca** se è contemporaneamente surgettiva e iniettiva. Schematicamente, tutti i punti del codominio saranno raggiunti da **una sola freccia** (Figura 0.6).

ESEMPIO 0.12. Per aiutare lo studente a comprendere meglio ed assimilare le definizioni 0.12, riportiamo alcuni semplici esempi di funzioni reali di variabile reale $f\colon \mathbb{R} \to \mathbb{R}$.

- Le funzioni $f(x) = x$, $f(x) = 2x$ e $f(x) = x^3$ sono iniettive. Verifichiamo ad esempio che $f(x) = x^3$ è iniettiva. Siano $x_1, x_2 \in \mathbb{R}$ numeri reali, proviamo che:
$$f(x_1) = f(x_2) \quad \implies \quad x_1 = x_2.$$
Infatti è ben noto che l'uguaglianza $(x_1)^3 = (x_2)^3$ implica che $x_1 = x_2$, da cui ricaviamo l'iniettività di f.
- Le funzioni $f(x) = x^2$ e $f(x) = \sin x$ non sono iniettive. Verifichiamo che $f(x) = x^2$ non è iniettiva. Basta provare l'esistenza di due numeri reali x_1 e x_2 tali che:
$$x_1 \neq x_2 \quad \text{e} \;\; f(x_1) = f(x_2).$$
Osserviamo che se x_1 e x_2 sono due numeri reali non nulli opposti allora si ha $(x_1)^2 = (x_2)^2$, per cui possiamo concludere che f non è iniettiva.

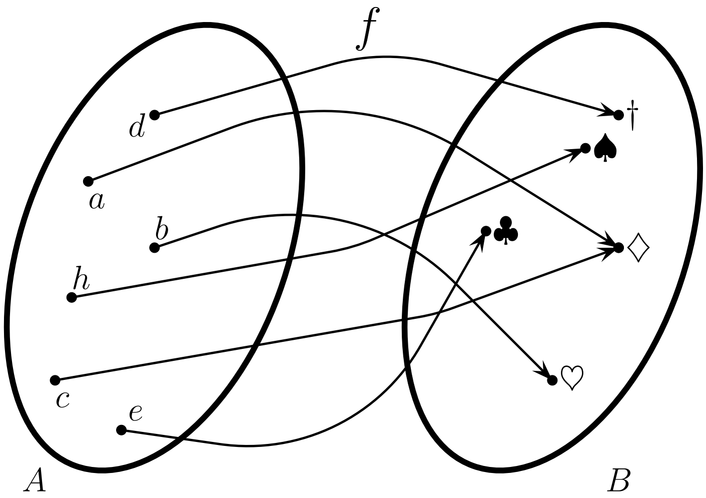

Figura 0.4. Rappresentazione schematica mediante i diagrammi di Eulero-Venn di una funzione $f\colon A \to B$ surgettiva. Tutti gli elementi del codominio B sono raggiunti da almeno una freccia.

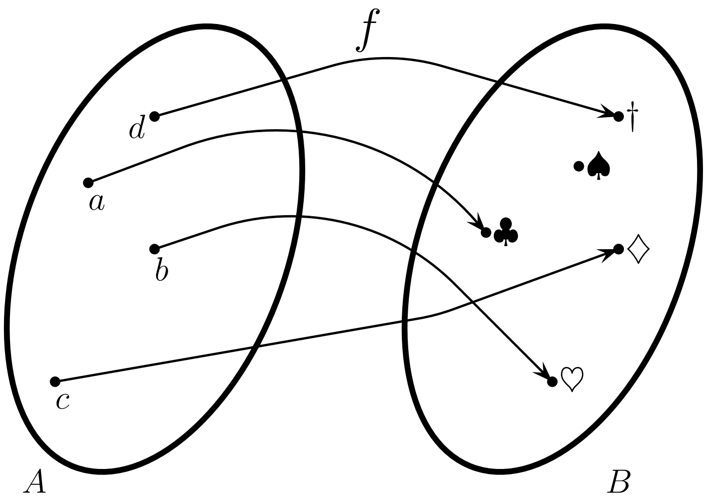

Figura 0.5. Rappresentazione schematica mediante i diagrammi di Eulero-Venn di una funzione $f\colon A \to B$ iniettiva. Tutti gli elementi del codominio B sono raggiunti al più da una freccia.

- Le funzioni $f(x) = 2x$ e $f(x) = x^3$ sono suriettive. Verifichiamo ad esempio che $f(x) = x^3$ è suriettiva. Proviamo che:

$$\forall y \in \mathbb{R} \quad \exists x \in \mathbb{R} \ \ | f(x) = y.$$

  Osserviamo che $\forall y \in \mathbb{R}$ l'equazione $x^3 = y$ ammette la soluzione $y = \sqrt[3]{x}$, da cui ricaviamo che f è suriettiva.
- Le funzioni $f(x) = x^2$ e $f(x) = \sin x$ non sono suriettive. Verifichiamo che $f(x) = x^2$ non è suriettiva. Basta provare l'esistenza di un numero

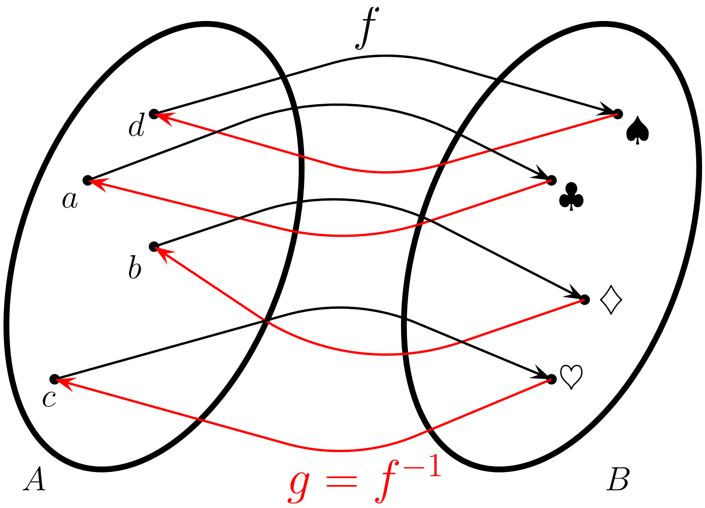

FIGURA 0.6. Rappresentazione schematica mediante i diagrammi di Eulero-Venn di una funzione $f\colon A \to B$ bigettiva. Tutti gli elementi del codominio B sono raggiunti da **una ed una sola** freccia. La funzione risulta invertibile: la funzione $g = f^{-1}\colon B \to A$ è definita in modo che $x = g(y) = f^{-1}(y)$ se $y = f(x)$.

reale $y \in \mathbb{R}$ per cui l'equazione $f(x) = y$ non ammetta soluzioni reali. Osserviamo che se $y < 0$ l'equazione $x^2 = y$ non amette soluzioni reali, per cui possiamo concludere che f non è suriettiva.

OSSERVAZIONE 0.13. Cambiando il codominio, è sempre possibile rendere surgettiva una funzione $f\colon A \to B$: basta usare come codominio l'immagine di f. La funzione $g\colon A \to \operatorname{Im} f$, definita in modo che $g(x) = f(x)$ per tutti gli x di A è surgettiva.

L'iniettività è in generale più problematica, ma sotto certe condizioni è possibile arrivare ad avere una funzione iniettiva da una generica. Per fare questo, può essere necessario *restringere* l'ambito della f, ossia il suo dominio.

DEFINIZIONE 0.13. Sia $f\colon A \to B$ una funzione, e sia $D \subset A$ un sottinsieme del dominio. Chiameremo **restrizione della f a D** o anche **funzione ristretta a D** la funzione

$$f_D\colon D \to B\,,$$

definita in modo che per ogni $x \in D$ si ha

$$f_D(X) = f(x)\,.$$

In altre parole, la funzione ristretta è definita solo sugli elementi del sottoinsieme D, ed ha gli stessi valori che aveva la f originale.

Nelle funzioni biunivoche ogni elemento del dominio ha la controimmagine non vuota, e questa contiene un solo elemento (controimmagine unica); in tal caso, possiamo "invertire" le frecce e definire una funzione in cui dominio e codominio si scambiano i ruoli (Figura 0.6).

Definizione 0.14. Sia $f\colon A \to B$ una funzione bigettiva; allora è **invertibile**, e la **funzione inversa** $g = f^{-1}\colon B \to A$ è definita da $g(y) = f^{-1}(y) = x$ se $y = f(x)$.

1.5. Composizione di funzioni.

Per finire, ricordiamo un concetto che permette di definire funzioni "in cascata":

Definizione 0.15. Siano $f\colon X \to Y$ e $g\colon Y \to Z$ due funzioni, definite in modo che il codominio della prima coincida con il dominio della seconda. Viene chiamata **funzione composta** $f \circ g$ la funzione che ha come dominio X come codominio Z e definita in modo che l'immagine di ogni elemento $x \in X$ sia l'immagine secondo g dell'immagine di x secondo f, ossia:

$$f \circ g\colon X \to Z, \qquad x \mapsto g\left(f(x)\right) \,. \tag{0.4}$$

Si noti che l'ordine di composizione è essenziale: $f \circ g \neq g \circ f$ (sempre che le due composizioni siano possibili, per la scelta dei domini e dei codomini delle f, g).

2. Insiemi numerici e operazioni interne.

In ciascuno degli insiemi numerici visti precedentemente (1.1) sono definite alcune **operazioni interne**.

Definizione 0.16. Sia A un insieme; chiamiamo **operazione interna (o "binaria")** una funzione

$$f\colon A \times A \to A$$

che associa ad una coppia ordinata a, b di elementi di A un elemento di A stesso; normalmente, l'operazione binaria viene indicata da un simbolo interposto fra i due termini a e b, ad esempio:

$$(a, b) \mapsto a \star b$$

In altre parole, un'operazione su un insieme A è detta interna se può essere **sempre** eseguita per ogni coppia di numeri di A ed il risultato è un numero che appartiene **allo stesso insieme**.

Esempio 0.14. Vediamo come esempio le due operazioni interne "per eccellenza".
L'operazione di addizione fra numeri interi è un'operazione interna:

$$\begin{aligned} +\colon & \mathbb{N} \times \mathbb{N} \to \mathbb{N} \\ & (n_1, n_2) \mapsto n_1 + n_2 \qquad n_1, n_2 \in \mathbb{N} \end{aligned}$$

L'operazione di moltiplicazione fra numeri interi è un'operazione interna:

$$\begin{aligned} .: \mathbb{N} \times \mathbb{N} &\to \mathbb{N} \\ (n_1, n_2) &\mapsto n_1.n_2 \qquad n_1, n_2 \in \mathbb{N} \end{aligned}$$

In effetti, l'idea di "ingrandire" un certo insieme numerico nasce proprio dall'esigenza di poter introdurre nuove operazioni interne, altrimenti non eseguibili sempre:

- per poter definire come operazione interna la sottrazione, vengono creati i numeri interi (relativi) $\mathbb{Z}$, di cui i numeri naturali $\mathbb{N}$ sono un sottoinsieme (quello dei numeri positivi o nulli);
- per poter sempre eseguire le divisioni (tranne che per 0), vengono creati i numeri razionali $\mathbb{Q}$, che sono le frazioni (con denominatore non nullo) fra numeri interi (anzi, meglio, le classi di equivalenza delle frazioni);
- per poter sempre estrarre la radice quadrata di numeri positivi nell'insieme in cui si opera, si introducono i numeri reali $\mathbb{R}$, un sottoinsieme dei quali è in corrispondenza biunivoca con i numeri razionali, e, pertanto, si identifica con $\mathbb{Q}$ stesso.

Ogni volta che un insieme numerico viene "allargato", si definiscono nel nuovo insieme le operazioni interne di addizione e moltiplicazione, in modo che i risultati delle operazioni ristrette ad un sottoinsieme "precedente" coincidano con quelli ottenuti con la definizione "precedente".

Proprietà 0.3. *Ricordiamo alcune proprietà di queste operazioni.*

(1) Proprietà commutativa dell'addizione*:*

$$\forall a, b \quad a + b = b + a,$$

è verificata in tutti gli insiemi $\mathbb{N}$*,* $\mathbb{Z}$*,* $\mathbb{Q}$ $\mathbb{R}$*.*

(2) Proprietà associativa dell'addizione*:*

$$\forall a, b, c \quad (a + b) + c = a + (b + c),$$

è verificata in tutti gli insiemi $\mathbb{N}$*,* $\mathbb{Z}$*,* $\mathbb{Q}$ *e* $\mathbb{R}$*.*

(3) Esistenza dell'elemento neutro per la somma*:*

$$\exists a_0 | \ \forall a \quad a + a_0 = a_0 + a = a,$$

tale elemento è il numero 0 *e la proprietà è verificata in tutti gli insiemi* $\mathbb{N}$*,* $\mathbb{Z}$*,* $\mathbb{Q}$ *e* $\mathbb{R}$*.*

(4) Esistenza dell'elemento opposto per la somma*:*

$$\forall a \quad \exists a' \mid a + a' = a' + a = 0,$$

$a' = -a$ *è l'opposto di* a *e tale proprietà è verificata in* $\mathbb{Z}$*,* $\mathbb{Q}$ *e* $\mathbb{R}$ *(ma non in* $\mathbb{N}$*).*

(5) Proprietà distributiva della moltiplicazione rispetto all'addizione*:*

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

è verificata in tutti gli insiemi $\mathbb{N}$*,* $\mathbb{Z}$*,* $\mathbb{Q}$ *e* $\mathbb{R}$*.*

(6) Proprietà commutativa della moltiplicazione*:*

$$\forall a, b \quad a.b = b.a,$$

è verificata in tutti gli insiemi $\mathbb{N}$*,* $\mathbb{Z}$*,* $\mathbb{Q}$ *e* $\mathbb{R}$*.*

(7) Proprietà associativa della moltiplicazione*:*

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

è verificata in tutti gli insiemi $\mathbb{N}$*,* $\mathbb{Z}$*,* $\mathbb{Q}$ *e* $\mathbb{R}$*.*

(8) Esistenza dell'elemento neutro per il prodotto*:*

$$\exists u \mid \forall a \quad a.u = u.a = a,$$

tale elemento è il numero 1 *e la proprietà è verificata in tutti gli insiemi* $\mathbb{N}$*,* $\mathbb{Z}$*,* $\mathbb{Q}$ *e* $\mathbb{R}$*.*

(9) Esistenza dell'elemento inverso per il prodotto*:*

$$\forall a \neq 0 \quad \exists a'' \mid a.a'' = a''.a = 1,$$

$a'' = a^{-1}$ *è l'inverso di* a *e tale proprietà è verificata in* $\mathbb{Q}$ *e* $\mathbb{R}$*, privati dello* 0 *(ma non in* $\mathbb{N}$ *o* $\mathbb{Z}$*, neanche escludendo lo* 0*).*

Per risolvere il problema dell'estrazione di radici quadrate di numeri negativi, vengono creati i numeri complessi $\mathbb{C}$. In essi, viene introdotta l'unità immaginaria i, fatta in modo che

$$i^2 = -1\,.$$

Così facendo, un numero *complesso* z può essere composto di una *parte reale* $a = \Re(z)$ ed una *parte immaginaria* $b = \Im(z)$, "sommate":

$$z = a + i\,b\,.$$

Un modo molto efficace di rappresentare i numeri complessi è quello dovuto a Gauss ed Argand, che realizza una corrispondenza biunivoca fra z ed i punti di un piano cartesiano in cui si riporta la parte reale sull'asse delle ascisse e la parte immaginaria su quello delle ordinate (Figura 0.7).

I numeri che hanno parte immaginaria nulla si trovano in corrispondenza biunivoca con i numeri reali stessi, e vengono identificati con questi ultimi.

Riprenderemo nel dettaglio le considerazioni sui numeri complessi nell'Appendice (sezione 5), dove impareremo alcuni strumenti essenziali per lavorare con essi. Per il paragrafi rimanenti, comunque, serviranno solo alcune operazioni e proprietà di base:

(1) l'operazione di coniugazione (o coniugio), che associa ad ogni numero $z = a + i\,b \in \mathbb{C}$ il suo complesso coniugato $\overline{z}$ che ha uguale parte reale e parte immaginaria cambiata di segno

$$\overline{z} = a - i\,b\,,$$

(2) il risultato (Proposizione 0.11) –abbastanza evidente– che un numero complesso è reale se e solo se la sua parte immaginaria è nulla:

$$z \in \mathbb{R} \iff \Im(z) = 0\,.$$

La scrittura algebrica consente di trattare i numeri complessi seguendo le regole dell'algebra "solite"; ad esempio, l'addizione fra due numeri complessi si ottiene sommando separatamente le parti reali ed immaginarie:

$$z_1 = a_1 + i\,b_1, \qquad z_2 = a_2 + i\,b_2 \implies z_1 + z_2 = (a_1 + a_2) + i(b_1 + b_2)\,,$$

e anche, per il prodotto

$$\begin{aligned} z_1 . z_2 &= (a_1 + i\,b_1).(a_2 + i\,b_2) = a_1a_2 + ia_1b_2 + ib_1a_2 + i^2\,b_1b_2 \\ &= (a_1a_2 - b_1b_2) + i(a_1b_2 + a_2b_1)\,, \end{aligned}$$

ossia, trattando formalmente le scritture come se i termini con i fossero dei monomi di primo grado in questa variabile. Così operando, ci si convince che le operazioni di addizione moltiplicazione in $\mathbb{C}$ godono di tutte le proprietà di cui godono in $\mathbb{R}$.

3. Strutture algebriche.

Nella sezione precedente 2 abbiamo visto, sostanzialmente, che, una volta che abbiamo un insieme dotato di certe operazioni interne che obbediscono a determinate leggi, abbiamo un modo "standard" di procedere con esse, che possiamo "riportare" in altri insiemi, con operazioni interne "sovrapponibili" alle prime: formalizziamo ora questo concetto.

Definizione 0.17. Chiamiamo **struttura algebrica** un insieme A non vuoto, dotato di una o più operazioni interne

$$(A, \diamond, \star, \dots)$$

con

$$\begin{aligned} \diamond \colon A \times A \to A, &\qquad (a,b) \mapsto a \diamond b \in A, \\ \star \colon A \times A \to A, &\qquad (a,b) \mapsto a \star b \in A, \dots \end{aligned}$$

Esempio 0.15. Sono strutture algebriche, per esempio: $(\mathbb{N}, +)$, $(\mathbb{R}, +, .)$, $(\mathbb{C}, .)$ ecc.

Cerchiamo ora di classificare le strutture algebriche in base ad alcune proprietà; cominciamo con un concetto molto importante nell'algebra astratta, che useremo diffusamente anche noi.

Definizione 0.18. Una struttura algebrica $(G, \star)$ è un **gruppo** (o, anche, "G con l'operazione interna $\star$ ha la struttura algebrica di gruppo") se l'operazione interna $\star$ gode delle seguenti proprietà.

- $\mathcal{G}_1$ (*Proprietà associativa*)

$$(g_1 \star g_2) \star g_3 = g_1 \star (g_2 \star g_3) \qquad \forall\, g_1, g_2, g_3 \in G.$$

- $\mathcal{G}_2$ (*Esistenza dell'elemento neutro*)

$$\exists\, u \in G \,|\, g \star u = u \star g = g \qquad \forall\, g \in G.$$

- $\mathcal{G}_3$ (*Esistenza dell'elemento inverso (o opposto)*)
$$\forall\, g \in G,\ \exists\, g^{-1} \in G \,|\, g \star g^{-1} = g^{-1} \star g = u.$$

Inoltre, il gruppo viene detto **abeliano** o **commutativo** se vale anche un'altra proprietà.

- $\mathcal{G}_4$ (*Proprietà commutativa*)
$$g_1 \star g_2 = g_2 \star g_1 \qquad \forall\, g_1, g_2 \in G.$$

Esempio 0.16. Sono esempi di gruppi (abeliani) le strutture algebriche $(\mathbb{Z}, +)$, $(\mathbb{Q}, +)$, $(\mathbb{Q}^*, .)$.

NON sono invece un gruppo le strutture $(\mathbb{N}, +)$ (manca in generale l'opposto in $\mathbb{N}$ per la somma: servono i numeri interi relativi: $5 + (-5) = 0$), $(\mathbb{Z}^*, .)$ (manca in generale l'inverso in $\mathbb{Z}$: servono le frazioni: $3.\frac{1}{3} + 1$).

Cerchiamo ora di capire come "combinare" due operazioni interne in una struttura algebrica.

Definizione 0.19. Chiamiamo **anello** una struttura algebrica $(A, \oplus, \star)$ dotata di due operazioni interne (che chiameremo "addizione", e "moltiplicazione") che è un *gruppo abeliano* additivo (ossia, rispetto alla prima operazione $+$) e con la seconda operazione che gode delle seguenti proprietà.

- $\mathcal{A}_1$ (*Proprietà associativa*)
$$(a \star b) \star c = a \star (b \star c) \qquad \forall\, a, b, c \in A.$$
- $\mathcal{A}_2$ (*Proprietà distributiva della somma sul prodotto*)
$$a \star (b \oplus c) = (a \star b) \oplus (a \star c) \qquad \forall\, a, b, c \in A.$$

Inoltre: l'anello è detto **commutativo** se vale anche la seguente proprietà.

- $\mathcal{A}_3$ (*Proprietà commutativa del prodotto*)
$$a \star b = b \star a \qquad \forall\, a, b \in A;$$

l'anello è detto **unitario** se vale

- $\mathcal{A}_4$ (*Esistenza dell'elemento neutro per il prodotto*)
$$\exists\, u\, (\neq 0) \,|\, u \star a = a \star u = a \qquad \forall\, a \in A.$$
(dove 0 è l'elemento neutro di $\oplus$, che esiste, visto che $(A, \oplus)$ è un gruppo).

Infine, l'anello è detto **integro** se non esistono divisori dello 0, ossia se

- $\mathcal{A}_5$ (*Legge di annullamento del prodotto*)
$$a \star b = 0 \implies a = 0 \vee b = 0\,;$$

Esempio 0.17. Le strutture $(\mathbb{Z}, +, .)$, $(\mathbb{Q}, +, .)$, $(\mathbb{R}, +, .)$, $(\mathbb{C}, +, .)$, con le ordinarie operazioni di addizione e moltiplicazione sono esempi di anello integro unitario commutativo (l'elemento neutro della moltiplicazione è il numero 1)

Ad un anello "manca poco" per essere una struttura algebrica molto importante, perché sarà quella che servirà in molti contesti del nostro corso:

DEFINIZIONE 0.20. Un anello commutativo unitario $(\mathbb{K}, +, .)$ viene detto **campo** se ogni elemento di $\mathbb{K}$, escluso lo 0 ha l'elemento inverso per l'operazione di moltiplicazione

PROPOSIZIONE 0.4. *La struttura algebrica di campo equivale ad un insieme $\mathbb{K}$ dotato di due operazioni interne di addizione (+) e moltiplicazione (.), che godono delle seguenti proprietà:*

$\mathcal{S}_1$) *(Commutativa di +)*

$$a + b = b + a \qquad \forall\, a, b \in \mathbb{K}.$$

$\mathcal{S}_2$) *(Associativa di +)*

$$(a + b) + c = a + (b + c) \qquad \forall\, a, b, c \in \mathbb{K}.$$

$\mathcal{S}_3$) *(Esistenza del neutro per +)*

$$\exists 0 \in \mathbb{K} \,|\, a + 0 = 0 + a \qquad \forall\, a \in \mathbb{K}.$$

$\mathcal{S}_4$) *(Esistenza dell'opposto per +)*

$$\forall\, a \in \mathbb{K}, \exists b \in \mathbb{K} \,|\, a + b = 0 \qquad (b = -a).$$

$\mathcal{P}_1$) *(Commutativa di +)*

$$a.b = b.a \qquad \forall\, a, b \in \mathbb{K}.$$

$\mathcal{P}_2$) *(Associativa di .)*

$$(a.b).c = a.(b.c) \qquad \forall\, a, b, c \in \mathbb{K}.$$

$\mathcal{P}_3$) *(Esistenza del neutro per .)*

$$\exists 1 \neq 0 \in \mathbb{K} \,|\, a.1 = 1.a = a \qquad \forall\, a \in \mathbb{K}.$$

$\mathcal{P}_4$) *(Esistenza dell'inverso per .)*

$$\forall\, a \neq 0 \in \mathbb{K}, \exists b \in \mathbb{K} \,|\, a.b = 1 \qquad (b = a^{-1} = \frac{1}{a}).$$

$\mathcal{PS}$) *(Proprietà distributiva di . rispetto a +)*

$$a.(b + c) = (a.b) + (a.c) \qquad \forall a, b, c \in \mathbb{K}.$$

La verifica dell'equivalenza è lasciata allo studente come esercizio.

ESEMPIO 0.18. Sono campi:

- $(\mathbb{Q}, +, .)$;
- $(\mathbb{R}, +, .)$;
- $(\mathbb{C}, +, .)$.

Infatti, per tutte queste strutture algebriche valgono le proprietà 0.3, che sono proprio quelle richieste dalla Proposizione 0.4.

Invece, $(\mathbb{Z}, +, .)$ non è un campo: non sempre c'è l'inverso di un numero intero **all'interno** dell'insieme dei numeri interi (anzi, in $\mathbb{Z}$ esiste solo per i numeri ± 1).

Resta da definire cosa intendiamo quando vogliamo dire che due insiemi, con le operazioni che godono delle proprietà viste sopra, possono essere "sovrapposti".

Definizione 0.21. Siano $(\mathbb{K}, +, .)$ e $(\mathbb{K}', \oplus, \star)$ due campi; diciamo che esiste un **isomorfismo** fra campi se esiste un'applicazione $f\colon \mathbb{K} \to \mathbb{K}'$ tale che, per ogni scelta di $a, b \in \mathbb{K}$:

$$\begin{aligned} f(a+b) &= f(a) \oplus f(b) \\ f(a.b) &= f(a) \star f(b). \end{aligned}$$

I campi $\mathbb{K}$ e $\mathbb{K}'$ sono detti **isomorfi**.

Sulla scorta di questo diremo, propriamente che

- $\mathbb{R}$ contiene un insieme isomorfo a $\mathbb{Q}$.
- $\mathbb{C}$ contiene in insieme isomorfo a $\mathbb{R}$.

Ecco, quindi, in quale senso diciamo che l'insieme dei numeri complessi che hanno parte immginaria nulla *si identifica* con i numeri reali.

4. Anello dei polinomi.

Un capitolo fondamentale dell'Algebra è lo studio dei polinomi e delle equazioni algebriche. Ricordiamo innanzittutto la definizione di polinomio a coefficienti reali in una indeterminata x:

Definizione 0.22. Un **polinomio** $p(x)$ a coefficienti reali in x è un'espressione algebrica costituita da una somma finita di monomi che contengono solo potenze (con esponenente ≥ 0) di x, cioè

$$p(x) = a_0 + a_1 x + a_2 x^2 + ... + a_n x^n, \quad a_i \in \mathbb{R},\ i = 0, .., n\,.$$

Il **grado** di $p(x)$ viene indicato con $\deg p(x)$ ed è il massimo esponente con cui compare l'indeterminata x, equivalentemente:

$$\deg p(x) = \max\{i \mid a_i \neq 0\}.$$

Sono esempi di polinomi in x:

$$p_1(x) = 2 - x^3 + 3x^4, \quad p_2(x) = 2x - 5, \quad p_3(x) = 3;$$

per i quali abbiamo:

$$\deg p_1(x) = 4 \quad \deg p_2(x) = 1 \quad \deg p_3(x) = 0.$$

Introduciamo una forma compatta per scrivere espressioni come quella che abbiamo usato per un polinomio generico nella definizione 0.22. Notiamo, infatti, che si devono sommare molti termini ($n+1$, in questo caso) legati direttamente (tramite indice o esponente i, per esempio) ad un numero naturale.

Definizione 0.23. Chiamiamo **sommatoria** una somma di termini che dipendono da un numero naturale i (detto *indice di sommatoria*) che può assumere tutti i valori a partire da un numero iniziale n_A fino ad un numero finale n_B (detti *estremi di sommatoria*). Usiamo per questo un simbolo formato da una lettera greca Σ, con indicazione in calce dell'indice e dell'estremo inferiore di sommatoria, ed in testa dell'estremo superiore di sommatoria. Ad esempio

$$\sum_{1\ i}^{6} c_i = \sum_{i=1}^{6} c_i = c_1 + c_2 + c_3 + c_4 + c_5 + c_6 \,. \tag{0.5}$$

Osservazione 0.19. Notiamo che l'indice di sommatoria è *muto*, ossia

$$\sum_{i=1}^{6} c_i = \sum_{k=1}^{6} c_k \,;$$

inoltre, oltre alla scrittura base, esistono numerose varianti del simbolo di sommatoria (ad esempio, si possono sommare solo i termini di indice pari o dispari, o altre condizioni ancora). Non svolgeremo qui una trattazione sistematica (e anche, forse, un po' noiosa) di tutti i numerosi casi riscontrabili, spesso chiari una volta compreso il meccanismo del simbolo stesso, e rimandiamo l'eventuale spiegazione dettagliata della variante del simbolo al momento in cui la si vorrà utilizzare.

Definizione 0.24. Usando il simbolo di sommatoria, un polinomio generico di grado n nella variabile x si potrà scrivere come

$$p(x) = \sum_{i=0}^{n} a_i \, x^i \,,$$

dove converremo che, per $i = 0$, diamo alla scrittura x^i il significato equivalente di 1 (anche per $x = 0$, anche se al simbolo 0^0 non attribuiamo normalmente un significato).

Al variare dei coefficienti in $\mathbb{R}$ i polinomi formano un insieme che possiamo dotare di una struttura algebrica.

Definizione 0.25. Indichiamo con $\mathbb{R}[x]$ l'insieme di tutti i polinomi a coefficienti reali in x. Inoltre, indichiamo con $\mathbb{R}_n[x]$ l'insieme dei polinomi a coefficienti reali *di grado minore o uguale a n*. Nell'insieme $\mathbb{R}[x]$ (e, analogamente, in $\mathbb{R}_n[x]$, per la prima) sono definite due operazioni interne.

(1) La prima operazione è la **somma di polinomi**: la somma di due polinomi $p(x)$ e $q(x)$ è il polinomio $p(x)+q(x)$ che si ottiene sommando i monomi simili tra loro.
Ad esempio, se $p(x) = 1 - x + x^3$ e $q(x) = 3 + 4x - x^2 - x^3 + 6x^4$, abbiamo:

$$\begin{aligned} p(x) + q(x) &= (1+3) + (-1+4)x + (0-1)x^2 + (1-1)x^3 + 6x^4 \\ &= 4 + 3x - x^2 + 6x^4. \end{aligned}$$

Vogliamo stabilire una relazione tra il grado del polinomio somma e i gradi dei singoli addendi.
Siano $p(x), q(x) \in \mathbb{R}[x]$ con $\deg p(x) \neq \deg q(x)$, allora si ha:

$$\deg(p(x) + q(x)) = \max(\deg p(x), \deg q(x));$$

se invece $\deg p(x) = \deg q(x) = d$, può succedere che $\deg(p(x)+q(x)) < d$.

Ad esempio, se $p(x) = 2-x+x^5$ e $q(x) = 3-x^5$, abbiamo $p(x)+q(x) = 5-x$, quindi $\deg(p(x)+q(x)) = 1 < 5$.

(2) La seconda operazione in $\mathbb{R}[x]$ è il **prodotto di polinomi**: il prodotto di due polinomi $p(x)$ e $q(x)$ è il polinomio $p(x).q(x)$ che si ottiene moltiplicando ciascun monomio di $p(x)$ per ciascun monomio di $q(x)$ e sommando poi tutti i prodotti ottenuti.

Ad esempio, se $p(x) = 3-2x+x^3$ e $q(x) = 1+4x^2$, abbiamo:

$$p(x).q(x) = (3-2x+x^3).(1+4x^2) =$$

$$= 3.1 + 3.4x^2 + (-2x).(1) + (-2x).4x^2 + x^3.1 + x^3.4x^2 =$$

$$= 3 + 12x^2 - 2x - 8x^3 + x^3 + 4x^5 = 3 - 2x + 12x^2 - 7x^3 + 4x^5.$$

Osserviamo che il grado del polinomio prodotto di $p(x)$ e $q(x)$ soddisfa la relazione

$$\deg(p(x).q(x)) = \deg p(x) + \deg q(x)\,,$$

è quindi la somma dei gradi dei fattori.

Esercizio 4.1*:* Consideriamo due polinomi qualunque

$$q(x) = \sum_{i=0}^{n} a_i\, x^i$$

e

$$r(x) = \sum_{i=0}^{m} a_i\, x^i\,.$$

Mostrare che il polinomio prodotto $p(x) = q(x).r(x)$ si può scrivere nella forma

$$p(x) = \sum_{i=0}^{m+n} c_i\, x^i\,,$$

in cui i coefficienti c_i sono dati da

$$c_i = \sum_{r+s=i} a_r b_s\,,$$

(l'ultimo simbolo va inteso come la sommatoria di tutti i termini in cui la somma degli indici r ed s è pari a i).

Proprietà 0.5. *Le operazioni introdotte in* $\mathbb{R}[x]$ *soddisfano le seguenti proprietà:*

(1) Proprietà commutativa dell'addizione*:*

$$\forall p(x), q(x) \in \mathbb{R}[x] \quad p(x) + q(x) = q(x) + p(x);$$

(2) Proprietà associativa dell'addizione*:*

$$\forall p(x), q(x), r(x) \in \mathbb{R}[x] \quad (p(x) + q(x)) + r(x) = p(x) + (q(x) + r(x));$$

(3) Esistenza dell'elemento neutro per la somma*:*

$$\forall p(x) \in \mathbb{R}[x] \quad p(x) + 0 = 0 + p(x) = p(x),$$

tale elemento è il polinomio di grado 0 *che è identicamente* 0*.*

(4) Esistenza dell'elemento opposto per la somma*:*

$$\forall p(x) \in \mathbb{R}[x] \quad p(x) + (-p(x)) = (-p(x)) + p(x) = 0,$$

$-p(x)$ è il polinomio opposto di $p(x)$ che si ottiene semplicemente cambiando segno a tutti i coefficienti di $p(x)$;

(5) Proprietà distributiva della moltiplicazione rispetto alla somma*:* $\forall p(x), q(x), r(x) \in \mathbb{R}[x]$

$$(p(x) + q(x)).r(x) = p(x).r(x) + q(x).r(x),$$

$$r(x).(p(x) + q(x)) = r(x).p(x) + r(x).q(x);$$

(6) Proprietà commutativa della moltiplicazione*:*

$$\forall p(x), q(x) \in \mathbb{R}[x] \quad p(x).q(x) = q(x).p(x);$$

(7) Proprietà associativa della moltiplicazione*:*

$$\forall p(x), q(x), r(x) \in \mathbb{R}[x] \quad (p(x).q(x)).r(x) = p(x).(q(x).r(x));$$

(8) Esistenza dell'elemento neutro per il prodotto*:*

$$\forall p(x) \in \mathbb{R}[x] \quad p(x).1 = 1.p(x) = p(x),$$

tale elemento è il polinomio di grado 0 *che coincide col numero* 1*.*

Osservazione 0.20. Osserviamo che nell'insieme $\mathbb{R}[x]$ dei polinomi a coefficienti reali in x valgono le stesse proprietà (1)-(8) elencate nella Proprietà 0.3 per l'addizione e la moltiplicazione di numeri interi. Per questo motivo, possiamo dire che $\mathbb{R}[x]$ è l'**anello dei polinomi** a coefficienti reali nell'indeterminata x (Definizione 0.19).

Osserviamo anche che $\mathbb{R}[x]$ non è un campo, infatti la proprietà 9 dell'elenco 0.3 (esistenza dell'inverso) è verificata solo per i polinomi di grado 0, diversi da 0.

Infine, limitandosi all'insieme $\mathbb{R}_n[x]$, notiamo che

- se $n \geq 1$, l'operazione di addizione fra polinomi è interna in $\mathbb{R}_n[x[$, mentre l'operazione di moltiplicazione tra polinomi *non* è interna: basta moltiplicare un polinomio di grado n con uno di grado 1 per convincersene ($p(x) = x^n \in \mathbb{R}_n[x]$, $q(x) = x \in \mathbb{R}_n[x]$: allora $p(x).q(x) = x^{n+1} \notin \mathbb{R}_n[x]$);
- se $n = 0$, $\mathbb{R}_0[x]$ è formato solo da polinomi di grado 0, ossia solo dai termini noti, ed è in banale corrispondenza biunivoca con $\mathbb{R}$; poiché le operazioni di addizione e moltiplicazione fra numeri reali e fra polinomi di grado nullo sono definite alla stessa maniera, possiamo dire che $\mathbb{R}_0[x]$ è isomorfo a $\mathbb{R}$ (Definizione 0.21).

La proprietà seguente dell'anello dei polinomi $\mathbb{R}[x]$ è la generalizzazione di un'analoga ben nota proprietà dei numeri interi.

Proprietà 0.6 (**Algoritmo della divisione**). *Dati due polinomi $a(x)$ **(dividendo)** e $b(x)$ **(divisore)** in $\mathbb{R}[x]$, con $b(x)$ non identicamente nullo, esistono e sono univocamente determinati due polinomi $(q(x), r(x))$ appartenenti a $\mathbb{R}[x]$ tali che*

$$a(x) = b(x).q(x) + r(x),$$

inoltre $\deg r(x) < \deg b(x)$ *oppure* $r(x)$ *è il polinomio nullo.*

Il polinomio $q(x)$ *è detto* ***quoziente*** *e il polinomio* $r(x)$ *è detto* ***resto****, infine il procedimento con il quale si determinano i polinomi* $(q(x), r(x))$ *è detto* ***algoritmo della divisione****.*

Se $r(x)$ *è il polinomio nullo, allora* $a(x) = b(x).q(x)$*, diciamo che* $b(x)$ *è un* ***fattore*** *di* $a(x)$*, o equivalentemente* $a(x)$ *è* ***divisibile*** *per* $b(x)$*.*

La divisione si esegue ordinando secondo le potenze decrescenti i polinomi dividendo e divisore, e procedendo con il classico algoritmo di divisione studiato nelle scuole superiori.

Esercizio 4.1. A titolo di esempio, consideriamo i polinomi

$$a(x) = x^4 - 2x^2 + 4x + 8 \quad \text{e} \quad b(x) = x - 2\,.$$

Verificare eseguendo la divisione tra i polinomi che si ottiene

$$q(x) = x^3 + 2x^2 + 2x + 8 \qquad \text{e} \qquad r(x) = 24\,.$$

ossia

$$a(x) = (x-2)(x^3 + 2x^2 + 2x + 8) + 24.$$

Introdurremo adesso un po' di terminologia che, oltre a fornire strumenti per elaborare criteri sotto i quali un polinomio è divisibile per un altro, diviene utile per molti argomenti che affronteremo nel nostro corso.

Definizione 0.26. Sia $p(x) \in \mathbb{R}[x]$ un polinomio di grado $n > 0$: diremo che $\alpha \in \mathbb{R}$ è una **radice** di $p(x)$ se sostituendo α al posto di x nel polinomio $p(x)$ si ottiene 0, cioè se

$$p(\alpha) = a_0 + a_1\alpha + a_2(\alpha)^2 + ... + a_n(\alpha)^n = 0.$$

Equivalentemente, α è una radice del polinomo $p(x)$ se e solo se è una soluzione dell'equazione che si ottiene uguagliando a zero il polinomio

$$p(x) = 0;$$

tale equazione viene chiamata equazione **algebrica** nell'incognita x di grado n.

Trovare le radici di un polinomio equivale, quindi, a risolvere un'equazione algebrica in x di grado n.

Ricordiamo la proprietà fondamentale che lega le radici di un polinomio con la sua divisibilità per binomi di primo grado:

Proprietà 0.7 (**Teorema di Ruffini**). *Sia* $p(x) \in \mathbb{R}[x]$*, un polinomio di grado* $n > 0$*:* $\alpha \in \mathbb{R}$ *è una radice di* $p(x)$ *se e solo se* $p(x)$ *è divisibile per il polinomio* $(x - \alpha)$*.*

Dimostrazione. Infatti, applicando l'algoritmo della divisione ai polinomi $p(x)$ e $(x - \alpha)$, otteniamo due polinomi $q(x)$ e $r(x)$ tali che:

$$p(x) = (x - \alpha)q(x) + r(x),$$

con $\deg(r(x)) = 0$ (dovendo essere il grado del resto inferiore al grado del dividendo), cioè $r(x) = r \in \mathbb{R}$.

Sostituendo α al posto di x in entrambi i membri dell'uguaglianza scritta otteniamo:

$$p(\alpha) = 0.q(\alpha) + r = r\,.$$

Ora, se il $p(x)$ è divisibile per $x - \alpha$, il resto della divisione deve essere $r = 0$, e quindi $p(\alpha) = 0$, ossia α è radice di $p(x)$. Viceversa, se α è radice di $p(x)$, $p(\alpha) = 0$, quindi $r = 0$, cioè $p(x)$ è divisibile per il polinomio $(x - \alpha)$. □

Ricordiamo inoltre la nozione di molteplicità algebrica.

Definizione 0.27. Sia $p(x) \in \mathbb{R}[x]$ un polinomio di grado $n > 0$ e sia $\alpha \in \mathbb{R}$ una **radice** di $p(x)$:

- α è una radice di **molteplicità algebrica** $\mu \geq 1$ se e solo se $p(x)$ è divisibile per il polinomio $(x - \alpha)^{\mu}$ e $p(x)$ non è divisibile per il polinomio $(x - \alpha)^{\mu+1}$, ossia, se
$$p(x) = (x - \alpha)^{\mu} q(x)\,, \qquad \text{con } q(\alpha) \neq 0\,.$$
- Se $\mu = 1$ allora diremo che α è una **radice semplice**.

Le radici reali di un polinomio $p(x) \in \mathbb{R}[x]$ di grado $n > 0$ si ottengono quindi determinando tutti i i fattori di primo grado in x, ossia i cosiddetti **fattori lineari reali** di $p(x)$.

- Se $n = 1$, allora $p(x) = a_0 + a_1 x$, con $a_1 \neq 0$: il polinomio ha un'unica radice che si ottiene risolvendo l'equazione di primo grado
$$a_0 + a_1 x = 0 \;\Leftrightarrow\; x = -a_0 . {a_1}^{-1}.$$
- Se $n = 2$, allora $p(x) = a_0 + a_1 x + a_2 x^2$, con $a_2 \neq 0$: ricordiamo che l'equazione di secondo grado
$$a_0 + a_1 x + a_2 x^2 = 0,$$
ammette soluzioni reali se e solo il suo discriminante risulta positivo:
$$\Delta = a_1^2 - 4a_0 a_2 \geq 0.$$
In tal caso, le due soluzioni reali dell'equazione α_1 e α_2 sono le radici del polinomio e $p(x)$ può essere scritto nel seguente modo:
$$p(x) = a_2(x - \alpha_1)(x - \alpha_2).$$
Nel caso in cui $\Delta = 0$, le radici coincidono, $\alpha_1 = \alpha_2$, e la scrittura scritta sopra diventa:
$$p(x) = a_2(x - \alpha_1)^2.$$
Se invece risulta $\Delta = a_1^2 - 4a_0 a_2 < 0$ l'equazione non ammette soluzioni reali, di conseguenza il polinomio $p(x)$ non ha radici reali, $p(x)$ non ammette fattori lineari reali: $p(x)$ è detto un **polinomio irriducibile** in $\mathbb{R}$ di grado 2.
- Se, invece, $n \geq 2$, allora proviamo che $p(x)$ ha al massimo n fattori lineari reali.

  Infatti, sia $\alpha_1 \in \mathbb{R}$ una radice di $p(x)$ allora
$$p(x) = (x - \alpha_1) q_1(x),$$

dove $q_1(x)$ è un polinomio di grado $n-1$. Osserviamo ora che se $\alpha_2 \in \mathbb{R}$ è una radice di $q_1(x)$, allora $(x-\alpha_2)$ è un fattore di $q_1(x)$, quindi otteniamo:

$$p(x) = (x-\alpha_1)(x-\alpha_2).q_2(x),$$

dove $q_2(x)$ è un polinomio di grado $n-2$. Il procedimento ha termine se troviamo un quoziente $q_i(x)$ che non ha radici reali, $i \leq n-1$, oppure dopo n passi e l'ultimo quoziente $q_n(x) = a_n$ ha grado zero.

Le considerazioni appena fatte ci portano a caratterizzare quei polinomi per i quali la fattorizzazione è "totale".

Definizione 0.28. Nel caso in cui un polinomio $p(x) \in \mathbb{R}$ di grado $n \geq 2$ è prodotto di n fattori lineari reali, non necessariamente distinti, $p(x)$ è detto **totalmente decomponibile** in $\mathbb{R}$:

$$p(x) = a_n(x-\alpha_1)(x-\alpha_2)....(x-\alpha_n), \qquad \alpha_i \in \mathbb{R},\ i = 1, \dots n\,.$$

Osservazione 0.21. Se un polinomio totalmente decomponibile ha k radici distinte α_i, $i = 1, \dots k$, ciascuna di molteplicità algebrica μ_k, potremo scrivere

$$p(x) = a_n(x-\alpha_1)^{\mu_1}(x-\alpha_2)^{\mu_2} \ \dots \ (x-\alpha_k)^{\mu_k}.$$

Esempio 0.22. Per chiarire le idee, sono utili alcuni esempi.

(1) Il polinomio $p(x) = x^2 + 1$ non ammette radici reali, quindi è un polinomio di grado 2 irriducibile in $\mathbb{R}$.

(2) Il polinomio $p(x) = 3x^2 - 6x + 3$ ammette la radice $x = 1$ con molteplicità algebrica $\mu = 2$ e può essere scritto nel seguente modo:

$$p(x) = 3(x-1)^2,$$

risulta quindi totalmente decomponibile in $\mathbb{R}$.

(3) Il polinomio $p(x) = x^3 + x - x^2 - 1$ può essere scritto come prodotto di due fattori:

$$p(x) = (x^2+1).(x-1),$$

le radici di tale polinomio si ottengono quindi determinando le radici di ciascun fattore. Osserviamo che il polinomio x^2+1 è irriducibile in $\mathbb{R}$, quindi non ha radici reali. Il polinomo $x-1$ è lineare ed ha un'unica radice reale:

$$x - 1 = 0 \iff x = 1.$$

Il polinomio $p(x)$ ha quindi un'unica radice semplice $x = 1$.

(4) Il polinomio $p(x) = -x^3 + 3x^2 - 3x + 1$ può essere scritto nel seguente modo:

$$p(x) = (1-x)^3,$$

risulta quindi totalmente decomponibile in $\mathbb{R}$, ammette la radice $x = 1$ con molteplicità algebrica $\mu = 3$.

Analogamente a quanto avviene nell'anello dei numeri interi $\mathbb{Z}$ per la scomposizione di un numero in fattori primi, esiste un teorema che garantisce l'unicità della scomposizione di un polinomio in fattori. Tale risultato è il **Teorema**

Fondamentale dell'Algebra e vale più in generale per polinomi a cofficienti nel campo complesso:

TEOREMA 0.8. *Sia $p(x) \in \mathbb{C}[x]$ un polinomio di grado $n \geq 1$, $p(x)$ è totalmente decomponibile in fattori lineari (non necessariamente distinti) in $\mathbb{C}$:*

$$p(x) = a_n(x - \alpha_1)(x - \alpha_2)....(x - \alpha_n), \quad \alpha_i \in \mathbb{C}, i = 1, .., n;$$

tale decomposizione è unica a meno dell'ordine dei fattori.

OSSERVAZIONE 0.23. Come conseguenza del Teorema Fondamentale dell'Algebra 0.8, ogni equazione algebrica a coefficienti complessi in x di grado n ammette n soluzioni (non necessariamente distinte) in $\mathbb{C}$,

DEFINIZIONE 0.29. Quanto riportato nell'Osservazione precedente 0.23 si esprime sinteticamente dicendo che il campo $\mathbb{C}$ è algebricamente chiuso.

Nel caso di polinomi a coefficienti reali vale un risultato molto più debole:

TEOREMA 0.9. *Sia $p(x) \in \mathbb{R}[x]$ un polinomio di grado $n \geq 1$, $p(x)$ si decompone in $\mathbb{R}$ nel prodotto di fattori di grado ≤ 2:*

$$p(x) = a_n p_1(x).p_2(x)....p_k(x), \quad k \leq n, \quad \deg p_i(x) \leq 2, \forall i = 1 \ldots k,$$

con fattori $p_i(x)$ tali che, se $\deg(p_i(x)) = 2$, allora $p_i(x)$ è irriducible. Tale decomposizione è unica a meno dell'ordine dei fattori.

ESEMPIO 0.24. Se consideriamo il polinomio

$$p(x) = 2x^3 - 2$$

si può scomporre come

$$p(x) = 2x^3 - 2 = 2(x - 1)(x^2 + x + 1) = 2\, p_1(x)\, p_2(x)$$

ed il polinomio $p_2(x) = (x^2 + x + 1)$ è irriducibile (non ha radici reali: il discriminante dell'equazione per le radici è $\Delta = 1 - 4 = -3 < 0$).

Esiste però un corollario molto utile del Teorema Fondamentale: per un polinomio reale di grado dispari possiamo garantire l'esistenza di una radice reale di un polinomio.

COROLLARIO 0.10. *Sia $p(x) \in \mathbb{R}[x]$ un polinomio di grado $n \geq 1$, con n dispari. Allora $p(x)$ ammette almeno una radice reale.*

DIMOSTRAZIONE. Scriviamo esplicitamente il polinomio $p(x)$:

$$p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0\,, \tag{0.6}$$

con $a_i \in \mathbb{R}$ per ogni $i = 0 \ldots n$. Poiché $\mathbb{R} \subseteq \mathbb{C}$, possiamo considerare $p(z)$ come un polinomio di $\mathbb{C}[z]$. Come abbiamo visto sopra, grazie al Teorema Fondamentale dell'Algebra (0.8), $p(z)$ ammette n radici in $\mathbb{C}$, contando le molteplicità, ossia

$$p(z) = a_n z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0 = a_n(z - \alpha_1)(z - \alpha_2) \ldots (z - \alpha_n)\,, \tag{0.7}$$

con gli α_i non necessariamente tutti distinti.

Mostriamo, anzitutto, che, a queste condizioni, se z_0 è una radice, anche il complesso coniugato $\overline{z_0}$ (Definizione 0.34) è una radice di $p(z)$. Infatti, se z_0 è una radice

$$p(z_0) = a_n z_0^n + a_{n-1}z_0^{n-1} + \cdots + a_1 z_0 + a_0 = 0\,; \tag{0.8}$$

ora, prendiamo il complesso coniugato di entrambi i membri della 0.8 otteniamo, sfruttando le proprietà del coniugio (Proprietà 0.12) ed il fatto che i coefficienti a_i sono reali, quindi $\overline{a_i} = a_i$ per ogni $i = 0 \dots n$:

$$
\begin{aligned}
\overline{p(z_0)} &= \overline{a_n z_0^n + a_{n-1} z_0^{n-1} + \cdots + a_1 z_0 + a_0} \\
&= \overline{a_n z_0^n} + \overline{a_{n-1} z_0^{n-1}} + \cdots + \overline{a_1 z_0} + \overline{a_0} \\
&= \overline{a_n}\,\overline{z_0^n} + \overline{a_{n-1}}\,\overline{z_0^{n-1}} + \cdots + \overline{a_1}\,\overline{z_0} + \overline{a_0} \\
&= \overline{a_n}(\overline{z_0})^n + \overline{a_{n-1}}(\overline{z_0})^{n-1} + \cdots + \overline{a_1}\,\overline{z_0} + \overline{a_0} \\
&= a_n(\overline{z_0})^n + a_{n-1}(\overline{z_0})^{n-1} + \cdots + a_1\overline{z_0} + a_0 \\
&= p(\overline{z_0}) = \overline{0} \\
&= 0\,,
\end{aligned}
\tag{0.9}
$$

quindi $\overline{p(z_0)} = 0$, cioè $p(\overline{z_0}) = 0$, ossia, $\overline{z_0}$ è radice di $p(z)$.

A questo punto, possiamo "scartare" le radici complesse di $p(z)$ a due a due, prendendo le coppie z_0 e $\overline{z_0}$: in altre parole, possiamo ordinare così le radici in $\mathbb{C}$ del polinomio: prendiamo una radice, α_1; anche $\overline{\alpha_1}$ deve essere radice, e diciamo che questa sia α_2; poi, prendiamo un'altra radice, α_3 e chiamiamo α_4 la radice $\overline{\alpha_3}$. Siccome le radici totali sono in numero dispari, restiamo almeno con l'ultima, che deve essere "accoppiata" con se stessa, ossia $\alpha_n = \overline{\alpha_n}$, che mostra che α_n è reale. Ovviamente, non necessariamente tutte le radici scartate sono complesse, anche fra loro si possono trovare altre radici reali autoconiugate (in questo caso, scartando le radici complesse coniugate a coppie, prima di "arrivare in fondo", troveremo una radice reale). □

Osservazione 0.25. Il Corollario 0.10 *non* afferma né che la radice reale sia unica, né che le radici non possono essere *tutte* reali: in effetti, entrambi questi casi possono verificarsi tranquillamente.

APPENDICE

5. Numeri complessi

Le prime idee sui numeri complessi sono venute a formarsi nell'epoca rinascimentale, in particolare grazie a G. Cardano, che introdusse il concetto di **unità immaginaria** $i = \sqrt{-1}$, ossia tale che

$$i^2 = -1 \tag{0.10}$$

per arrivare ad una formula risolutiva di alcune equazioni di terzo grado; per esempio, in una formula che stava elaborando per risolvere l'equazione $x^3 = 15x + 4$ (di cui è nota la soluzione $x = 4$), Cardano si trova a dover usare la quantità $\sqrt{-121} = 11\sqrt{-1}$. "Immaginando" che la radice quadrata di -1 esista e sia i, e procedendo, riesce a trovare il numero reale 4 come soluzione dell'equazione.

Con il tempo, si sviluppa l'idea di considerare l'insieme dei numeri "complessi", ossia quelli "composti" da una *parte reale* e da una *parte immaginaria.*

Definizione 0.30. Una maniera per "distinguere" le due parti è ovviamente quella di rappresentare un numero complesso z tramite la cosiddetta **rappresentazione algebrica**

$$z = a + i\,b \qquad a = \Re(z) \in \mathbb{R}\,, \quad b = \Im(z) \in \mathbb{R}\,, \tag{0.11}$$

dove chiameremo i numeri reali a, b, rispettivamente, la **parte reale** e la **parte immaginaria** del numero complesso z. Indicheremo con $\mathbb{C}$ l'insieme di tutti i numeri complessi.

Fino a questo punto, il simbolo $+$ serve solo a distinguere le due parti, assieme alla i che identifica la parte immaginaria. Una maniera alternativa per fare questo, più formale, sarebbe quella di scrivere il numero complesso z come la coppia ordinata (a, b) in cui il primo elemento è $\Re(z)$ ed il secondo elemento è $\Im(z)$.

Vogliamo dotare $\mathbb{C}$ di una struttura algebrica, in particolare occorre definire le operazioni interne di addizione e moltiplicazione, che indicheremo con i consueti simboli.

Definizione 0.31. Siano $z_1 = (a_1, b_1) \equiv a_1 + i\,b_1$ e $z_2 = (a_2, b_2) \equiv a_2 + i\,b_2$ due numeri complessi.

- L'operazione di addizione fra i due numeri è definita in modo che la somma abbia parte reale coincidente con la somma delle parti reali e parte immaginaria coincidente con la somma delle parti immaginarie:

$$z_1 + z_2 = (a_1 + a_2, b_1 + b_2) \equiv (a_1 + a_2) + i\,(b_1 + b_2). \tag{0.12}$$

- L'operazione di moltiplicazione fra i due numeri è definita in modo che il prodotto sia dato da

$$z_1.z_2 = (ac - bd, ad + bc) \equiv (ac - bd) + i\,(ad + bc). \tag{0.13}$$

ESERCIZIO 5.1. Si lascia allo studente il compito (un po' noioso) di verificare che, con le operazioni definite in 0.31:

- l'addizione gode delle proprietà commutativa e associativa;
- $(0,0)$ è l'elemento neutro per la somma: lo indicheremo semplicemente come 0;
- dato $(a,b) \in \mathbb{C}$, il numero $(-a,-b)$ è il suo opposto per la somma;
- la moltiplicazione gode delle proprietà commutativa e associativa;
- vale la proprietà distributiva del prodotto rispetto alla somma;
- $(1,0)$ è l'elemento neutro per il prodotto: lo indicheremo semplicemente come 1.

Quindi, $(\mathbb{C},+,.)$ è un **anello commutativo unitario** (Definizione 0.19).

OSSERVAZIONE 0.26. Appare subito chiaro che, per come abbiamo definito le due operazioni, la rappresentazione algebrica è pienamente giustificata dal fatto che si possono eseguire le operazioni di addizione e moltiplicazione sfruttando le consuete proprietà, trattando i come una lettera in un polinomio, e ricordando che $i^2 = -1$.

Ad esempio, se $z_1 = a + i\,b$ e $z_2 = c + i\,d$:

$$\begin{aligned} z_1.z_2 &= (a + i\,b).(c + i\,d) \\ &= a.(c + i\,d) + i\,b.(c + i\,d) \\ &= a.c + i\,a.d + i\,b.c + i^2 b.d \\ &= a.c + i\,(b.c + a.d) - b.d \\ &= (a.c - b.d) + i\,(b.c + a.d). \end{aligned}$$

Quindi, il simbolo di somma nella rappresentazione algebrica $a + i\,b$ può essere, a tutti gli effetti trattato come la somma tra i numeri complessi $(a,0) = a + i\,0$ e $(0,b) = 0 + i\,b$. Risulta pertanto di immediata verifica anche che

- dato $(a,b) \in \mathbb{C}$ diverso da 0, il numero $(\frac{a}{a^2+b^2}, -\frac{b}{a^2+b^2})$ è il suo inverso per il prodotto, infatti, usando la rappresentazione algebrica:

$$\begin{aligned} &(a+ib)(\frac{a}{a^2+b^2} - i\frac{b}{a^2+b^2}) \\ &= \frac{a^2}{a^2+b^2} - i\frac{ab}{a^2+b^2} + i\frac{ba}{a^2+b^2} - i^2\frac{b^2}{a^2+b^2} \\ &= \frac{a^2}{a^2+b^2} + \frac{b^2}{a^2+b^2} + i.0 \\ &= \frac{a^2+b^2}{a^2+b^2} = 1\,. \end{aligned}$$

Quindi, $(\mathbb{C},+,.)$ è un **campo** (Definizione 0.20).

OSSERVAZIONE 0.27. Conviene sviluppare alcune ulteriori osservazioni.

(1) Esiste una corrispondenza biunivoca immediata fra i numeri complessi $\mathbb{C}$ ed i punti del piano cartesiano $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$; al numero complesso $z = a + ib$ corrisponde il punto del piano di coordinate (a,b). Questo permetterà di avere una rappresentazione grafica molto efficace dei numeri complessi.

(2) Questa corrispondenza biunivoca è un isomorfismo tra campi, se in $\mathbb{R}^2$ definiamo le operazioni di addizione e moltiplicazione fra le coppie ordinate esattamente come abbiamo fatto per i numeri complessi (Definizione 0.31).
(3) I numeri $z \in \mathbb{C}$ che hanno parte immaginaria nulla sono in corrispondenza biunivoca con i numeri reali; basta prendere il numero reale coincidente con la parte reale di z.

$$(a, 0) = a + i\,0 \simeq a.$$

Questa corrispondenza è un isomorfismo con il campo $\mathbb{R}$; se $a, b \in \mathbb{R}$:

$$(a, 0) + (b, 0) = ((a + b), 0) \simeq a + b \tag{0.14}$$

$$(a, 0).(b, 0) = (a.b - 0.0, a.0 + 0.b) = ((a.b), 0) \simeq a.b\,. \tag{0.15}$$

In base all'osservazione 0.27 (1), possiamo **identificare i numeri reali con i numeri complessi con parte immaginaria nulla**, ossia vale la

Proprietà 0.11. *un numero complesso è reale se e solo se la sua parte immaginaria è nulla:*

$$z \in \mathbb{R} \iff \Im(z) = 0\,.$$

Definizione 0.32. Analogamente, chiameremo **numeri immaginari puri** i numeri complessi che hanno parte reale nulla.

5.1. Piano di Gauss-Argand.

La corrispondenza biunivoca che abbiamo stabilito fra numeri complessi e coppie ordinate di numeri reali, ci consente di visualizzare in modo efficace i numeri complessi come punti del piano cartesiano, secondo una rappresentazione suggerita da Gauss e Argand (Figura 0.7). I due assi coordinati, in questa rappresentazione vengono chiamati, rispettivamente **asse reale** ed **asse immaginario**.

5.2. Modulo e coniugio.

In questa sezione vogliamo caratterizzare ulteriormente un numero complesso. Notiamo che ad ogni numero complesso $z = a + ib$ viene naturalmente associata la distanza dall'origine del punto $P = (a, b)$ che lo rappresenta nel piano di Gauss-Argand; chiameremo **modulo** di z, quindi, la lunghezza del segmento OP:

Definizione 0.33. Dato un numero complesso $z = a+ib$ chiamiamo *modulo* di z il numero reale

$$|z| = \sqrt{a^2 + b^2}\,.$$

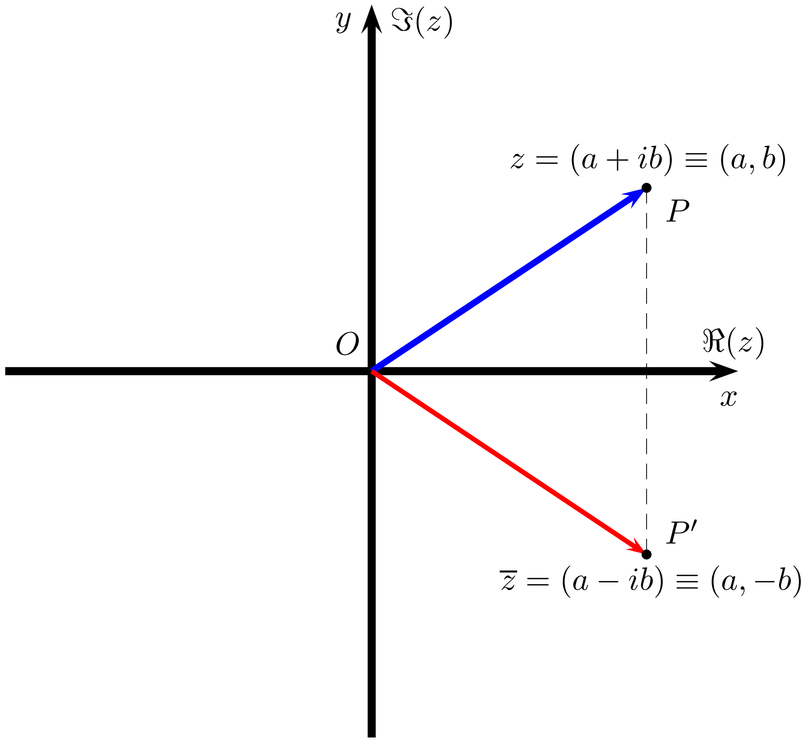

FIGURA 0.7. Il piano di Gauss-Argand, che realizza l'isomorfismo fra il piano cartesiano ed i numeri complessi. Ogni punto P del piano, identificato dalla coppia ordinata delle sue coordinate cartesiane (a, b) è in corrispondenza biunivoca con il numero complesso $z = a + i\,b$. Il numero complesso coniugato $\overline{z} = a - i\,b$ corrisponde, pertanto, al punto P' simmetrico di P rispetto all'asse delle ascisse.

OSSERVAZIONE 0.28. Per il modulo valgono le seguenti proprietà:

$$|z| \geq 0\,;$$
$$|z| = 0 \iff a^2 + b^2 = 0 \iff a = 0 \wedge b = 0 \iff z = 0\,.$$

Inoltre, possiamo associare a P il suo simmetrico rispetto all'asse reale: l'operazione che effettua questa associazione viene detta **coniugio** o coniugazione:

DEFINIZIONE 0.34. Dato un numero complesso $z = a + ib$, chiamiamo **numero complesso coniugato** di z, il numero $\overline{z} = a - ib$, che ha uguale parte reale e parte immaginaria opposta

PROPRIETÀ 0.12. *In base a questo, valgono alcune proprietà, la cui verifica, quasi immediata, viene lasciata come esercizio.*

(1) *Un numero complesso è reale se e solo se coincide con il suo complesso coniugato*
$$z \in \mathbb{R} \iff z = \overline{z}\,.$$

(2) *La parte reale di un numero complesso è la semisomma del numero con il suo complesso coniugato; la parte immaginaria di un numero complesso è la semidifferenza fra il numero ed il suo complesso coniugato, divisa per i:*

$$\Re(z) = \frac{z + \overline{z}}{2},$$
$$\Im(z) = \frac{z - \overline{z}}{2i}.$$

(3) *Il quadrato del modulo di un numero complesso è dato dal prodotto del numero con il suo complesso coniugato:*

$$z.\overline{z} = (a + ib)(a - ib) = a^2 - (ib)^2 = a^2 + b^2 = |z|^2 .$$

(4) *Il coniugato del coniugato di un numero complesso è il numero stesso:*

$$\overline{(\overline{z})} = z .$$

(5) *Il complesso coniugato di una somma o di un prodotto di numeri complessi è la somma o il prodotto dei complessi coniugati dei singoli numeri*

$$\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2},$$
$$\overline{z_1.z_2} = \overline{z_1}.\overline{z_2}.$$

(6) *Un numero ed il suo complesso coniugato hanno lo stesso modulo:*

$$|\overline{z}| = |z|.$$

(7) *Il modulo del prodotto di due numeri complessi è il prodotto dei moduli:*

$$|z_1.z_2| = |z_1|.|z_2| .$$

ATTENZIONE! L'ultima proprietà non vale con la somma; infatti, vale la seguente

PROPOSIZIONE 0.13 (Disuguaglianza triangolare). *Siano*

$$z = a + ib \quad e \quad w = c + id$$

due numeri complessi; in generale, vale la seguente disuguaglianza

$$|z + w| \leq |z| + |w|. \tag{0.16}$$

La disuguaglianza ha un'immediata interpretazione geometrica (da cui il nome), rappresentata nella Figura 0.8.

DIMOSTRAZIONE. Poiché i moduli sono numeri reali positivi (Osservazione 0.28), equivale a dimostrare la disuguaglianza:

$$|z + w|^2 \leq (|z| + |w|)^2, \tag{0.17}$$

ossia, in base alla definizione di modulo e di somma di numeri complessi, che

$$(a + c)^2 + (b + d)^2 \leq (\sqrt{a^2 + b^2} + \sqrt{c^2 + d^2})^2, \tag{0.18}$$

cioè, sviluppando i quadrati

$$a^2 + c^2 + 2ac + b^2 + d^2 + 2bd \leq a^2 + b^2 + c^2 + d^2 + 2\sqrt{a^2 + b^2}\sqrt{c^2 + d^2},$$

o anche, togliendo a destra e sinistra le quantità uguali e dividendo per 2

$$ac + bd \leq \sqrt{a^2 + b^2}\sqrt{c^2 + d^2}.$$

Ora, se $ac + bd \leq 0$, la disugaglianza è vera, essendo il secondo mebbro al più nullo; altrimenti, possiamo ancora elevare al quadrato entrambi i membri e resta da dimostrare che

$$a^2c^2 + b^2d^2 + 2abcd \leq (a^2 + b^2)(c^2 + d^2) = a^2c^2 + b^2d^2 + a^2d^2 + b^2c^2,$$

equivalente a

$$2abcd \leq a^2d^2 + b^2c^2,$$

o anche

$$(ad - bc)^2 = a^2d^2 + b^2c^2 - 2abcd \geq 0,$$

che è vera, avendo scritto il primo membro di questa nuova disequazione equivalente a quella di partenza come un quadrato. □

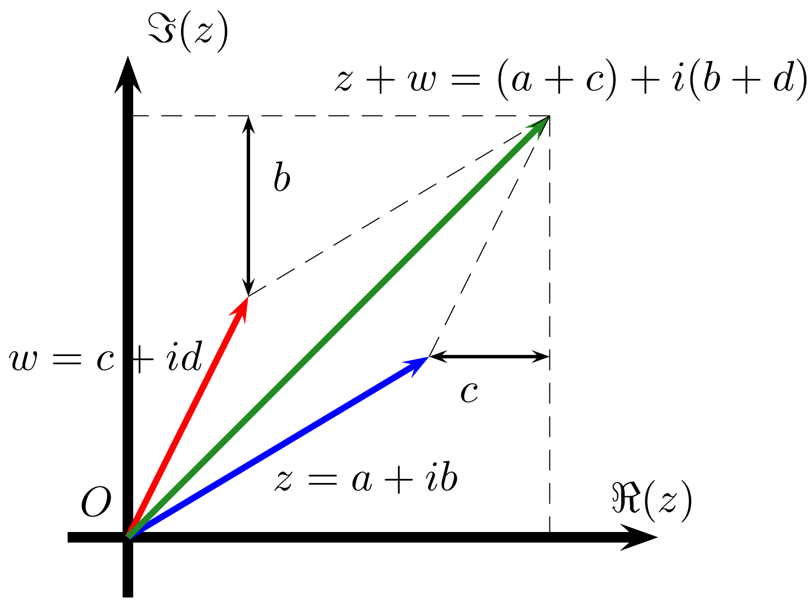

FIGURA 0.8. La rappresentazione geometrica della somma di due numeri complessi $z = a + ib$ e $w = c + id$ nel piano di Gauss-Argand, dalla quale si comprende il fondamento della disuguaglianza triangolare (Equazione 0.16); il modulo della somma $z + w$ è la diagonale del parallelogramma che ha per lati i segmenti di lunghezza pari ai moduli di z e w: poiché in un triangolo ogni lato è minore della somma degli altri si ottiene la disuguaglianza. Lo studente è invitato a convicersi della validità della disuguaglianza per tutte le combinazioni di appartenenza di z e w ai differenti quadranti coordinati.

La definizione di numero complesso coniugato ci permette di scrivere in modo compatto l'inverso di un numero complesso non nullo (ultimo punto dell'Osservazione 0.27): osservato che

$$z \neq 0 \iff |z| \neq 0, \tag{0.19}$$

verifichiamo che, per $z \neq 0$:

$$z^{-1} = \frac{\overline{z}}{|z|^2}. \tag{0.20}$$

Infatti (punto 3 delle Proprietà 0.12):

$$z.z^{-1} = \frac{z\overline{z}}{|z|^2} = \frac{|z|^2}{|z|^2} = 1. \tag{0.21}$$

Esempio 0.29. Calcoliamo l'inverso del numero $z = 3 - 2i$.
Si ha $|z| = \sqrt{9+4} = \sqrt{13}$ e, quindi

$$z^{-1} = \frac{3+2i}{13} = \frac{3}{13} + i\frac{2}{13}$$

Verifichiamo:

$$(3-2i)(\frac{3+2i}{13}) = \frac{3^2-(2i)^2}{13} = \frac{9+4}{13} = \frac{13}{13} = 1.$$

Osservazione 0.30. Se $z = a + ib \in \mathbb{R}$, allora $b = 0$, e $|z| = \sqrt{a^2+0} = |a|$, ossia il modulo definito nei complessi per i numeri reali coincide con il valore assoluto (o modulo, appunto) sui reali.

Inoltre, se $a \neq 0$, $z^{-1} = \frac{a-i0}{a^2} = \frac{1}{a}$, e anche l'inverso coincide con quello definito per i reali.

6. Rappresentazione trigonometrica

Esiste un modo alternativo di rappresentare i numeri complessi, utile per calcolare le potenze di un numero complesso o per risolvere equazioni del tipo

$$z^n = w \in \mathbb{C}, \qquad \text{per } z \in \mathbb{C}.$$

Si tratta della **rappresentazione trigonometrica**; in base a questa, anziché individuare il punto P nel piano di Gauss-Argand che rappresenta il numero complesso $z = a+ib$ con la coppia ordinata di coordinate (a, b), lo individuiamo –come si fa in coordinate polari– mediante

- la distanza di P dall'origine (ossia il **modulo** del numero complesso $\rho = |z|$);
- l'angolo che il segmento OP forma con il semiasse delle x positive, che chiameremo **argomento** del numero complesso $\arg z$.

(Figura 0.9). Assegnati un modulo ρ ed un argomento $\theta = \arg z$, la parte reale e la parte immaginaria si calcolano immediatamente, in maniera univoca:

$$a = \Re(z) = \rho\cos\theta, \tag{0.22}$$

$$b = \Im(z) = \rho\sin\theta. \tag{0.23}$$

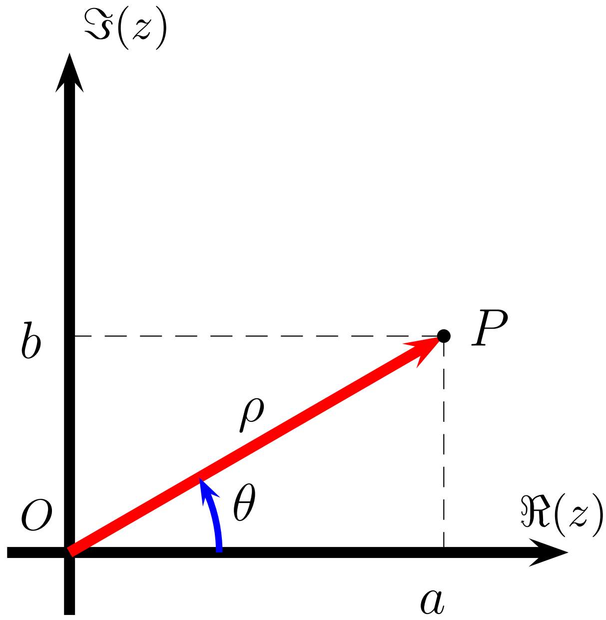

Figura 0.9. La rappresentazione trigonometrica di un numero complesso a confronto con quella algebrica.

Al contrario, assegnati a, parte reale e b, parte immaginaria di un numero complesso, risulta immediata la determinazione del modulo:

$$\rho = \sqrt{a^2 + b^2}\,. \tag{0.24}$$

Invece, per l'argomento di $z \in \mathbb{C}$ dati la parte reale e la parte immaginaria occorre un minimo di attenzione in più. Anzitutto, osserviamo che esistono infiniti angoli che indiduano P.

Osservazione 0.31. Due angoli che individuano la direzione di OP rispetto al semiasse x positivo differiscono fra di loro per un numero intero di angolo giro $2k\pi$, con $k \in \mathbb{Z}$.

Esiste, cioè, una *relazione di equivalenza* fra argomenti di un numero complesso, che consentono di raggrupparli in classi di equivalenza; questo consente di definire un argomento “standard” come la classe di equivalenza degli argomenti.

Conveniamo pertanto, quando opportuno, di ricondurre l'angolo dell'argomento all'intervallo $(-\pi, \pi]$:

Definizione 0.35. Chiamiamo **argomento principale** di un numero complesso $z = a + ib$ l'argomento di z ricondotto all'intervallo $(-\pi, \pi]$, e verrà indicato con $\operatorname{Arg} z$.

Sinteticamente, l'argomento principale di $z = a + ib$ è dato dalla funzione $\operatorname{arctg}^*(a, b)$, che nel dettaglio risulta definita secondo lo schema seguente:

Algoritmo 0.14 (Calcolo di $\operatorname{Arg} z = \operatorname{arctg}^*(a, b)$). Sia $z = a + ib \in \mathbb{C}$

- se $b = 0$ e $a = 0$ (cioè, se $\rho = 0$), $\operatorname{Arg} z = 0$ (per convenzione);

- se $b = 0$ e $a < 0$, $\operatorname{Arg} z = \pi$;
- se $a = 0$ e $b > 0$, $\operatorname{Arg} z = \frac{\pi}{2}$;
- se $a = 0$ e $b < 0$, $\operatorname{Arg} z = -\frac{\pi}{2}$;
- se $a > 0$, $\operatorname{Arg} z = \operatorname{arctg}\left(\frac{b}{a}\right)$;
- se $a < 0$ e $b > 0$, $\operatorname{Arg} z = \operatorname{arctg}\left(\frac{b}{a}\right) + \pi$;
- se $a < 0$ e $b < 0$, $\operatorname{Arg} z = \operatorname{arctg}\left(\frac{b}{a}\right) - \pi$;

L'algoritmo può risultare un po' articolato, ma si può riassumere brevemente dicendo che l'argomento principale è essenzialmente dato dall'arcotangente del rapporto b/a, con l'aggiunta dell'informazione circa il quadrante in cui cade il punto corrispondente a z (con questo accorgimento, includiamo anche il caso in cui il rapporto non è definito, cioè $a = 0$).

Esempio 0.32. Qualche esempio aiuta a comprendere meglio quanto detto.

(1) Sia $z = \sqrt{3} + 3i$; calcolare $\rho = |z|$ e $\operatorname{Arg} z$.
Per il modulo, abbiamo
$$\rho = \sqrt{(\sqrt{3})^2 + 3^2} = \sqrt{3+9} = 2\sqrt{3}.$$
Il numero complesso cade nel primo quadrante del piano di Gauss-Argand, quindi
$$\operatorname{Arg} z = \operatorname{arctg}(\frac{3}{\sqrt{3}}) = \operatorname{arctg}(\sqrt{3}) = \frac{\pi}{3}.$$

(2) Sia $z = -2i$; calcolare $\rho = |z|$ e $\operatorname{Arg} z$.
Per il modulo, abbiamo
$$\rho = \sqrt{0^2 + (-2)^2} = 2.$$
Il numero è immaginario puro, con parte reale negativa, quindi:
$$\operatorname{Arg} z = -\frac{\pi}{2}$$

(3) Sia $z = -1 + i$; calcolare $\rho = |z|$ e $\operatorname{Arg} z$.
Per il modulo, abbiamo
$$\rho = \sqrt{(-1)^2 + 1^2} = \sqrt{2}.$$
Il numero complesso cade nel secondo quadrante del piano di Gauss-Argand, quindi
$$\operatorname{Arg} z = \operatorname{arctg}(-1) + \pi = -\frac{\pi}{4} + \pi = \frac{3}{4}\pi$$

(4) Sia $z = -(\sqrt{3} + i)$; calcolare $\rho = |z|$ e $\operatorname{Arg} z$.
Per il modulo, abbiamo
$$\rho = \sqrt{(\sqrt{3})^2 + 1^2} = 2.$$
Il numero complesso cade nel terzo quadrante del piano di Gauss-Argand, quindi
$$\operatorname{Arg} z = \operatorname{arctg}(\frac{1}{\sqrt{3}}) - \pi = \frac{\pi}{6} - \pi = -\frac{5}{6}\pi.$$

La rappresentazione trigonimetrica risulta particolarmente utile quando si devono moltiplicare o dividre fra loro numeri complessi: vale infatti la seguente regola.

TEOREMA 0.15. *Siano z_1, z_2 due numeri complessi, e siano ρ_1, θ_1 e ρ_2, θ_2, rispettivamente, i moduli e gli argomenti dei numeri.*

- *Il prodotto $z_1.z_2$ ha per modulo il prodotto dei moduli, e per argomento la somma degli argomenti.*
- *Se $z_2 \neq 0$, il rapporto $\frac{z_1}{z_2}$ ha per modulo il rapporto dei moduli e per argomento la differenza fra il primo ed il secondo argomento.*

In altre parole:

$$|z_1.z_2| = \rho_1\rho_2 \qquad \operatorname{Arg}(z_1.z_2) = \operatorname{Arg} z_1 + \operatorname{Arg} z_2; \tag{0.25}$$

$$|\frac{z_1}{z_2}| = \frac{\rho_1}{\rho_2} \qquad \operatorname{Arg}(\frac{z_1}{z_2}) = \operatorname{Arg} z_1 - \operatorname{Arg} z_2; \tag{0.26}$$

DIMOSTRAZIONE. Scriviamo esplicitamente in forma algebrica (Equazioni 0.22) i due numeri:

$$z_1 = \rho_1 \cos\theta_1 + i\rho_1 \sin\theta_1 \qquad = \rho_1(\cos\theta_1 + i\sin\theta_1), \tag{0.27}$$

$$z_2 = \rho_2 \cos\theta_2 + i\rho_2 \sin\theta_2 \qquad = \rho_2(\cos\theta_2 + i\sin\theta_2). \tag{0.28}$$

Calcoliamo il prodotto; raccogliendo i termini opportunamente, abbiamo:

$$z_1.z_2 = \rho_1\rho_2\left[(\cos\theta_1\cos\theta_2 - \sin\theta_1\sin\theta_2) + i(\sin\theta_1\cos\theta_2 + \sin\theta_2\cos\theta_1)\right] \tag{0.29}$$

ossia, usando le note formule della trigonometria per gli angoli somma:

$$z_1.z_2 = \rho_1\rho_2\left[\cos(\theta_1+\theta_2) + i\sin(\theta_1+\theta_2)\right], \tag{0.30}$$

che mostra esattamente il primo risultato.

Analogamente, si può procedere per il secondo punto, osservando che la divisione per z_2 equivale alla moltiplicazione per il numero inverso $\frac{1}{z_2}$, e che questo numero deve avere modulo pari a $1/\rho_2$ ed argomento $-\theta_2$. Si lascia allo studente di seguire il dettaglio della verifica. □

6.1. Formula di Eulero.

La rappresentazione trigonometrica si può scrivere in forma compatta usando la notazione dovuta ad Eulero:

DEFINIZIONE 0.36. Sia $\theta \in \mathbb{R}$; definiamo l'esponenziale di argomento immaginario $i\theta$ come

$$\mathrm{e}^{i\theta} = \cos\theta + i\sin\theta. \tag{0.31}$$

La scrittura introdotta nell'Equazione 0.31 consente di scrivere un numero complesso di modulo ρ ed argomento θ come

$$z = a + ib = \rho\,\mathrm{e}^{i\theta},$$

dove a, b sono legate a ρ, θ come visto prima. La scrittura è molto compatta, e si può usare per il calcolo dei prodotto e dei rapporti, usando formalmente le consuete regole per la moltiplicazione/divisione di potenze di uguale base.

ESEMPIO 0.33. Sia $z = \sqrt{3} + 3i$ e $w = -2i$ (i primi due numeri complessi dell'esempio 0.32). Possiamo scrivere

$$z = 2\sqrt{3}\,\mathrm{e}^{i\frac{\pi}{3}} \qquad w = 2\,\mathrm{e}^{-i\frac{\pi}{2}}.$$

Il prodotto risulta:

$$z.w = 4\sqrt{3}\,\mathrm{e}^{-i\frac{\pi}{6}} = 4\sqrt{3}(\frac{\sqrt{3}}{2} - i\frac{1}{2}) = 6 - 2\sqrt{3}\,i = 2(3 - \sqrt{3}\,i)$$

(verificare il risultato nella forma algebrica).

6.2. Formula di DeMoivre.

Per il calcolo delle potenze di un numero complesso, viene molto utile la cosiddetta formula di De Moivre:

$$(\cos\theta + i\sin\theta)^n = \cos(n\theta) + i\sin(n\theta)\,, \tag{0.32}$$

con $n \in \mathbb{N}$. La dimostrazione della formula si conduce per induzione, e viene lasciata come esercizio.

In base alla formula, se un numero complesso è noto in forma trigonometrica, possiamo calcolare agevolmente la sua $n-$esima potenza. Ricorrendo alla notazione di Eulero (Formula 0.31). Se

$$z = (\cos\theta + i\sin\theta) = \rho\,\mathrm{e}^{i\theta} \tag{0.33}$$

allora

$$z^n = \rho^n\,(\cos\theta + i\sin\theta)^n = \rho^n\,\left(\cos(n\theta) + i\sin(n\theta)\right) = \rho^n\,\mathrm{e}^{in\theta} \tag{0.34}$$

Sfruttando la formula di De Moivre, possiamo affrontare problemi come quelli di calcolo di potenze in modo particolarmente efficace.

ESEMPIO 0.34. Calcolare $(1+i)^{17}$.

Ovviamente, sarebbe possibile procedere algebricamente, moltiplicando la base $(1+i)$ per se stessa 17 volte; possiamo però procedere più speditamente nella seguente maniera. Poniamo $z = 1 + i$.

Il nostro compito è quello di calcolare z^{17}; calcoliamo modulo ed argomento di z:

$$|z| = \sqrt{1+1} = \sqrt{2}, \qquad \operatorname{Arg} z = \operatorname{arctg} 1 = \frac{\pi}{4}.$$

Quindi, usando la formula di De Moivre

$$\begin{aligned}
(1+i)^{17} &= z^{17} = (\sqrt{2})^{17}\,\mathrm{e}^{i\,17\frac{\pi}{4}}\\
&= (\sqrt{2})^{16}\sqrt{2}\left(\cos(\frac{17}{4}\pi) + i\sin(\frac{17}{4}\pi)\right)\\
&= 256\sqrt{2}\left[\cos\left(\frac{16}{4}\pi + \frac{1}{4}\pi\right) + i\sin\left(\frac{16}{4}\pi + \frac{1}{4}\pi\right)\right]\\
&= 256\sqrt{2}\left[\cos\left(4\pi + \frac{1}{4}\pi\right) + i\sin\left(4\pi + \frac{1}{4}\pi\right)\right]\\
&= 256\sqrt{2}\left(\frac{\sqrt{2}}{2} + i\frac{\sqrt{2}}{2}\right) = 256(1+i).
\end{aligned}$$

Concludiamo con un'applicazione classica: l'utilizzo della formula di De Moivre per calcolare radici $n-$esime di numeri complessi, o meglio le soluzioni di equazioni del tipo

$$z^n = w \qquad z, w \in \mathbb{C}$$

con w noto.

Osservazione 0.35. Ricordiamo che, per numeri reali x, la scrittura $\sqrt[n]{x}$ va intesa come la radice **aritmetica** di x, che esiste sempre se n è dispari, e per $x \geq 0$ se n è pari. La radice aritmetica è sempre positiva; in altre parole, per esempio $\sqrt{4} = 2 \geq 0$; invece l'equazione $x^2 = 4$ ha due souzioni (ossia, il polinomio $x^2 - 4$ ha due radici (Definizione 0.26), che sono $x = \pm\sqrt{4} = \pm 2$.

In questo senso, a volte si parla di "radici" di un numero come sinonimo di soluzioni dell'equazione algebrica corrispondente.

In base a quanto detto nell'osservazione precedente, affrontiamo un esempio fondamentale.

Applicazione 0.36 (**Le radici dell'unità**). L'obiettivo è quello di determinare in $\mathbb{C}$ le radici $n-$esime di 1, ossia **tutte** le soluzioni dell'equazione

$$z^n = 1, \qquad z \in \mathbb{C},\ n \in \mathbb{N}\,. \tag{0.35}$$

Se scriviamo z in forma trigonometrica, trovare le soluzioni complesse dell'Equazione 0.35 equivale a trovare i numeri reali ρ, θ per i quali

$$\rho^n\left(\cos(n\theta) + i\sin(n\theta)\right) = 1, \tag{0.36}$$

in base alla formula di De Moivre.

Ora, occorre tenere presente che esistono infiniti argomenti φ equivalenti per il numero reale 1: essi sono tutti i numeri della forma

$$\varphi = 2k\pi \qquad k \in \mathbb{Z}. \tag{0.37}$$

Quindi, dall'Eq. 0.36 ricaviamo subito che

$$\rho = 1, \tag{0.38a}$$

$$\theta = \frac{1}{n}2k\pi, \tag{0.38b}$$

e appare chiaro che la soluzione 0.38b produce n argomenti θ distinti e **non equivalenti**: questi si ottengono dando i valori $k = 0, 1, \dots n - 1$ al parametro k.

In altre parole, per $n = 1$ la radice è 1, per $n = 2$ le radici sono ± 1, e per $n > 2$ le **radici $n-$esime dell'unità** sono i numeri complessi che corrispondono ai **vertici di un poligono regolare di n lati**, inscritto nella circonferenza unitaria centrata in O e con un vertice in $(1, 0)$.

Sulla scorta di questo, è facile mostrare il seguente risultato generale.

Proposizione 0.16. *Sia $w \in \mathbb{C}$ un numero complesso noto; siano $\sigma = |w|$ e $\alpha = \operatorname{Arg} w$ la sua rappresentazione trigonometrica.*

Le radici dell'equazione

$$z^n = w \qquad z \in \mathbb{C}\ n \in \mathbb{N} \tag{0.39}$$

sono date da

$$z_k = \sqrt[n]{\sigma}\left(\cos\frac{\alpha + 2k\pi}{n} + i\sin\frac{\alpha + 2k\pi}{n}\right) = \sqrt[n]{\sigma}\,\mathrm{e}^{i\frac{\alpha+2k\pi}{n}} \tag{0.40}$$

con $k = 1, 2, \dots n - 1$.

Concludiamo con un esempio applicativo:

Esempio 0.37. Trovare le soluzioni dell'equazione

$$(z + i)^4 = i.$$

Cominciamo a risolvere l'equazione ausiliaria

$$w^4 = i.$$

In base alla Proposizione 0.16, otteniamo

$$w_1 = \mathrm{e}^{i\pi/8} \quad w_2 = i\,\mathrm{e}^{i\pi/8} \quad w_3 = -\mathrm{e}^{i\pi/8} \quad w_4 = -i\,\mathrm{e}^{i\pi/8}$$

Ossia,

$$w_1 = \frac{1}{2}\left(\sqrt{2 + \sqrt{2}} + i\sqrt{2 - \sqrt{2}}\right) \tag{0.41}$$

$$w_2 = -\frac{1}{2}\left(\sqrt{2 - \sqrt{2}} - i\sqrt{2 + \sqrt{2}}\right) \tag{0.42}$$

$$w_3 = -\frac{1}{2}\left(\sqrt{2 + \sqrt{2}} + i\sqrt{2 - \sqrt{2}}\right) \tag{0.43}$$

$$w_4 = \frac{1}{2}\left(\sqrt{2 - \sqrt{2}} - i\sqrt{2 + \sqrt{2}}\right), \tag{0.44}$$

da cui ricaviamo $z = w_k - i$ per $k = 1 \dots 4$.

CAPITOLO 1

Vettori applicati e geometria dello spazio.

1. Vettori applicati.

Ogni studente ha incontrato nei propri studi grandezze che non si possono esprimere con un solo numero reale, come ad esempio lo spostamento e la velocità istantanea di un punto che si muove nello spazio, o la forza che agisce su un corpo. Tali grandezze si posssono rappresentare con i vettori. Fissato un punto O nello spazio euclideo $\mathcal{E}$ introduciamo il concetto di vettore applicato in O.

Definizione 1.1. Un *vettore applicato in* O è un segmento orientato $\boldsymbol{v}$ con primo estremo O, se P è il secondo estremo di $\boldsymbol{v}$ scriviamo $\boldsymbol{v} = \overrightarrow{OP}$ (vedasi Figura 1.1).

Indichiamo con $\mathbb{E}^3_O$ *l'insieme dei vettori applicati in* O.

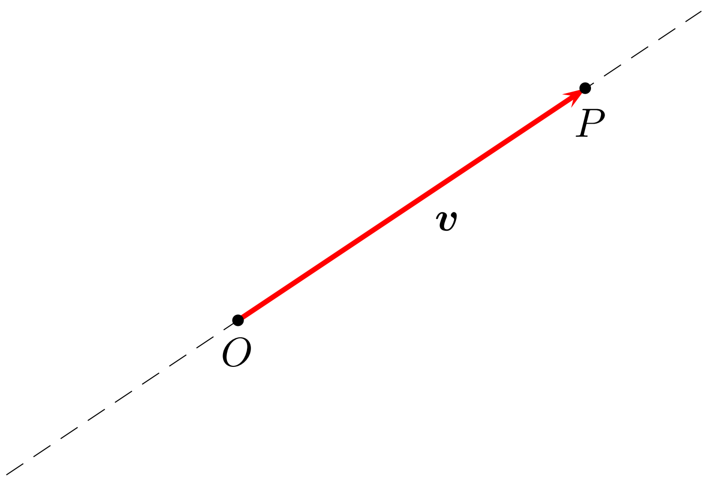

Figura 1.1. Il segmento orientato $\boldsymbol{v} = \overrightarrow{OP}$.

Osservazione 1.1. Fissata un'unità di misura per le lunghezze, ogni vettore $\boldsymbol{v} = \overrightarrow{OP}$, con $P \neq O$, determina univocamente:

(1) una direzione: quella della retta per O e P,
(2) un verso di percorrenza: da O a P,
(3) un numero reale $|\boldsymbol{v}|$, detto *modulo di* $\boldsymbol{v}$: misura della lunghezza del segmento $\overline{OP}$ rispetto all'unità fissata.

Se $P = O$, la direzione ed il verso di $\boldsymbol{v}$ non sono determinati e $|\boldsymbol{v}| = 0$: il vettore $\overrightarrow{OO}$ è il *vettore nullo* di $\mathbb{E}^3_O$, che indicheremo anche con $\mathbf{0}_{E^3_O}$, o anche semplicemente con $\mathbf{0}$.
Infine se $|\boldsymbol{v}| = 1$: $\boldsymbol{v}$ è *un versore* di $\mathbb{E}^3_O$ (convenzionalmente, indicheremo talora con un "cappuccio" i versori: $\hat{\boldsymbol{v}}$).

È noto a tutti come si ottiene la risultante di due forze che agiscono sullo stesso corpo, usando la *regola del parallelogramma*:

Definizione 1.2. Siano $\overrightarrow{OA}$ e $\overrightarrow{OB}$ due vettori applicati in O, la loro *somma* $\overrightarrow{OA}+\overrightarrow{OB}$ è il vettore $\overrightarrow{OC}$, dove C è il secondo estremo del vettore applicato nel punto A, con uguale direzione, verso e modulo del vettore $\overrightarrow{OB}$ (Figura 1.2).

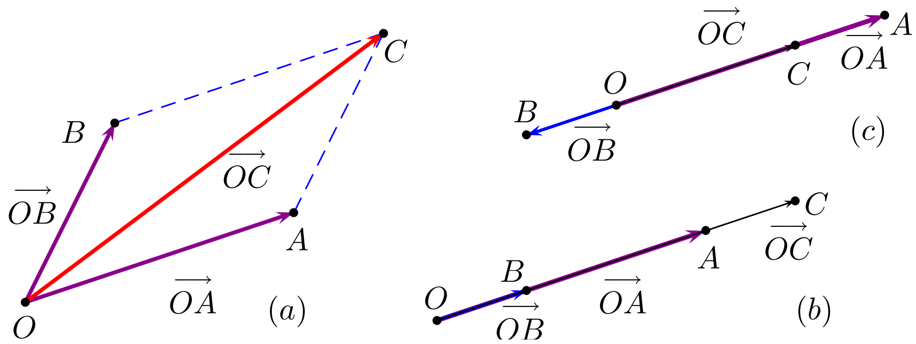

Figura 1.2. La somma di due vettori $\overrightarrow{OA}$ e $\overrightarrow{OB}$ non aventi la stessa direzione si ottiene costruendo il parallelogramma $OACB$ che ha per lati OA e OB; il vettore somma corrisponde a $\overrightarrow{OC}$, ossia alla diagonale del parallelogramma con un estremo in O (a). Se i due vettori hanno la stessa direzione e lo stesso verso, il vettore somma ha la stessa direzione e verso, e modulo $|\overrightarrow{OC}| = |\overrightarrow{OA}| + |\overrightarrow{OB}|$ (b). Se i due vettori hanno la stessa direzione e verso opposto, il vettore somma ha la stessa direzione, modulo $|\overrightarrow{OC}| = ||\overrightarrow{OA}| - |\overrightarrow{OB}||$, e verso uguale a quello dei due di modulo maggiore (c).

Osservazione 1.2. Se i punti A, B, O non sono allineati: C è il quarto vertice del parallelogramma individuato da A, B, O. Applicando le note disuguaglianze triangolari al triangolo di vertici O, A, C si ottiene:

$$\big|\ |\overrightarrow{OA}| - |\overrightarrow{OB}|\ \big| \leq |\overrightarrow{OC}| \leq |\overrightarrow{OA}| + |\overrightarrow{OB}|.$$

Se i punti A, B, O sono allineati: C giace sulla retta contenente i punti e si ha:

(1) se $\overrightarrow{OA}$ e $\overrightarrow{OB}$ hanno lo stesso verso: $\overrightarrow{OC}$ ha il verso di $\overrightarrow{OA}$ e $\overrightarrow{OB}$,

$$|\overrightarrow{OC}| = |\overrightarrow{OA}| + |\overrightarrow{OB}|;$$

(2) se $\overrightarrow{OA}$ e $\overrightarrow{OB}$ hanno versi opposti: $\overrightarrow{OC}$ ha il verso dell'addendo con modulo maggiore e
$$|\overrightarrow{OC}| = |\ \ |\overrightarrow{OA}| - |\overrightarrow{OB}|\ \ |.$$

Proposizione 1.1. *La somma di vettori introdotta gode delle seguenti proprietà:*

(1) *proprietà commutativa: per ogni* $\overrightarrow{OA}, \overrightarrow{OB} \in \mathbb{E}^3_O$
$$\overrightarrow{OA} + \overrightarrow{OB} = \overrightarrow{OB} + \overrightarrow{OA},$$
(2) *proprietà associativa: per ogni* $\overrightarrow{OA}, \overrightarrow{OB}, \overrightarrow{OC} \in \mathbb{E}^3_O$
$$(\overrightarrow{OA} + \overrightarrow{OB}) + \overrightarrow{OC} = \overrightarrow{OA} + (\overrightarrow{OB} + \overrightarrow{OC}),$$
(3) *il vettore nullo è l'elemento neutro: per ogni* $\overrightarrow{OP} \in \mathbb{E}^3_O$
$$\overrightarrow{OP} + \mathbf{0} = \mathbf{0} + \overrightarrow{OP} = \overrightarrow{OP},$$
(4) *ogni vettore* $\overrightarrow{OP} \in \mathbb{E}^3_O$ *ammette opposto:*
$$\overrightarrow{OP} + \overrightarrow{OP'} = \overrightarrow{OP'} + \overrightarrow{OP} = \mathbf{0},$$
dove P' è il simmetrico del punto P rispetto ad O.

Possiamo concludere che $(\mathbb{E}^3_O, +)$ è un gruppo abeliano.

Le proprietà (1), (3) e (4) sono conseguenza immediata della definizione data di somma, la verifica della proprietà (2) richiede invece maggior impegno, ed è lasciata al lettore come esercizio.

Nell'insieme $\mathbb{E}^3_O$ è possibile definire anche un'operazione "esterna", la moltiplicazione di un vettore per un numero reale:

Definizione 1.3. Siano $\alpha \in \mathbb{R}$ e $\boldsymbol{v} \in \mathbb{E}^3_O$, il *prodotto* di α e $\boldsymbol{v}$ è il vettore applicato in O che indicheremo $\alpha\ \boldsymbol{v}$ così definito:

(1) se $\alpha = 0$ o $\boldsymbol{v} = \mathbf{0}$: $\alpha\ \boldsymbol{v} = \mathbf{0}$;
(2) se $\alpha \neq 0$ e $\boldsymbol{v} = \overrightarrow{OP}$: $\alpha\ \overrightarrow{OP} = \overrightarrow{OS}$ dove S appartiene alla retta OP,
$$|\overrightarrow{OS}| = |\alpha|\ \ |\overrightarrow{OP}|;$$
infine il verso di $\overrightarrow{OS}$ è uguale od opposto al verso di $\overrightarrow{OP}$ a seconda del segno di α (si veda per esempio la Figura 1.3).

Proposizione 1.2. *La moltiplicazione di un vettore per un numero reale gode delle seguenti proprietà:*

(1) $\alpha\ (\overrightarrow{OA}+\overrightarrow{OB}) = \alpha\ \overrightarrow{OA}+\alpha\overrightarrow{OB}$, *per ogni* $\overrightarrow{OA}, \overrightarrow{OB} \in \mathbb{E}^3_O$, *per ogni* $\alpha \in \mathbb{R}$;
(2) $(\alpha + \beta)\overrightarrow{OA} = \alpha\ \overrightarrow{OA} + \beta\overrightarrow{OA}$, *per ogni* $\overrightarrow{OA} \in \mathbb{E}^3_O$, *per ogni* $\alpha, \beta \in \mathbb{R}$;
(3) $(\alpha\beta)\overrightarrow{OA} = \alpha(\beta\overrightarrow{OA}) = \beta(\alpha\overrightarrow{OA})$, *per ogni* $\overrightarrow{OA} \in \mathbb{E}^3_O$, *per ogni* $\alpha, \beta \in \mathbb{R}$;
(4) $1\ \overrightarrow{OA} = \overrightarrow{OA}$, *per ogni* $\overrightarrow{OA} \in \mathbb{E}^3_O$.

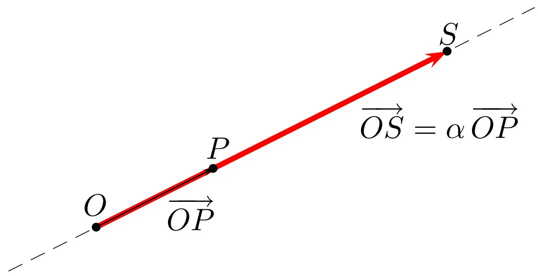

Figura 1.3. Il segmento orientato $\boldsymbol{w} = \overrightarrow{OS} = \alpha\ \overrightarrow{OP} = \alpha\ \boldsymbol{v}$, con $\alpha \in \mathbb{R}$, ha la stessa direzione di $\overrightarrow{OP}$ e verso concorde a quest'ultimo se $\alpha > 0$.

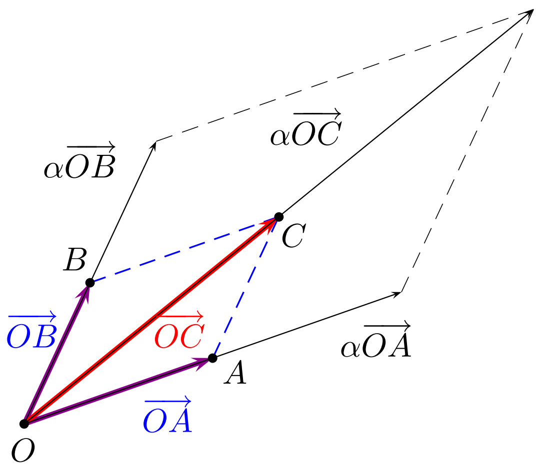

Figura 1.4. $\alpha\ (\overrightarrow{OA} + \overrightarrow{OB}) = \alpha\ \overrightarrow{OA} + \alpha\overrightarrow{OB}$.

Le proprietà (2), (3) e (4) sono di immediata verifica, la proprietà (1) si può verificare sfruttando le proprietà della similitudine (si usi come traccia per la dimostrazione la Figura 1.4).

Esercizio 1.1.
Siano $\boldsymbol{v} \in E^3_O$ e $\alpha \in \mathbb{R}$, verificare che:

$$\alpha \boldsymbol{v} = \mathbf{0} \Longleftrightarrow \alpha = 0 \ \ o \ \ \boldsymbol{v} = \mathbf{0}.$$

Fissato nello spazio euclideo $\mathcal{E}$ un punto O, che chiamiamo *origine*, esiste un' applicazione naturale tra lo spazio $\mathcal{E}$ e lo spazio $\mathbb{E}^3_O$ dei vettori applicati nel punto O:

$$\mathcal{E} \to \mathbb{E}^3_O \tag{1.1}$$

che associa ad ogni punto P il segmento orientato $\overrightarrow{OP}$. Osserviamo che per ogni vettore $\boldsymbol{v} \in \mathbb{E}^3_O$ esiste un unico punto P che gli corrisponde: P è il punto finale di $\boldsymbol{v}$. Otteniamo quindi una corrispondenza biunivoca tra $\mathcal{E}$ e $\mathbb{E}^3_O$, che ci consente di identificare i punti dello spazio con i vettori applicati in O. Introduciamo la seguente notazione:

$$P = O + \boldsymbol{v}$$

per indicare che il punto P corrisponde al vettore $\boldsymbol{v}$ applicato in O.
In generale, dati un punto $P \in \mathcal{E}$ e un vettore $\boldsymbol{v} \in \mathbb{E}^3_O$, indichiamo con

$$P' = P + \boldsymbol{v}$$

il punto finale del segmento orientato con origine P avente direzione, verso e modulo uguali a quelli di $\boldsymbol{v}$. Fissato un vettore $\boldsymbol{v} \in \mathbb{E}^3_O$, consideriamo la trasformazione dello spazio

$$\tau_{\boldsymbol{v}} : \mathcal{E} \to \mathcal{E}$$

che trasforma il punto P nel punto $P' = P + \boldsymbol{v}$.

Definizione 1.1*:* $\tau_{\boldsymbol{v}} : \mathcal{E} \to \mathcal{E}$ è detta *traslazione di vettore* $\boldsymbol{v}$.

Si verifica facilmente che se $\boldsymbol{v}$ è il vettore nullo, $\tau_{\boldsymbol{v}}$ è l'identità, ovvero la trasformazione che trasforma ogni punto in se stesso. Per ogni vettore non nullo $\boldsymbol{v}$, $\tau_{\boldsymbol{v}}$ è biunivoca e la trasformazione inversa (ovvero quella che riporta i punti nella posizione originale) è la traslazione del vettore $-\boldsymbol{v}$.

Esercizio 1.2.
Siano O, A, B tre punti non allineati e sia $\boldsymbol{w} = \overrightarrow{OB} - \overrightarrow{OA}$. Verificare che $B = A + \boldsymbol{w}$. (Figura 1.5).

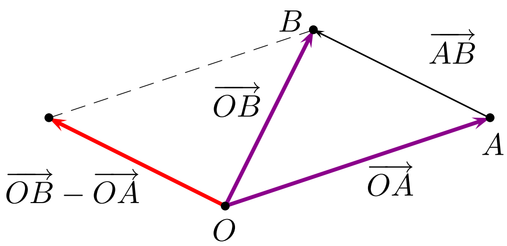

Figura 1.5. Il vettore differenza $\boldsymbol{w} = \overrightarrow{OB} - \overrightarrow{OA}$ si ottiene costruendo il parallelogramma avente per lati i vettori $\overrightarrow{OB}$ e $-\overrightarrow{OA}$.

2. Sistemi di riferimento cartesiani.

Fissato un vettore non nullo $\boldsymbol{u} \in \mathbb{E}^3_O$, possiamo considerare l'insieme di tutti i vettori applicati che si ottengono moltiplicando $\boldsymbol{u}$ per un numero reale, indichiamo tale insieme con

$$\text{Span}(\boldsymbol{u}) = \{\boldsymbol{v} \in \mathbb{E}^3_O \; | \boldsymbol{v} = \alpha \boldsymbol{u}, \; \alpha \in \mathbb{R}\}, \tag{1.2}$$

ci chiediamo quale sia il significato geometrico. Identificando $\mathbb{E}^3_O$ con lo spazio $\mathcal{E}$, Span$(\boldsymbol{u})$ corrisponde al sottoinsieme dei punti $P \in \mathcal{E}$ tali che:

$$P = O + \alpha \boldsymbol{u}, \quad \alpha \in \mathbb{R}.$$

Proprietà 1.3. Span($\boldsymbol{u}$) *rappresenta la retta* r *per l'origine individuata dal vettore* $\boldsymbol{u}$.

Sia $\boldsymbol{v} \in \text{Span}(\boldsymbol{u})$, il punto finale di $\boldsymbol{v}$ è : $P = O + \alpha\boldsymbol{u}$, $\alpha \in \mathbb{R}$. Osserviamo che, per definizione del prodotto $\alpha\boldsymbol{u}$, P è un punto della retta r.

Sia ora P un punto della retta r: proviamo che esiste un unico numero reale α tale che $P = O + \alpha\boldsymbol{u}$, cioè $\overrightarrow{OP} = \alpha\boldsymbol{u}$. Osserviamo che:

$$|\alpha|\ |\boldsymbol{u}| = |\overrightarrow{OP}| \Leftrightarrow |\alpha| = \frac{|\overrightarrow{OP}|}{|\boldsymbol{u}|},$$

ed il segno di α è positivo o negativo a seconda se i vettori $|\overrightarrow{OP}|$ e $\boldsymbol{u}$ hanno lo stesso verso oppure versi opposti.

Nel seguito, diremo brevemente che Span($\boldsymbol{u}$) è la retta passante per l'origine generata dal vettore non nullo $\boldsymbol{u}$. Osserviamo infine che, viceversa, ogni retta r passante per l'origine O, può essere identificata con Span($\boldsymbol{u}$), per un vettore non nullo $\boldsymbol{u} = \overrightarrow{OP}$, $P \in r$.

Dalla proprietà precedente (1.3) si ha che se O, A, B sono punti distinti, essi sono allineati se e solo se $\text{Span}(\overrightarrow{OB}) = \text{Span}(\overrightarrow{OA})$, se e solo se $\overrightarrow{OB} \in \text{Span}(\overrightarrow{OA})$. Traduciamo questa proprietà geometrica in una proprieà algebrica sui vettori applicati:

Definizione 1.4. Due vettori $\boldsymbol{u}$ e $\boldsymbol{v}$ di $\mathbb{E}^3_O$ sono detti:

- *linearmente dipendenti* se $\boldsymbol{v} \in \text{Span}(\boldsymbol{u})$,
- *linearmente indipendenti* se $\boldsymbol{v} \notin \text{Span}(\boldsymbol{u})$.

Siano ora $\{\boldsymbol{u}, \boldsymbol{v}\} \in \mathbb{E}^3_O$ due vettori linearmente indipendenti, allora Span($\boldsymbol{u}$) e Span($\boldsymbol{v}$) rappresentano due rette distinte e incidenti in O, indichiamo con π il piano che le contiene. Consideriamo il sottoinsieme di $\mathbb{E}^3_O$ dato da tutti i vettori ottenuti sommando un vettore di Span($\boldsymbol{u}$) e un vettore di Span($\boldsymbol{v}$), che indichiamo con

$$\text{Span}(\boldsymbol{u}, \boldsymbol{v}) = \{\boldsymbol{w} \in \mathbb{E}^3_O |\ \boldsymbol{w} = \alpha\boldsymbol{u} + \beta\boldsymbol{v}, \quad \alpha, \beta \in \mathbb{R}\ \}, \tag{1.3}$$

ci chiediamo quale sia il significato geometrico di tale insieme. Identificando lo spazio $\mathbb{E}^3_O$ con $\mathcal{E}$, Span($\boldsymbol{u}, \boldsymbol{v}$) corrisponde al sottoinsieme dei punti $P \in \mathcal{E}$ tali che:

$$P = O + \alpha\boldsymbol{u} + \beta\boldsymbol{v}, \quad \alpha, \beta \in \mathbb{R}.$$

Proprietà 1.4. Span($\boldsymbol{u}, \boldsymbol{v}$) *rappresenta il piano* π *per l'origine individuato dalle rette* Span($\boldsymbol{u}$) *e* Span($\boldsymbol{v}$).

Sia $\boldsymbol{w} \in \text{Span}(\boldsymbol{u}, \boldsymbol{v})$, il punto finale di $\boldsymbol{w}$ è : $P = O + \alpha\boldsymbol{u} + \beta\boldsymbol{v}$, $\alpha, \beta \in \mathbb{R}$. Per definizione, P è il punto finale della diagonale per O del parallelogrammo di lati $\alpha\boldsymbol{u}$ e $\beta\boldsymbol{v}$, quindi è un punto del piano π che contiene le rette Span($\boldsymbol{u}$) e Span($\boldsymbol{v}$).

Sia ora $P \in \pi$, proviamo che P è il vertice di un unico parallelogramma avente un vertice in O e due lati rispettivamente sulle rette Span($\boldsymbol{u}$) e Span($\boldsymbol{v}$): basta tracciare le due rette nel piano π passanti per P e parallele alle rette

Span($\boldsymbol{u}$) e Span($\boldsymbol{v}$). Detti A e B rispettivamente i punti di intersezione di tali rette con Span($\boldsymbol{u}$) e Span($\boldsymbol{v}$), si ha:

$$\overrightarrow{OP} = \overrightarrow{OA} + \overrightarrow{OB}, \quad \overrightarrow{OA} \in \mathrm{Span}(\boldsymbol{u}) \quad \overrightarrow{OB} \in \mathrm{Span}(\boldsymbol{v}),$$

quindi esistono due numeri reali α e β tali che $\overrightarrow{OP} = \alpha\boldsymbol{u} + \beta\boldsymbol{v}$. (Figura 1.6).

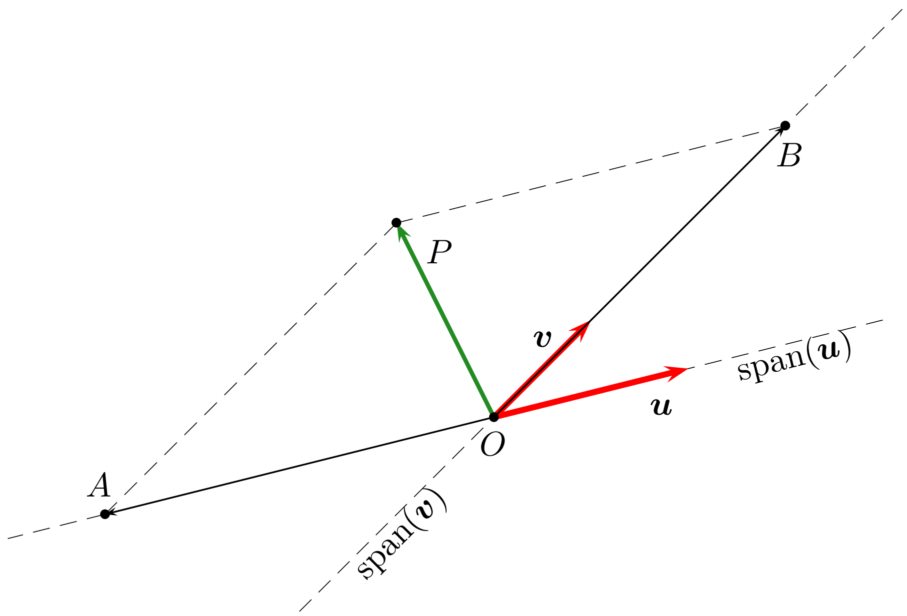

FIGURA 1.6. Preso un punto P qualunque nel piano π, è sempre possibile decomporre il vettore $\overrightarrow{OP}$ in un vettore $\overrightarrow{OA} \in \mathrm{Span}(\boldsymbol{u})$ ed un altro $\overrightarrow{OB} \in \mathrm{Span}(\boldsymbol{v})$.

In seguito, diremo brevemente che Span($\boldsymbol{u}, \boldsymbol{v}$) è il piano generato dai vettori linearmente indipendenti $\boldsymbol{u}$ e $\boldsymbol{v}$. Osserviamo infine che, viceversa, ogni piano π passante per l'origine O, può essere identificato con Span($\boldsymbol{u}, \boldsymbol{v}$), per una coppia di vettori linearmente indipendenti $\boldsymbol{u}$ e $\boldsymbol{v}$.

OSSERVAZIONE 1.3. Siano $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{E}^3_O$ due vettori linearmente indipendenti, dalla proprietà precedente segue che ogni vettore $\boldsymbol{w} \in \mathrm{Span}(\boldsymbol{u}, \boldsymbol{v})$ si può scrivere in un unico modo come somma di un vettore di Span($\boldsymbol{u}$) e un vettore di Span($\boldsymbol{v}$):

$$\boldsymbol{w} = \alpha\boldsymbol{u} + \beta\boldsymbol{v}, \quad \alpha, \beta \in \mathbb{R}.$$

Diciamo per questo che $\{\boldsymbol{u}, \boldsymbol{v}\}$ costituiscono una *base* $\mathcal{B}$ di Span($\boldsymbol{u}, \boldsymbol{v}$). Fissata una base $\mathcal{B}$ possiamo associare ad ogni vettore $\boldsymbol{w} \in \mathrm{Span}(\boldsymbol{u}, \boldsymbol{v})$ un'unica coppia di numeri reali, (α, β), detti *coordinate* del vettore $\overrightarrow{OP}$ rispetto alla base fissata. Le coordinate identificano in modo univoco il vettore $\boldsymbol{w}$, infatti conoscendo i valori α e β posso ricostruire il vettore $\boldsymbol{w}$.

Sia π un piano passante per l'origine O, fissiamo due rette ortogonali per O, scegliamo infine due versori $\hat{\boldsymbol{\imath}}$ e $\hat{\boldsymbol{\jmath}}$ su di esse: possiamo identificare il piano π con Span($\hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}$). La base $\mathcal{B} = \{\hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}\}$ è detta *base ortonormale* per il piano

Span$(\hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}})$. Abbiamo in questo modo fissato un *sistema di riferimento cartesiano ortonormale nel piano.* Per ogni punto P del piano si ha $P = O + x\hat{\boldsymbol{\imath}} + y\hat{\boldsymbol{\jmath}}$, i numeri reali x e y, univocamente determinati, sono detti *coordinate cartesiane del punto* P sul piano π nel sistema di riferimento cartesiano ortonormale fissato.

DEFINIZIONE 1.5. $\mathcal{R}(O, \hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}})$ *indica un sistema di riferimento cartesiano ortonormale nel piano.*

ESERCIZIO 2.1. Siano $\boldsymbol{u}$,$\boldsymbol{v}$ due vettori linearmente indipendenti di $\mathbb{E}^3_O$ e sia $\boldsymbol{w} \in \mathbb{E}^3_O$. Verificare che i vettori $\boldsymbol{u}$,$\boldsymbol{v}$ e $\boldsymbol{w}$ sono complanari (cioè esiste un piano che li contiene) se e solo se $\boldsymbol{w} \in \mathrm{Span}(\boldsymbol{u}, \boldsymbol{v})$.

Siano $\boldsymbol{u}$,$\boldsymbol{v}$ e $\boldsymbol{w}$ tre vettori non complanari di $\mathbb{E}^3_O$. Consideriamo il sottoinsieme di $\mathbb{E}^3_O$ dato dai vettori applicati che sono somma di vettori rispettivamente di Span$(\boldsymbol{u})$, Span$(\boldsymbol{v})$ e Span$(\boldsymbol{w})$:

$$\mathrm{Span}(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}) = \{\boldsymbol{z} \in \mathbb{E}^3_O \colon \boldsymbol{z} = \alpha\boldsymbol{u} + \beta\boldsymbol{v} + \gamma\boldsymbol{w}, \quad \alpha, \beta, \gamma \in \mathbb{R}\}, \tag{1.4}$$

vogliamo determinare il significato geometrico di tale insieme. A tale scopo proviamo la seguente:

PROPRIETÀ 1.5. *Ogni vettore* $\overrightarrow{OP} \in \mathbb{E}^3_O$ *ammette un'unica scrittura del seguente tipo:*

$$\overrightarrow{OP} = \overrightarrow{OA} + \overrightarrow{OB}, \quad \overrightarrow{OA} \in \mathrm{Span}(\boldsymbol{w}) \quad \overrightarrow{OB} \in \mathrm{Span}(\boldsymbol{u}, \boldsymbol{v}).$$

Osserviamo che la proprietà è vera se $\overrightarrow{OP} \in \mathrm{Span}(\boldsymbol{u}, \boldsymbol{v})$: in tal caso si ha $\overrightarrow{OA} = \mathbf{0}$ e $B = P$. Supponiamo ora che P non appartenga al piano individuato da Span$(\boldsymbol{u}, \boldsymbol{v})$. Proviamo che P è vertice di un unico parallelogrammma avente un lato $\overrightarrow{OA} \in \mathrm{Span}(\boldsymbol{w})$ ed un lato $\overrightarrow{OB} \in \mathrm{Span}(\boldsymbol{u}, \boldsymbol{v})$: basta tracciare la retta per P parallela alla retta Span$(\boldsymbol{w})$ ed il piano per P parallelo al piano Span$(\boldsymbol{u}, \boldsymbol{v})$. Detti A e B rispettivamente i punti di intersezione del piano tracciato con la retta Span$(\boldsymbol{w})$ e della retta tracciata con Span$(\boldsymbol{u}, \boldsymbol{v})$ (vedi Figura 1.7) si ottiene il parallelogramma O, A, P, B. Otteniamo quindi:

$$\overrightarrow{OP} = \overrightarrow{OA} + \overrightarrow{OB} \quad \overrightarrow{OA} \in \mathrm{Span}(\boldsymbol{w}) \quad \overrightarrow{OB} \in \mathrm{Span}(\boldsymbol{u}, \boldsymbol{v}).$$

Dall'unicità della costruzione del parallelogramma O, A, P, B e poiché $\overrightarrow{OB}$ si scrive in un unico modo come somma di vettori di Span$(\boldsymbol{u})$ e Span$(\boldsymbol{v})$, otteniamo la seguente:

PROPRIETÀ 1.6. *Per ogni* $\boldsymbol{z} \in \mathbb{E}^3_O$, *sono univocamente determinati tre numeri reali* α, β, γ *tali che:*

$$\boldsymbol{z} = \alpha\boldsymbol{u} + \beta\boldsymbol{v} + \gamma\boldsymbol{w}.$$

Si ha quindi $\mathrm{Span}(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}) = \mathbb{E}^3_O$.

OSSERVAZIONE 1.4. Abbiamo quindi provato che una terna di vettori non complanari $\{\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}\}$ di $\mathbb{E}^3_O$ costituisce una *base* $\mathcal{B}$ di $\mathbb{E}^3_O$. I numeri reali (α, β, γ) sono detti *coordinate del vettore* $\boldsymbol{z}$ rispetto alla base fissata.

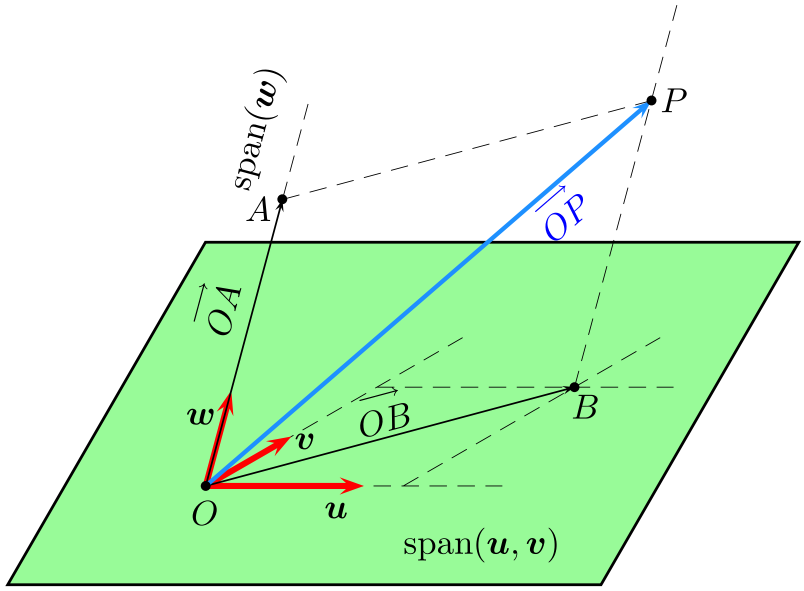

FIGURA 1.7. La costruzione per decomporre ogni vettore $\overrightarrow{OP} \in \mathbb{E}^3_O$ in un vettore in Span($\boldsymbol{w}$) ed uno in Span($\boldsymbol{u}, \boldsymbol{v}$).

Fissate nello spazio tre rette per O, mutuamente ortogonali e tre versori $\hat{\boldsymbol{\imath}}$, $\hat{\boldsymbol{\jmath}}$ e $\hat{\boldsymbol{k}}$ su di esse, la base $\mathcal{B} = \{\hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}, \hat{\boldsymbol{k}}\}$ è detta *base ortonormale* di $\mathbb{E}^3_O$. Ad ogni punto P associamo una terna , univocamente determinata, di numeri reali (x, y, z), tali che

$$P = O + x\hat{\boldsymbol{\imath}} + y\hat{\boldsymbol{\jmath}} + z\hat{\boldsymbol{k}},$$

dette *coordinate* del punto P nel sistema di riferimento cartesiano *ortonormale* fissato.

DEFINIZIONE 1.6. *$\mathcal{R}(O, \hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}, \hat{\boldsymbol{k}})$ è un sistema di riferimento cartesiano ortonormale nello spazio.*

NOTAZIONE 1.7. Fissato nello spazio un sistema di riferimento cartesiano ortonormale $\mathcal{R}(O, \hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}, \hat{\boldsymbol{k}})$:

- le rette orientate Span($\hat{\boldsymbol{\imath}}$), Span($\hat{\boldsymbol{\jmath}}$) e Span($\hat{\boldsymbol{k}}$) sono gli assi coordinati;
- i piani Span($\hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}$), Span($\hat{\boldsymbol{\imath}}, \hat{\boldsymbol{k}}$) e Span($\hat{\boldsymbol{\jmath}}, \hat{\boldsymbol{k}}$) sono i piani coordinati;
- $P = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ indica il punto P di coordinate (x, y, z) in $\mathcal{R}(O, \hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}, \hat{\boldsymbol{k}})$:

$$P = O + x\hat{\boldsymbol{\imath}} + y\hat{\boldsymbol{\jmath}} + z\hat{\boldsymbol{k}};$$

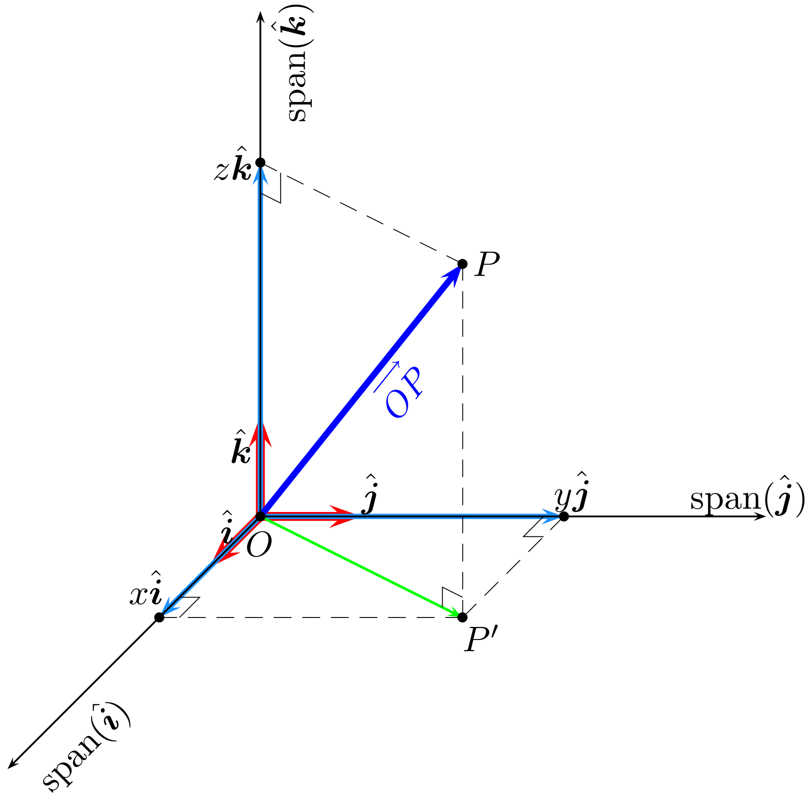

FIGURA 1.8. Il riferimento cartesiano ortogonale monometrico $\mathcal{R}(O, \hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}, \hat{\boldsymbol{k}})$.

- $\overrightarrow{OP} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ indica il vettore applicato in O di coordinate (x, y, z) rispetto alla base $\mathcal{B} = \{\hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}, \hat{\boldsymbol{k}}\}$:

$$\overrightarrow{OP} = x\hat{\boldsymbol{\imath}} + y\hat{\boldsymbol{\jmath}} + z\hat{\boldsymbol{k}}.$$

OSSERVAZIONE 1.5. Sia $\mathcal{B} = \{\hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}, \hat{\boldsymbol{k}}\}$ una base ortonormale di $\mathbb{E}^3_O$. Consideriamo i seguenti vettori $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{E}^3_O$:

$$\boldsymbol{u} = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} \quad \boldsymbol{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}.$$

Proviamo che le coordinate dei vettori $\boldsymbol{u} + \boldsymbol{v}$ e $\alpha\boldsymbol{v}$ si ottengono rispettivamente sommando le coordinate di $\boldsymbol{u}$ e $\boldsymbol{v}$ e moltiplicando le coordinate di $\boldsymbol{v}$ per α:

$$\boldsymbol{u} + \boldsymbol{v} = \begin{pmatrix} u_1 + v_1 \\ u_2 + v_2 \\ u_3 + v_3 \end{pmatrix} \quad \alpha\boldsymbol{v} = \begin{pmatrix} \alpha v_1 \\ \alpha v_2 \\ \alpha v_3 \end{pmatrix}.$$

Ricordiamo che per definizione abbiamo

$$\boldsymbol{u} = u_1\hat{\boldsymbol{\imath}} + u_2\hat{\boldsymbol{\jmath}} + u_3\hat{\boldsymbol{k}}, \quad \boldsymbol{v} = v_1\hat{\boldsymbol{\imath}} + v_2\hat{\boldsymbol{\jmath}} + v_3\hat{\boldsymbol{k}};$$

sommando i vettori $\boldsymbol{u}$ e $\boldsymbol{v}$ ed eseguendo la moltiplicazione di α e $\boldsymbol{v}$, applicando le proprietà delle operazioni della somma e del prodotto per un numero reale, otteniamo:

$$\boldsymbol{u}+\boldsymbol{v}=(u_1+v_1)\hat{\boldsymbol{\imath}}+(u_2+v_2)\hat{\boldsymbol{\jmath}}+(u_3+v_3)\hat{\boldsymbol{k}},$$
$$\alpha\boldsymbol{v}=\alpha(v_1\hat{\boldsymbol{\imath}}+v_2\hat{\boldsymbol{\jmath}}+v_3\hat{\boldsymbol{k}})=(\alpha v_1)\hat{\boldsymbol{\imath}}+(\alpha v_2)\hat{\boldsymbol{\jmath}}+(\alpha v_3)\hat{\boldsymbol{k}}.$$

Per l'unicità della scrittura di ogni vettore in una base fissata, queste sono le decomposizioni rispettivamente di $\boldsymbol{u}+\boldsymbol{v}$ e $\alpha\boldsymbol{v}$ nella base $\mathcal{B}$, da cui possiamo leggere le coordinate rispettivamente di $\boldsymbol{u}+\boldsymbol{v}$ e $\alpha\boldsymbol{v}$ nella base $\mathcal{B}$.

Fissata una base $\mathcal{B}$ in $\mathbb{E}^3_O$, le coordinate (x,y,z) di un vettore $\boldsymbol{v}$ identificano in modo univoco il vettore, cioè conoscendo le coordinate posso ricostruire il vettore $\boldsymbol{v}$. Inoltre le proprietà che abbiamo verificato ora, ci consentono di eseguire le operazioni tra i vettori di $\mathbb{E}^3_O$ direttamente sulle loro coordinate, dove tali operazioni sono più semplici da eseguire.

Esercizio 2.2. Siano $\boldsymbol{v}\in\mathbb{E}^3_O$ e $\tau_{\boldsymbol{v}}\colon\mathcal{E}\to\mathcal{E}$ la traslazione di vettore $\boldsymbol{v}$, che trasforma il punto P nel punto $P'=P+\boldsymbol{v}$.
Verificare che, per ogni punto $P\in\mathcal{E}$, risulta: $\overrightarrow{OP'}=\overrightarrow{OP}+\boldsymbol{v}$.
Fissato nello spazio un sistema di riferimento $\mathcal{R}(O,\hat{\boldsymbol{\imath}},\hat{\boldsymbol{\jmath}},\hat{\boldsymbol{k}})$ cartesiano ortonormale, siano

$$\boldsymbol{v}=\begin{pmatrix}v_1\\v_2\\v_3\end{pmatrix}\quad P=\begin{pmatrix}x\\y\\z\end{pmatrix},$$

verificare che le coordiante cartesiane del punto $P'=\tau_{\boldsymbol{v}}(P)$ sono le seguenti:

$$\begin{pmatrix}x'\\y'\\z'\end{pmatrix}=\begin{pmatrix}x+v_1\\y+v_2\\z+v_3\end{pmatrix}.$$

3. Rappresentazioni parametriche di rette e piani.

Iniziamo a studiare la geometria dello spazio in cui abbiamo fissato un sistema di riferimento cartesiano ortonormale $\mathcal{R}(O,\hat{\boldsymbol{\imath}},\hat{\boldsymbol{\jmath}},\hat{\boldsymbol{k}})$.

Una retta r è univocamente determinata assegnando un punto $P_0\in r$ ed un vettore non nullo $\boldsymbol{v}\in\mathbb{E}^3_O$ parallelo ad r, che dà la direzione di r:

Definizione 1.8. $\boldsymbol{v}$ è un *vettore direttore* di r (Figura 1.9).

Osservazione 1.6. Osserviamo che ogni vettore non nullo $\boldsymbol{w}\in\operatorname{Span}(\boldsymbol{v})$ è un vettore direttore di r.

Proprietà 1.7. *La retta r è l'insieme di tutti e soli i punti P che soddisfano la seguente equazione parametrica vettoriale:*

$$P=P_0+t\,\boldsymbol{v},\quad t\in\mathbb{R}. \tag{1.5}$$

Se la retta r passa per l'origine O del riferimento, possiamo scegliere $P_0=O$, abbiamo visto nella lezione precedente che $r=\operatorname{Span}(\boldsymbol{v})$, quindi è il luogo dei punti P tali che

$$P=t\,\boldsymbol{v},\quad t\in\mathbb{R}.$$

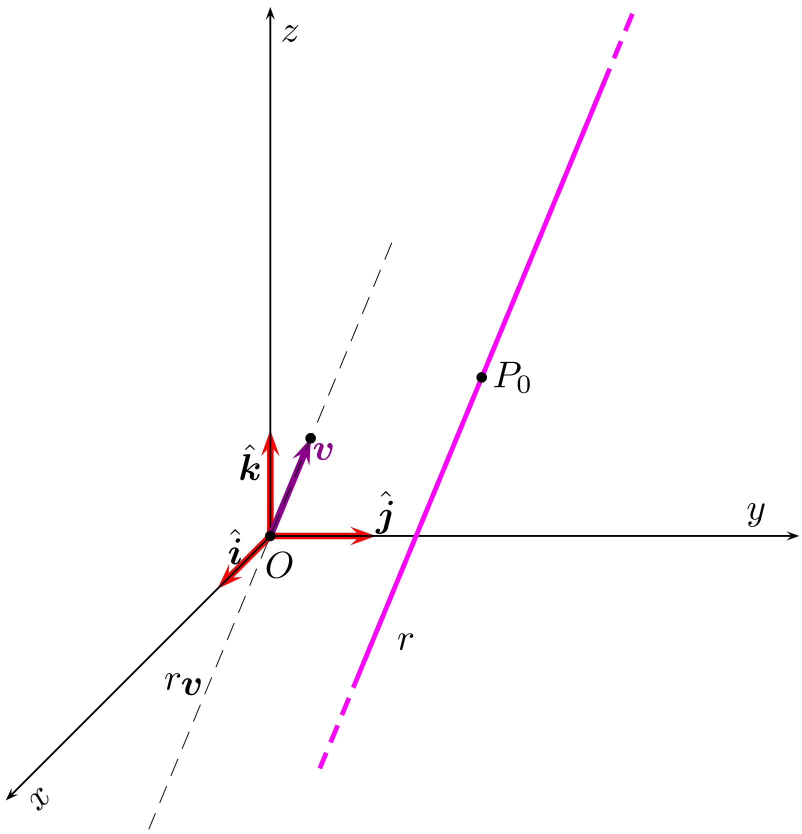

FIGURA 1.9. $\boldsymbol{v} \in \mathbb{E}^3_O$ è un *vettore direttore* della retta r se $r_{\boldsymbol{v}} = \mathrm{Span}(\boldsymbol{v}) \parallel r$.

Se la retta r non passa per l'origine, questa scelta non è possibile. Osserviamo che r risulta essere la traslazione di $\mathrm{Span}(\boldsymbol{v})$ del vettore $\overrightarrow{OP_0}$: infatti un punto $P \in r$ se e solo si ha (vedi Figura 1.10):

$$P = P_0 + \overrightarrow{OQ}, \quad \overrightarrow{OQ} \in \mathrm{Span}(\boldsymbol{v}),$$

che conclude la verifica.

OSSERVAZIONE 1.7. L'equazione parametrica di r consiste nel descrivere i punti della retta in funzione di un parametro reale a partire da un punto e da un vettore direttore, per questo motivo non è univocamente determinata. Infatti l'equazione seguente

$$P = P_1 + s\,\boldsymbol{w}, \quad s \in \mathbb{R}$$

rappresenta r per ogni punto $P_1 \in r$ e per ogni vettore non nullo $\boldsymbol{w} \in \mathrm{Span}(\boldsymbol{v})$. Esistono quindi infinite rappresentazioni parametriche di r.

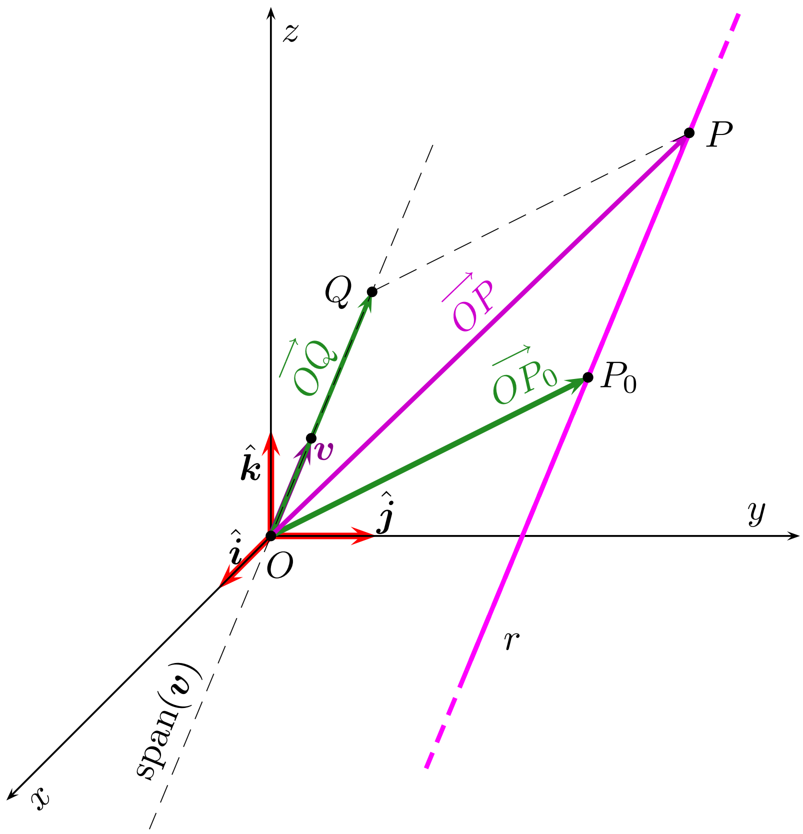

FIGURA 1.10. La retta r si ottiene traslando di un vettore $\overrightarrow{OP_0}$ la retta Span($\boldsymbol{v}$) passante per l'origine, con $P_0 \in r$, dove $\boldsymbol{v}$ è un direttore della retta. Preso, infatti un punto P qualunque su r, basta costruire il parallelogramma OP_0PQ che ha OP come diagonale ed OP_0 come lato; per costruzione, il lato OQ è parallelo a PP_0, dunque giace sulla retta passante per l'origine diretta come $\boldsymbol{v}$, ossia $\overrightarrow{OQ} \in \text{Span}(\boldsymbol{v})$.

Scriviamo ora le coordinate dei vettori coinvolti nell'equazione parametrica trovata. Siano

$$\boldsymbol{v} = \begin{pmatrix} l \\ m \\ n \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \quad P_0 = \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix},$$

indichiamo con $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ le coordinate di un punto $P \in r$, sostituendo nell'equazione precedente otteniamo

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix} + t \begin{pmatrix} l \\ m \\ n \end{pmatrix}. \tag{1.6}$$

Possiamo concludere che il punto P appartiene alla retta r se e solo se le sue coordinate sono le seguenti, al variare del parametro reale t:

$$\begin{cases} x = x_0 + t\, l, \\ y = y_0 + t\, m, \\ z = z_0 + t\, n, \qquad t \in \mathbb{R}. \end{cases} \tag{1.7}$$

Abbiamo ottenuto le *equazioni parametriche della retta* r.

APPLICAZIONE 1.8. *Retta r per due punti distinti A e B.*
Osserviamo che il vettore $\overrightarrow{OB} - \overrightarrow{OA}$ è non nullo ed è parallelo alla retta r, (vedi Esercizio 1.2), la retta r è quindi la retta per il punto A avente come direzione il vettore $\overrightarrow{OB} - \overrightarrow{OA}$, ha quindi equazione parametrica:

$$P = A + t(\overrightarrow{OB} - \overrightarrow{OA}), \quad t \in \mathbb{R}.$$

ESEMPIO 1.9. Fissiamo nello spazio un sistema di riferimento cartesiano ortonornormale $\mathcal{R}(O, \hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}, \hat{\boldsymbol{k}})$.

(1) Troviamo una rappresentazione parametrica per l'asse x del sistema di riferimento. L'asse x è la retta $\mathrm{Span}(\hat{\boldsymbol{\imath}})$, quindi ha equazione parametrica:
$$P = t\hat{\boldsymbol{\imath}}, \quad t \in \mathbb{R}.$$
Passando alle coordinate abbiamo:
$$\begin{cases} x = t, \\ y = 0, \\ z = 0, \qquad t \in \mathbb{R}. \end{cases}$$

(2) Troviamo una rappresentazione parametrica per la retta r passante per il punto $A = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$ e avente come direzione il vettore $\boldsymbol{v} = 2\hat{\boldsymbol{\imath}} - \hat{\boldsymbol{\jmath}}$.
La retta ha equazione parametrica:
$$P = A + t\,\boldsymbol{v}, \quad t \in \mathbb{R},$$
con $\boldsymbol{v} = \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix}$. Sostituendo e passando alle coordinate si ha:
$$\begin{cases} x = 1 + 2t, \\ y = 1 - t, \\ z = -1, \qquad t \in \mathbb{R}. \end{cases}$$

(3) Troviamo una rappresentazione parametrica per la retta r passamte per i punti $A = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$ e $B = \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix}$. La retta ha equazione parametrica:
$$P = A + t(\overrightarrow{OB} - \overrightarrow{OA}), \quad t \in \mathbb{R}.$$

Determiniamo le coordinate del vettore $\overrightarrow{OB} - \overrightarrow{OA}$:

$$\overrightarrow{OB} - \overrightarrow{OA} = \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} = \begin{pmatrix} -2 \\ 1 \\ 2 \end{pmatrix},$$

sostituendo e passando alle coordinate si ha:

$$\begin{cases} x = 1 - 2t, \\ y = 1 + t, \\ z = -1 + 2t, \qquad t \in \mathbb{R}. \end{cases}$$

Vogliamo ora determinare una rappresentazione parametrica per i piani nello spazio in cui abbiamo fissato un sistema di riferimento cartesiano ortonormale $\mathcal{R}(O, \hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}, \hat{\boldsymbol{k}})$.
Un piano π è univocamente determinato assegnando un punto $P_0 \in \pi$ ed una coppia di vettori linearmente indipendenti $\{\boldsymbol{u}, \boldsymbol{v}\} \in \mathbb{E}^3_O$ paralleli ad r:

Definizione 1.9. $\{\boldsymbol{u}, \boldsymbol{v}\}$ costituiscono una *giacitura* di π.

Osservazione 1.10. Osserviamo che per ogni coppia di vettori $\{\boldsymbol{u}', \boldsymbol{v}'\}$ tale che $\mathrm{Span}(\boldsymbol{u}', \boldsymbol{v}') = \mathrm{Span}(\boldsymbol{u}, \boldsymbol{v})$, anche $\{\boldsymbol{u}', \boldsymbol{v}'\}$ sono una giacitura di π.

Proprietà 1.8. *Il piano π è l'insieme di tutti e soli i punti P che soddisfano la seguente equazione parametrica vettoriale:*

$$P = P_0 + t\ \boldsymbol{u} + s\ \boldsymbol{v}, \quad t, s \in \mathbb{R}. \tag{1.8}$$

Se il piano π passa per l'origine O del riferimento, possiamo scegliere $P_0 = O$, abbiamo visto nella lezione precedente che $\pi = \mathrm{Span}(\boldsymbol{u}, \boldsymbol{v})$, quindi è il luogo dei punti P tali che

$$P = t\ \boldsymbol{u} + s\ \boldsymbol{v}, \quad t, s \in \mathbb{R}.$$

Se il piano π non passa per l'origine, tale scelta non è possibile. osserviamo che π risulta essere la traslazione di $\mathrm{Span}(\boldsymbol{u}, \boldsymbol{v})$ del vettore $\overrightarrow{OP_0}$: infatti un punto $P \in \pi$ se e solo si ha (vedi Figura 1.12):

$$P = P_0 + \overrightarrow{OQ}, \quad \overrightarrow{OQ} \in \mathrm{Span}(\boldsymbol{u}, \boldsymbol{v}),$$

che conclude la verifica.

Osservazione 1.11. L'equazione parametrica di π consiste nel descrivere i punti del piano in funzione di due parametri reali a partire da un punto e una coppia di vettori linearmente indipendenti che danno una giacitura del piano, perciò non è univocamente determinata. Infatti, l'equazione seguente:

$$P = P_1 + t'\boldsymbol{w}_1 + s'\boldsymbol{w}_2, \quad t', s' \in \mathbb{R}$$

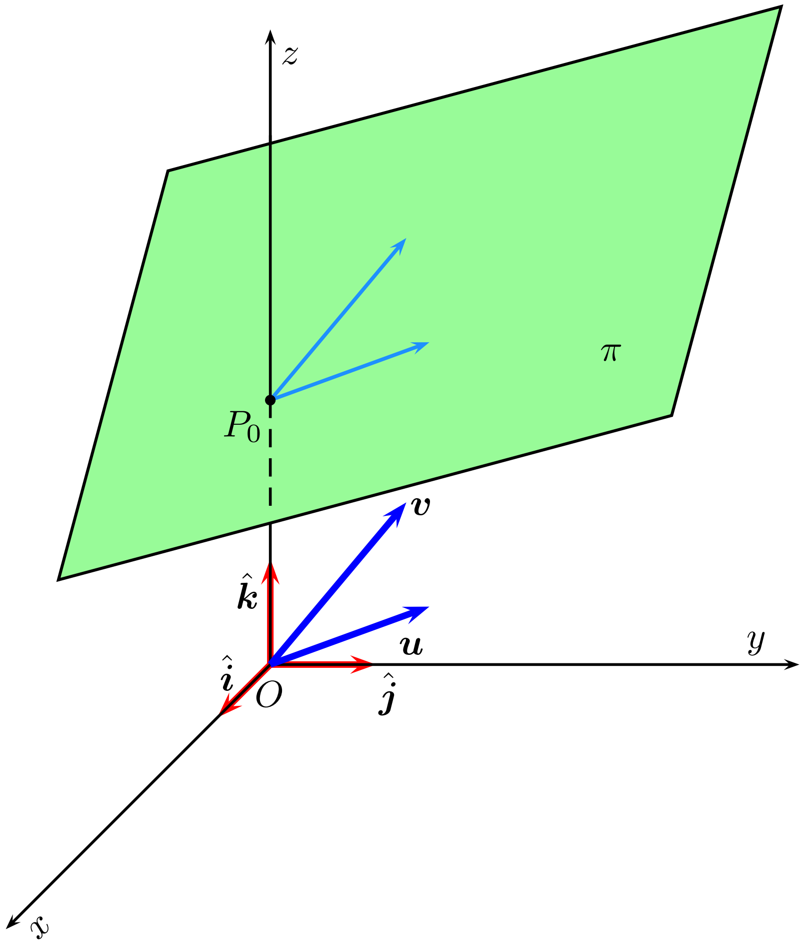

FIGURA 1.11. Un piano π è univocamente individuato da un suo punto P_0 e da una coppia di vettori $\{\boldsymbol{u}, \boldsymbol{v}\} \in \mathbb{E}^3_O$ linearmente indipendenti paralleli ad esso (ossia, tali che $\mathrm{Span}(\boldsymbol{u}, \boldsymbol{v}) \parallel \pi$).

rappresenta π per ogni punto $P_1 \in \pi$ e per ogni coppia di vettore linearmente indipendenti $\{\boldsymbol{w}_1, \boldsymbol{w}_2\}$ tali che

$$\mathrm{Span}(\boldsymbol{u}, \boldsymbol{v}) = \mathrm{Span}(\boldsymbol{w}_1, \boldsymbol{w}_2).$$

Esistono quindi infinite rappresentazioni parametriche di π.

Scriviamo ora le coordinate dei vettori coinvolti nell'equazione parametrica trovata. Siano

$$\boldsymbol{u} = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}, \quad \boldsymbol{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}, \quad P_0 = \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix},$$

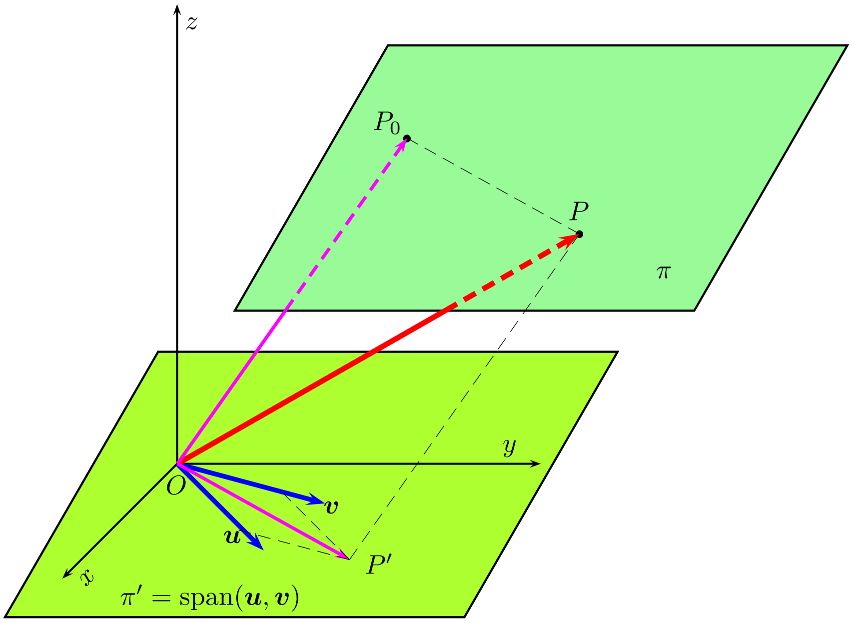

FIGURA 1.12. Qualunque sia il punto P nel piano π ottenuto traslando il piano $\pi' = \mathrm{Span}(\boldsymbol{u}, \boldsymbol{v})$ del vettore $\overrightarrow{OP_0}$, il vettore $\overrightarrow{OP}$ è sempre decomponibile nella somma di $\overrightarrow{OP_0}$ con un vettore $\overrightarrow{OP'}$ in π'.

indichiamo con $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ le coordinate di un punto $P \in \pi$, sostituendo nell'equazione precedente otteniamo

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix} + t \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} + s \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}, \quad t, s \in \mathbb{R}. \tag{1.9}$$

Possiamo concludere che il punto P appartiene al piano π se e solo se le sue coordinate sono le seguenti, al variare dei parametri reali t e s:

$$\begin{cases} x = x_0 + tu_1 + sv_1, \\ y = y_0 + tu_2 + sv_2, \\ z = z_0 + tu_3 + sv_3. \end{cases} \qquad t, s \in \mathbb{R} \tag{1.10}$$

Abbiamo ottenuto le *equazioni parametriche del piano* π.

APPLICAZIONE 1.12. *Piano per tre punti non allineati A, B e C.*
Consideriamo tre punti distinti A, B e C: osserviamo che
i punti A, B e C sono allineati se e solo se le rette AB e AC coincidono $\Longleftrightarrow$ le rette AB e AC hanno la stessa direzione.
Poiché la direzione della retta AB è data dal vettore $\overrightarrow{OB} - \overrightarrow{OA}$, e la direzione della retta AC è data dal vettore $\overrightarrow{OC} - \overrightarrow{OA}$, abbiamo che

i punti A, B e C sono allineati se e solo se i vettori $\overrightarrow{OB}-\overrightarrow{OA}$ e $\overrightarrow{OC}-\overrightarrow{OA}$ sono linearmente dipendenti se e solos se $\overrightarrow{OB}-\overrightarrow{OA} \in \mathrm{Span}(\overrightarrow{OC}-\overrightarrow{OA})$.
Supponiamo ora che i punti A, B e C non siano allineati, allora la coppia di vettori linearmente indipendenti $\{\overrightarrow{OB}-\overrightarrow{OA}, \overrightarrow{OC}-\overrightarrow{OA}\}$ danno una giacitura per il piano π determinato dai punti. Il piano π è quindi il piano per il punto A avente come giacitura $\{\overrightarrow{OB}-\overrightarrow{OA}, \overrightarrow{OC}-\overrightarrow{OA}\}$, ha quindi equazione parametrica:

$$P = A + t(\overrightarrow{OB}-\overrightarrow{OA}) + s(\overrightarrow{OC}-\overrightarrow{OA}), \quad t, s \in \mathbb{R}.$$

Esempio 1.13. Fissiamo nello spazio un sistema di riferimento cartesiano ortonormale $\mathcal{R}(O, \hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}, \hat{\boldsymbol{k}})$.

(1) Troviamo una rappresentazione parametrica per il piano x, y del sistema di riferimento. Il piano x, y è il piano $\mathrm{Span}(\hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}})$, per cui ha equazione parametrica:

$$P = t\hat{\boldsymbol{\imath}} + s\hat{\boldsymbol{\jmath}}, \quad t, s \in \mathbb{R}.$$

Passando alle coordinate abbiamo:

$$\begin{cases} x = t, \\ y = s, \\ z = 0, \qquad t, s \in \mathbb{R}. \end{cases}$$

(2) Troviamo una rappresentazione parametrica per il piano π passante per il punto $A = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$ e avente come giacitura i vettori $\boldsymbol{u} = \hat{\boldsymbol{\imath}} - \hat{\boldsymbol{\jmath}}$ e $\boldsymbol{v} = 2\hat{\boldsymbol{\imath}} + \hat{\boldsymbol{k}}$.
Ricordiamo che i vettori $\{\boldsymbol{u}, \boldsymbol{v}\}$ sono una giacitura per il piano $\pi \iff \boldsymbol{u}$ e $\boldsymbol{v}$ sono linearmente indipendenti $\iff \boldsymbol{v} \notin \mathrm{Span}(\boldsymbol{u})$. Abbiamo

$$\boldsymbol{u} = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \quad \boldsymbol{v} = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix},$$

inoltre

$$\mathrm{Span}(\boldsymbol{u}) = \{\alpha\boldsymbol{u}, \alpha \in \mathbb{R}\} = \{\begin{pmatrix} \alpha \\ -\alpha \\ 0 \end{pmatrix}, \alpha \in \mathbb{R}\}.$$

Il vettore $\boldsymbol{v} \in \mathrm{Span}(\boldsymbol{u})$ se esiste $\alpha \in \mathbb{R}$ tale che:

$$\begin{pmatrix} \alpha \\ -\alpha \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix},$$

che equivale al seguente sistema di equazioni nell'incognita α:

$$\begin{cases} \alpha = 2 \\ -\alpha = 0 \\ 0 = 1 \end{cases}$$

che non ha soluzioni reali. I vettori $\{\boldsymbol{u}, \boldsymbol{v}\}$ sono quindi linearmente indipendenti e il piano π ha equazione parametrica:

$$P = A + t\ \boldsymbol{u} + s\ \boldsymbol{v}, \quad t, s \in \mathbb{R}.$$

Sostituendo e passando alle coordinate otteniamo:

$$\begin{cases} x = 1 + t + 2s, \\ y = 1 - t, \\ z = -1 + s, \end{cases} \qquad t, s \in \mathbb{R}.$$

(3) Consideriamo i punti

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

Ci proponiamo di verificare che i punti non sono allineati e di trovare una rappresentazione parametica per il piano da essi determinato. Abbiamo visto che i punti sono allineati $\Longleftrightarrow$ i vettori $\overrightarrow{OB} - \overrightarrow{OA}$ e $\overrightarrow{OC} - \overrightarrow{OA}$ sono linearmente dipendenti. Nel nostro caso abbiamo:

$$\overrightarrow{OB} - \overrightarrow{OA} = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} - \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ -2 \end{pmatrix},$$

$$\overrightarrow{OC} - \overrightarrow{OA} = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} - \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix},$$

infine risulta

$$\text{Span}(\overrightarrow{OC} - \overrightarrow{OA}) = \{\alpha \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}, \alpha \in \mathbb{R}\} = \{\begin{pmatrix} \alpha \\ \alpha \\ -\alpha \end{pmatrix}, \alpha \in \mathbb{R}\}.$$

Il vettore $\overrightarrow{OB} - \overrightarrow{OA}$ appartiene a $\text{Span}(\overrightarrow{OC} - \overrightarrow{OA})$ se e solo se esiste $\alpha \in \mathbb{R}$ tale che $\begin{pmatrix} -1 \\ 1 \\ -2 \end{pmatrix} = \alpha \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$, passando alle coordinate si ha il sistema

$$\begin{cases} -1 = \alpha \\ 1 = \alpha \\ -2 = -\alpha \end{cases}$$

che risulta essere impossibile. Possiamo quindi concludere che i vettori sono linearmente indipendenti, quindi i punti A, B e C non sono allineati e determinano un unico piano π. L'equazione parametrica del piano π è:

$$P = A + t(\overrightarrow{OB} - \overrightarrow{OA}) + s(\overrightarrow{OC} - \overrightarrow{OA}), \quad t, s \in \mathbb{R}.$$

Sostituendo e passando alle coordinate otteniamo:

$$\begin{cases} x = 1 - t + s, \\ y = 0 + t + s, \\ z = 1 - 2t - s, \qquad t, s \in \mathbb{R}. \end{cases}$$

4. Prodotto scalare in $\mathbb{E}^3_O$.

Vogliamo ora occuparci di questioni di carattere metrico in $\mathbb{E}^3_O$, come ad esempio la misura di distanze ed angoli. Iniziamo a richiamare il concetto di proiezione ortogonale.
Siano $r = \text{Span}(\boldsymbol{u})$ una retta per O e P un punto non appartenente ad r. Consideriamo il piano Σ contenente r e P e nel piano Σ la retta n passante per P e perpendicolare a r, abbiamo: $r \cap n = P'$.

Definizione 1.10. Il punto P' è la proiezione ortogonale di P sulla retta $r = \text{Span}(\boldsymbol{u})$, il vettore $\overrightarrow{OP'}$ è la proiezione ortogonale di $\overrightarrow{OP}$ su r.

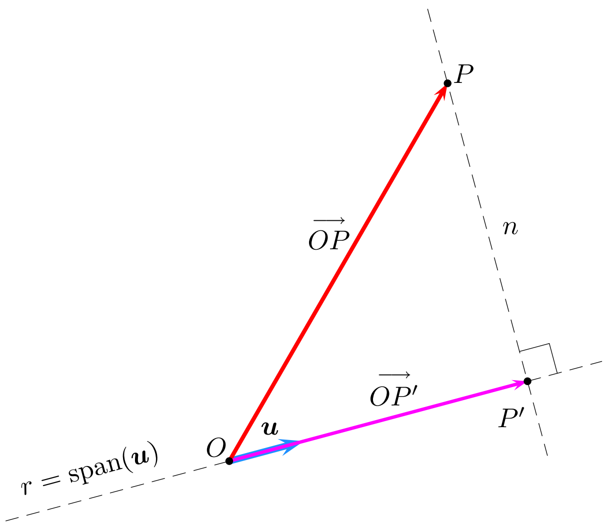

Figura 1.13. La proiezione ortogonale del vettore $\overrightarrow{OP}$ sulla retta $r = \text{Span}(\boldsymbol{u})$ è il vettore $\overrightarrow{OP'}$, con $P' = r \cap n$, dove n è la retta ortogonale ad r passante per P .

Osservazione 1.14. Sia $\overrightarrow{OP'}$ la proiezione ortogonale di $\overrightarrow{OP}$ su r. Osserviamo che:

(1) $\overrightarrow{OP'} = \mathbf{0}$ se e solo se OP è perpendicolare a r;
(2) $\overrightarrow{OP'} \in \text{Span}(\boldsymbol{u})$, quindi esiste un'unico numero reale k tale che

$$\overrightarrow{OP'} = k\boldsymbol{u};$$

(3) se $\hat{\boldsymbol{u}}$ è un versore, allora $k = |\overrightarrow{OP}| \cos\vartheta$, dove $0 \leq \vartheta \leq \pi$ è l'angolo convesso formato da $\overrightarrow{OP}$ e $\hat{\boldsymbol{u}}$ (Figura 1.14).

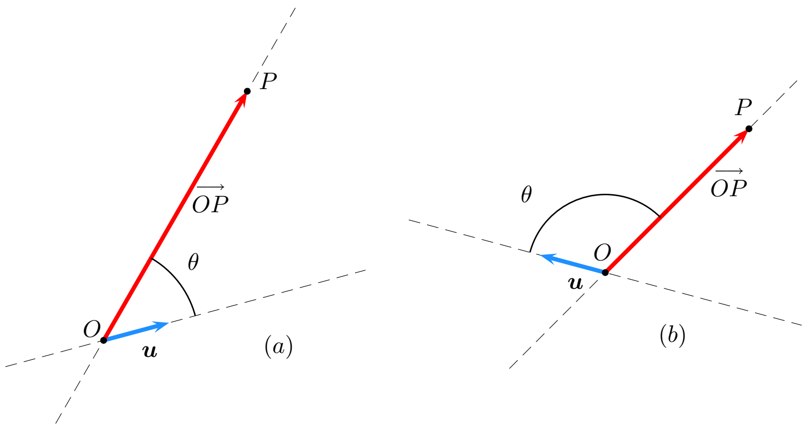

FIGURA 1.14. L'angolo convesso ϑ formato da $\overrightarrow{OP}$ e $\boldsymbol{u}$ può essere acuto (a) o ottuso (b).

Siano $\pi = \text{Span}(\boldsymbol{v}, \boldsymbol{w})$ un piano per O e P un punto non appartenente a π. Consideriamo la retta n passante per P e perpendicolare a π, abbiamo:

$$\pi \cap n = P''.$$

DEFINIZIONE 1.11. Il punto P'' è la proiezione ortogonale di P sul piano $\pi = \text{Span}(\boldsymbol{v}, \boldsymbol{w})$, il vettore $\overrightarrow{OP''}$ è la proiezione ortogonale di $\overrightarrow{OP}$ su π (Figura1.15).

Data una retta $r = \text{Span}(\boldsymbol{u})$ possiamo considerare il piano ortogonale a r e passante per l'origine O: indichiamo tale piano con $\pi_r^\perp$. Consideriamo ora le proiezioni P' e P'' di un punto P, rispettivamente sulla retta r e sul piano $\pi_r^\perp$. Abbiamo il seguente risultato:

PROPOSIZIONE 1.9. *Ogni vettore $\overrightarrow{OP} \in \mathbb{E}^3_O$ si può scrivere nel seguente modo:*

$$\overrightarrow{OP} = \overrightarrow{OP'} + \overrightarrow{OP''},$$

dove $\overrightarrow{OP'}$ è la proiezione ortogonale di $\overrightarrow{OP}$ su r e $\overrightarrow{OP''}$ è la proiezione ortogonale di $\overrightarrow{OP}$ su $\pi_r^\perp$. Infine, tale scrittura è unica.

Basta osservare che esiste un'unico rettangolo con un vertice in P, un vertice in O e avente due lati rispettivamente sulla retta $r = \text{Span}(\boldsymbol{u})$ e sul piano $\pi_r^\perp = \text{Span}(\boldsymbol{v}, \boldsymbol{w})$: per costruzione i lati $\overrightarrow{OP'}$ e $\overrightarrow{OP''}$ sono le proiezioni ortogonali del vettore $\overrightarrow{OP}$ su r e $\pi_r^\perp$ (Figura 1.16).

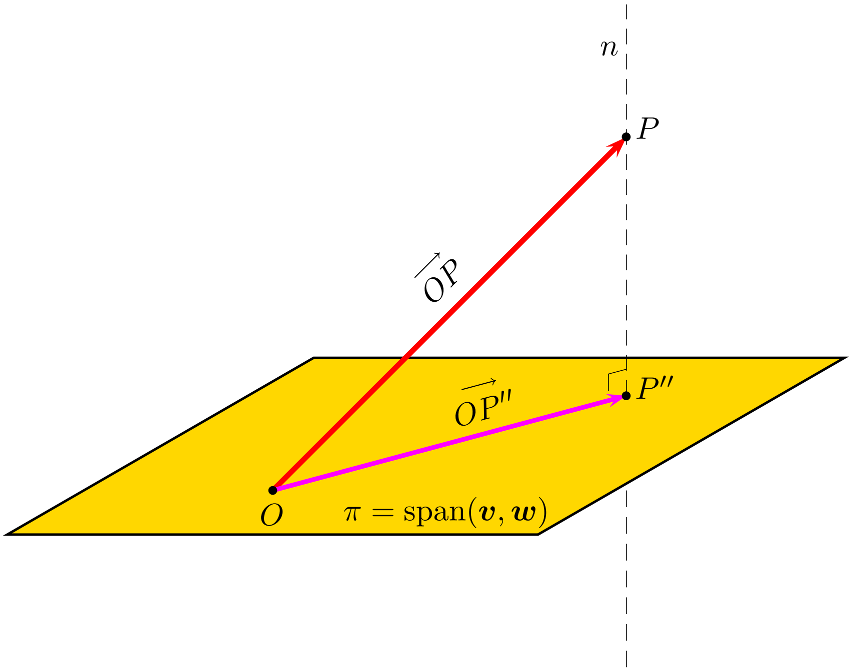

FIGURA 1.15. La proiezione ortogonale del vettore $\overrightarrow{OP}$ sul piano $\pi = \mathrm{Span}(\boldsymbol{v}, \boldsymbol{w})$ è il vettore $\overrightarrow{OP''}$, dove P'' è l'intersezione fra il piano e la retta n passante per P, ortogonale a π (per maggiore chiarezza, la figura omette di illustrare i vettori $\boldsymbol{v}$ e $\boldsymbol{w}$).

Una proprietà importante della proiezione ortogonale è che si comporta "bene" rispetto alla somma di vettori e rispetto alla moltiplicazione di un vettore per un numero reale.

PROPRIETÀ 1.10. *Siano r e $\pi_r^\perp$ come sopra, proviamo che:*

(1) *se $\overrightarrow{OR} = \overrightarrow{OP} + \overrightarrow{OQ}$, allora si ha:*
$$\overrightarrow{OR'} = \overrightarrow{OP'} + \overrightarrow{OQ'} \quad \overrightarrow{OR''} = \overrightarrow{OP''} + \overrightarrow{OQ''};$$
(2) *se $\overrightarrow{OR} = \alpha\overrightarrow{OP}$, allora si ha:*
$$\overrightarrow{OR'} = \alpha\overrightarrow{OP'} \quad \overrightarrow{OR''} = \alpha\overrightarrow{OP''}.$$

Verifichiamo la prima proprietà. Notiamo che dalla proprietà precedente abbiamo:
$$\begin{gathered}\overrightarrow{OP} = \overrightarrow{OP'} + \overrightarrow{OP''}\\ \overrightarrow{OQ} = \overrightarrow{OQ'} + \overrightarrow{OQ''}\end{gathered}$$
e sommando queste due uguaglianze troviamo
$$\overrightarrow{OP} + \overrightarrow{OQ} = (\overrightarrow{OP'} + \overrightarrow{OQ'}) + (\overrightarrow{OP''} + \overrightarrow{OQ''}).$$

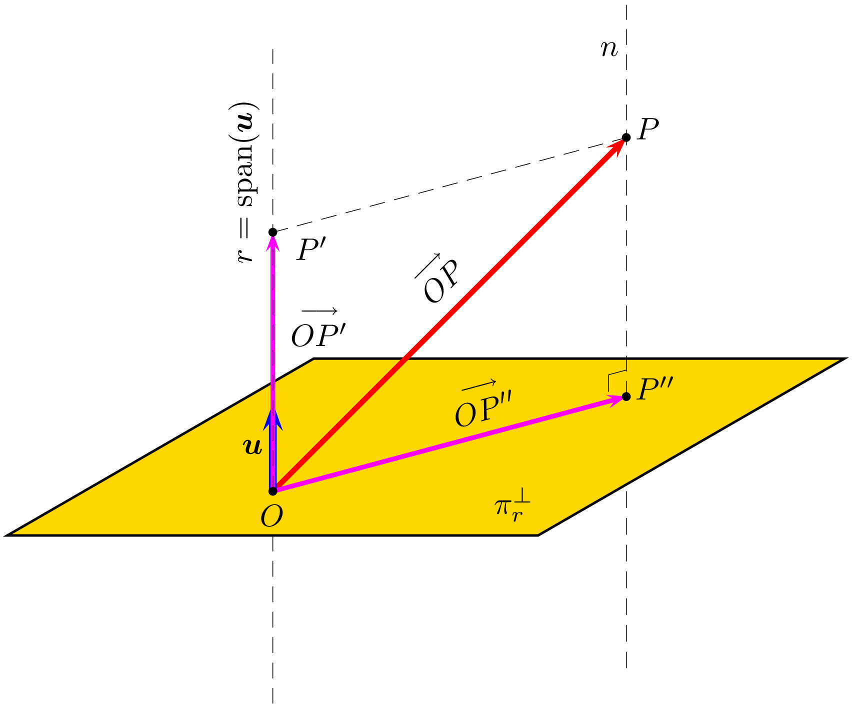

FIGURA 1.16. Il vettore $\overrightarrow{OP}$ si decompone in un vettore $\overrightarrow{OP'}$, proioezione ortogonale lungo $r = \mathrm{Span}(\boldsymbol{u})$, ed uno $\overrightarrow{OP''}$ proiezione ortogonale nel piano $\pi_r^\perp = \mathrm{Span}(\boldsymbol{v}, \boldsymbol{w})$ perpendicolare ad $\boldsymbol{u}$. La decomposizione è unica.

Ora $(\overrightarrow{OP'} + \overrightarrow{OQ'}) \in r$ e $(\overrightarrow{OP''} + \overrightarrow{OQ''}) \in r^\perp$, per l'unicità della scrittura possiamo concludere.

Introduciamo finalmente lo strumento fondamentale nello studio della geometria analitica, strettamante legato alle proiezioni ortogonali: il *prodotto scalare*, un'operazione che associa a due vettori $\boldsymbol{u}, \boldsymbol{v}$ di $\mathbb{E}^3_O$ un numero reale, nel seguente modo:

DEFINIZIONE 1.12. Il prodotto scalare di $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{E}^3_O$ è:

$$\langle \boldsymbol{u}, \boldsymbol{v} \rangle = |\boldsymbol{u}||\boldsymbol{v}| \cos \vartheta \tag{1.11}$$

dove ϑ è l'angolo convesso compreso tra i due vettori, come nel caso descritto sopra della proiezione ortogonale (Figura 1.14).

OSSERVAZIONE 1.15. Osserviamo che dalla definizione segue immediatamente che:
- se $\hat{\boldsymbol{u}}$ è un versore, $r = \mathrm{Span}(\hat{\boldsymbol{u}})$ e $\boldsymbol{v} = \overrightarrow{OP}$, allora la proiezione ortogonale del vettore $\overrightarrow{OP}$ sulla retta r è:

$$\overrightarrow{OP'} = \langle \hat{\boldsymbol{u}}, \overrightarrow{OP} \rangle \hat{\boldsymbol{u}};$$

- se $\boldsymbol{u} = \boldsymbol{v}$, si ha: $\langle \boldsymbol{v}, \boldsymbol{v} \rangle = |\boldsymbol{v}|^2$;
- $\langle \boldsymbol{u}, \boldsymbol{v} \rangle = 0$ se e solo se o uno dei due vettori è il vettore nullo di $\mathbb{E}^3_O$ oppure se i vettori $\boldsymbol{u}$ e $\boldsymbol{v}$ sono perpendicolari.

Le proprietà seguenti del prodotto scalare sono molto importanti:

Proprietà 1.11. *$\forall \boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in \mathbb{E}^3_O$ e $\forall k \in \mathbb{R}$ valgono le seguenti proprietà:*
(1) $\langle \boldsymbol{u}, \boldsymbol{v} \rangle = \langle \boldsymbol{v}, \boldsymbol{u} \rangle$ *(proprietà* commutativa*);*
(2) $\langle k\boldsymbol{u}, \boldsymbol{v} \rangle = \langle \boldsymbol{v}, k\boldsymbol{u} \rangle = k\langle \boldsymbol{u}, \boldsymbol{v} \rangle$*;*
(3) $\langle (\boldsymbol{u} + \boldsymbol{v}), \boldsymbol{w} \rangle = \langle \boldsymbol{u}, \boldsymbol{w} \rangle + \langle \boldsymbol{v}, \boldsymbol{w} \rangle$ *(proprietà* distributiva*).*

La prima proprietà segue dalla definizione.
La seconda proprietà segue dal fatto che: $|k\boldsymbol{u}| = |k||\boldsymbol{u}|$ e fissato $k \neq 0$, i vettori $k\boldsymbol{u}$ e $\boldsymbol{v}$, $\boldsymbol{u}$ e $k\boldsymbol{v}$ formano lo stesso angolo convesso β; inoltre, è facile verificare che si ha:

$$\beta = \begin{cases} \vartheta & k > 0 \\ \pi - \vartheta & k < 0 \end{cases} ,$$

da cui si ricava $\cos\beta = \frac{k}{|k|}\cos\vartheta$. Si ha quindi:

$$\langle k\boldsymbol{u}, \boldsymbol{v} \rangle = |k||\boldsymbol{u}||\boldsymbol{v}|\cos\beta = k|\boldsymbol{u}||\boldsymbol{v}|\cos\vartheta = k\langle \boldsymbol{u}, \boldsymbol{v} \rangle.$$

La terza proprietà è un po' più delicata. Se $\boldsymbol{w} = \boldsymbol{0}$ non c'è nulla da dimostrare. Indichiamo con r la retta generata da $\boldsymbol{w}$ e consideriamo il versore $\hat{\boldsymbol{w}} = \boldsymbol{w}/|\boldsymbol{w}|$. Osserviamo che

$$\langle \boldsymbol{u} + \boldsymbol{v}, \boldsymbol{w} \rangle = |\boldsymbol{w}|\langle \boldsymbol{u} + \boldsymbol{v}, \hat{\boldsymbol{w}} \rangle$$

$$\langle \boldsymbol{u}, \boldsymbol{w} \rangle + \langle \boldsymbol{v}, \boldsymbol{w} \rangle = |\boldsymbol{w}|(\langle \boldsymbol{u}, \hat{\boldsymbol{w}} \rangle + \langle \boldsymbol{v}, \hat{\boldsymbol{w}} \rangle),$$

dunque è sufficiente provare che $\langle (\boldsymbol{u} + \boldsymbol{v}), \hat{\boldsymbol{w}} \rangle = \langle \boldsymbol{u}, \hat{\boldsymbol{w}} \rangle + \langle \boldsymbol{v}, \hat{\boldsymbol{w}} \rangle$.
Poiché $\hat{\boldsymbol{w}}$ è un versore, dall'osservazione precedente sappiamo che $\langle \boldsymbol{u}+\boldsymbol{v}, \hat{\boldsymbol{w}} \rangle \hat{\boldsymbol{w}}$ è la proiezione ortogonale del vettore $\boldsymbol{u}+\boldsymbol{v}$ sulla retta r, dalla proprietà dimostrata sulle proiezioni ortogonali si ha

$$\langle \boldsymbol{u} + \boldsymbol{v}, \hat{\boldsymbol{w}} \rangle \hat{\boldsymbol{w}} = \langle \boldsymbol{u}, \hat{\boldsymbol{w}} \rangle \hat{\boldsymbol{w}} + \langle \boldsymbol{v}\hat{\boldsymbol{w}} \rangle \hat{\boldsymbol{w}},$$

da cui segue ciò che volevamo dimostrare.

Le proprietà (2) e (3) costituiscono globalmente la *proprietà di bilinearità del prodotto scalare.*

Osservazione 1.16. Fissiamo ora in $\mathbb{E}^3_O$ una base ortonormale $\mathcal{B} = \{\hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}, \hat{\boldsymbol{k}}\}$. Sfruttando il fatto che i vettori sono mutualmente ortogonali otteniamo le relazioni:

$$\langle \hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}} \rangle = \langle \hat{\boldsymbol{\imath}}, \hat{\boldsymbol{k}} \rangle = \langle \hat{\boldsymbol{\jmath}}, \hat{\boldsymbol{k}} \rangle = 0.$$

Sfruttando il fatto che sono versori ricaviamo invece

$$\langle \hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\imath}} \rangle = \langle \hat{\boldsymbol{\jmath}}, \hat{\boldsymbol{\jmath}} \rangle = \langle \hat{\boldsymbol{k}}, \hat{\boldsymbol{k}} \rangle = 1.$$

Utilizzando le proprietà di bilinearità e queste relazioni possiamo ricavare l'espressione del prodotto scalare in coordinate. Ovvero rispondere alla domanda: conoscendo le coordinate di $\boldsymbol{u}$ e di $\boldsymbol{v}$, in una base ortonormale $\mathcal{B}$, come posso calcolare $\langle \boldsymbol{u}, \boldsymbol{v} \rangle$?

PROPRIETÀ 1.12. *Fissata in $\mathbb{E}^3_O$ la base ortonormale $\mathcal{B} = \{\hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}, \hat{\boldsymbol{k}}\}$, consideriamo i seguenti vettori:*

$$\boldsymbol{u} = x\hat{\boldsymbol{\imath}} + y\hat{\boldsymbol{\jmath}} + z\hat{\boldsymbol{k}} \quad \boldsymbol{v} = x'\hat{\boldsymbol{\imath}} + y'\hat{\boldsymbol{\jmath}} + z'\hat{\boldsymbol{k}}.$$

Risulta:

(1) $\langle \boldsymbol{u}, \boldsymbol{v} \rangle = xx' + yy' + zz'$;
(2) $|\boldsymbol{u}| = \sqrt{x^2 + y^2 + z^2}$.

Verifichiamo la proprietà (1): sostituendo al posto di $\boldsymbol{u}$ e $\boldsymbol{v}$ la loro scrittura nella base $\mathcal{B}$ abbiamo:

$$\langle \boldsymbol{u}, \boldsymbol{v} \rangle = \langle x\hat{\boldsymbol{\imath}} + y\hat{\boldsymbol{\jmath}} + z\hat{\boldsymbol{k}}, x'\hat{\boldsymbol{\imath}} + y'\hat{\boldsymbol{\jmath}} + z'\hat{\boldsymbol{k}} \rangle,$$

per la linearità del prodotto scalare abbiamo:

$$\begin{aligned} \langle \boldsymbol{u}, \boldsymbol{v} \rangle =& xx'\langle \hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\imath}} \rangle + (xy' + x'y)\langle \hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}} \rangle + (xz' + x'z)\langle \hat{\boldsymbol{\imath}}, \hat{\boldsymbol{k}} \rangle \\ &+ yy'\langle \hat{\boldsymbol{\jmath}}, \hat{\boldsymbol{\jmath}} \rangle + (yz' + y'z)\langle \hat{\boldsymbol{\jmath}}, \hat{\boldsymbol{k}} \rangle + zz'\langle \hat{\boldsymbol{k}}, \hat{\boldsymbol{k}} \rangle, \end{aligned} \tag{1.12}$$

sostituendo le relazioni trovate abbiamo:

$$\langle \boldsymbol{u}, \boldsymbol{v} \rangle = xx' + yy' + zz'.$$

La proprietà (2) segue immediatamente dalla (1).

OSSERVAZIONE 1.17. Nelle ipotesi precedenti, sia ϑ l'angolo convesso formato dai vettori $\boldsymbol{u}$ e $\boldsymbol{v}$. Dalla definizione di prodotto scalare e dalle proprietà (1) e (2) precedenti risulta:

$$\cos\vartheta = \frac{xx' + yy' + zz'}{\sqrt{x^2 + y^2 + z^2}\sqrt{x'^2 + y'^2 + z'^2}}.$$

ESEMPIO 1.18. Fissiamo nello spazio un sistema di riferimento cartesiano ortonormale $\mathcal{R}(O, \hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}, \hat{\boldsymbol{k}})$.

(1) Dati i vettori $\boldsymbol{u} = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}$ e $\boldsymbol{v} = \begin{pmatrix} -6 \\ 3 \\ 4 \end{pmatrix}$, calcoliamo il loro prodotto scalare per stabilire se sono ortogonali.
Ricordiamo che i vettori $\boldsymbol{u}$ e $\boldsymbol{v}$ sono perpendicolari se e solo se $\langle \boldsymbol{u}, \boldsymbol{v} \rangle = 0$, abbiamo

$$\langle \boldsymbol{u}, \boldsymbol{v} \rangle = xx' + yy' + zz' = 1 \cdot (-6) + 2 \cdot (3) + 0 \cdot (4) = -6 + 6 = 0,$$

quindi i due vettori sono ortogonali.

(2) Dati i vettori $\boldsymbol{u} = \begin{pmatrix} -2 \\ 0 \\ 0 \end{pmatrix}$ e $\boldsymbol{v} = \begin{pmatrix} 3 \\ 3 \\ 0 \end{pmatrix}$, determiniamo l'angolo convesso da essi formato.
Abbiamo:

$$\langle \boldsymbol{u}, \boldsymbol{v} \rangle = xx' + yy' + zz' = -2 \cdot (3) + 0 \cdot (3) + 0 \cdot 0 = -6 \neq 0,$$

quindi i vettori non sono perpendicolari. Calcoliamo i loro moduli:

$$|\boldsymbol{u}| = \sqrt{x^2 + y^2 + z^2} = \sqrt{(-2)^2} = \sqrt{4} = 2,$$

$$|\boldsymbol{v}| = \sqrt{x'^2 + y'^2 + z'^2} = \sqrt{3^2 + 3^2} = \sqrt{18} = 3\sqrt{2}.$$

Indichiamo con ϑ l'angolo convesso formato dai vettori $\boldsymbol{u}$ e $\boldsymbol{v}$, abbiamo:

$$\cos\vartheta = \frac{\langle \boldsymbol{u}, \boldsymbol{v} \rangle}{|\boldsymbol{u}||\boldsymbol{v}|} = \frac{-6}{6\sqrt{2}} = -\frac{\sqrt{2}}{2},$$

da cui deduciamo che $\vartheta = \frac{3}{4}\pi$.

(3) Dati i vettori $\boldsymbol{u} = \hat{\boldsymbol{\imath}} + \hat{\boldsymbol{\jmath}}$ e $\boldsymbol{v} = 2\hat{\boldsymbol{\imath}} + 3\hat{\boldsymbol{\jmath}} - 1\hat{\boldsymbol{k}}$, determiniamo il vettore proiezione ortogonale di $\boldsymbol{v}$ sulla retta $\mathrm{Span}(\boldsymbol{u})$.

Scegliamo un versore sulla retta $r = \mathrm{Span}(\boldsymbol{u})$: abbiamo $\boldsymbol{u} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$, quindi il modulo di $\boldsymbol{u}$ è:

$$|\boldsymbol{u}| = \sqrt{x^2 + y^2 + z^2} = \sqrt{1^2 + 1^2} = \sqrt{2},$$

allora possiamo considerare sulla retta r il versore

$$\hat{\boldsymbol{u}} = \frac{1}{\sqrt{2}}\boldsymbol{u} = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix},$$

così $\mathrm{Span}(\boldsymbol{u}) = \mathrm{Span}(\hat{\boldsymbol{u}})$. Ricordiamo il significato geometrico del prodotto scalare, poiché $\hat{\boldsymbol{u}}$ è un versore, il vettore proiezione ortogonale di $\boldsymbol{v}$ sulla retta r è il vettore $\boldsymbol{v}' \in \mathrm{Span}(\hat{\boldsymbol{u}})$ dato da:

$$\boldsymbol{v}' = \langle \hat{\boldsymbol{u}}, \boldsymbol{v} \rangle \hat{\boldsymbol{u}}.$$

Calcoliamo il prodotto scalare di $\hat{\boldsymbol{u}}$ e $\boldsymbol{v}$:

$$\langle \hat{\boldsymbol{u}}, \boldsymbol{v} \rangle = \frac{1}{\sqrt{2}} \cdot 2 + \frac{1}{\sqrt{2}} \cdot 3 + 0 \cdot (-1) = \frac{5}{\sqrt{2}},$$

da cui otteniamo:

$$\boldsymbol{v}' = \frac{5}{\sqrt{2}}\hat{\boldsymbol{u}} = \frac{5}{2}\boldsymbol{u} = \frac{5}{2}\hat{\boldsymbol{\imath}} + \frac{5}{2}\hat{\boldsymbol{\jmath}}.$$

(4) Dati i vettori $\boldsymbol{u} = \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}$ e $\boldsymbol{v} = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}$, verifichiamo che non sono ortogonali e determiniamo i vettori $\boldsymbol{w} \in \mathrm{Span}(\boldsymbol{u}, \boldsymbol{v})$ ortogonali a $\boldsymbol{v}$.

Calcoliamo il prodotto scalare di $\boldsymbol{u}$ e $\boldsymbol{v}$:

$$\begin{aligned}\langle \boldsymbol{u}, \boldsymbol{v}\rangle &= xx' + yy' + zz' \\ &= 1 \cdot 1 + 2 \cdot (-1) + -1 \cdot (2) = 1 - 2 - 2 = -3 \neq 0,\end{aligned} \tag{1.13}$$

da cui deduciamo che i vettori non sono ortogonali.
Osserviamo che i vettori $\boldsymbol{u}$ e $\boldsymbol{v}$ sono linearmente indipendenti: infatti $\boldsymbol{v} \in \mathrm{Span}(\boldsymbol{u})$ se e solo se esiste un numero reale non nullo k tale che $\boldsymbol{v} = k\boldsymbol{u}$, ma ciò è impossibile poiché il seguente sistema

$$\begin{cases} k = 1 \\ 2k = -1 \\ -k = 2 \end{cases}$$

non ha soluzioni. Quindi $\mathrm{Span}(\boldsymbol{u}, \boldsymbol{v})$ è un piano per l'origine e un vettore $\boldsymbol{w} \in \mathrm{Span}(\boldsymbol{u}, \boldsymbol{v})$ se e solo se si scrive nel seguente modo:

$$\boldsymbol{w} = a\boldsymbol{u} + b\boldsymbol{v}, \quad a, b \in \mathbb{R};$$

infine $\boldsymbol{w}$ è ortogonale a $\boldsymbol{v}$ se il loro prodotto scalare è zero. Per la linearità del prodotto scalare abbiamo:

$$\langle \boldsymbol{w}, \boldsymbol{v}\rangle = \langle a\boldsymbol{u} + b\boldsymbol{v}, \boldsymbol{v}\rangle = a\langle \boldsymbol{u}, \boldsymbol{v}\rangle + b\langle \boldsymbol{v}, \boldsymbol{v}\rangle = 0,$$

calcoliamo

$$\langle \boldsymbol{v}, \boldsymbol{v}\rangle = x^2 + y^2 + z^2 = 1^2 + (-1)^2 + 2^2 = 6,$$

sostituendo otteniamo l'equazione:

$$-3a + 6b = 0,$$

le cui soluzioni sono (a, b) tali che $a = 2b$. Le soluzioni sono quindi i seguenti vettori:

$$\boldsymbol{w} = 2b\boldsymbol{u} + b\boldsymbol{v} = b(2\boldsymbol{u} + \boldsymbol{v}) \quad b \in \mathbb{R},$$

l'insieme ottenuto è $\mathrm{Span}(2\boldsymbol{u} + \boldsymbol{v})$, quindi la retta per O contenuta nel piano $\mathrm{Span}(\boldsymbol{u}, \boldsymbol{v})$ perpendicolare alla retta $\mathrm{Span}(\boldsymbol{v})$.

(5) Siano $\boldsymbol{u}$ e $\boldsymbol{v}$ due vettori non nulli linearmente indipendenti. Verifichiamo che se $\boldsymbol{w}$ è un vettore non nullo perpendicolare a $\boldsymbol{u}$ e a $\boldsymbol{v}$, allora $\boldsymbol{w}$ è perpendicolare a tutti i vettori di $\mathrm{Span}(\boldsymbol{u}, \boldsymbol{v})$.
Un vettore del piano $\mathrm{Span}(\boldsymbol{u}, \boldsymbol{v})$ si scrive $a\boldsymbol{u} + b\boldsymbol{v}$, con $a, b \in \mathbb{R}$, per la linearità del prodotto scalare abbiamo quindi:

$$\langle \boldsymbol{w}, a\boldsymbol{u} + b\boldsymbol{v}\rangle = a\langle \boldsymbol{w}, \boldsymbol{u}\rangle + b\langle \boldsymbol{w}, \boldsymbol{v}\rangle = a \cdot 0 + b \cdot 0 = 0,$$

poiché per ipotesi $\langle \boldsymbol{w}, \boldsymbol{u}\rangle = \langle \boldsymbol{w}, \boldsymbol{v}\rangle = 0$.

Applicazione 1.19. Distanza di due punti.

Dati due punti distinti A e B, la loro distanza è la lunghezza del segmento che ha per estremi i punti A e B, quindi il modulo del vettore $\overrightarrow{AB}$:

$$d(A, B) = |\overrightarrow{AB}|.$$

Fissato nello spazio un sistema di riferimento $\mathcal{R}(O, \hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}, \hat{\boldsymbol{k}})$ cartesiano ortonormale, osserviamo che il vettore applicato $\overrightarrow{OB} - \overrightarrow{OA}$ ha direzione, verso e modulo del segmento orientato $\overrightarrow{AB}$ (vedi Esercizio (1.2). Quindi abbiamo:

$$d(A, B) = |\overrightarrow{OB} - \overrightarrow{OA}|.$$

Posto $A = \begin{pmatrix} x_A \\ y_A \\ z_A \end{pmatrix}$ e $B = \begin{pmatrix} x_B \\ y_B \\ z_B \end{pmatrix}$, poiché abbiamo $\overrightarrow{OB} - \overrightarrow{OA} = \begin{pmatrix} x_B - x_A \\ y_B - y_A \\ z_B - z_A \end{pmatrix}$, otteniamo la formula:

$$d(A, B) = \sqrt{(x_B - x_A)^2 + (y_B - y_A)^2 + (z_B - z_A)^2},$$

che generalizza l'analoga ben nota formula della distanza di due punti nel piano.

5. Equazioni cartesiane di rette e piani.

È noto a tutti che, fissato nel piano un sistema di riferimento cartesiano ortonormale $\mathcal{R}(O, \hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}})$, le rette sono descrivibili come luogo dei punti le cui coordinate sono le soluzioni di un'equazione lineare

$$ax + by + c = 0, \quad (a, b) \neq (0, 0).$$

Proviamo che, analogamente, fissato nello spazio un sistema di riferimento cartesiano ortonormale $\mathcal{R}(O, \hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}, \hat{\boldsymbol{k}})$, i piani sono definiti da un'equazione lineare. Per fare ciò utilizzeremo la nozione di ortogonalità.

Dato un piano π esiste un'unica retta l passante per O ortogonale a π: l è detta *retta normale* al piano per O. Sia $\boldsymbol{n} \in \mathbb{E}^3_O$ un vettore tale che $\mathrm{Span}(\boldsymbol{n}) = l$:

Definizione 1.13. $\boldsymbol{n}$ è un *vettore normale* al piano π (Figura 1.17).

Osservazione 1.20. Osserviamo che ogni vettore non nullo $\boldsymbol{n}' \in \mathrm{Span}(\boldsymbol{n})$ è un vettore normale al piano π.

Supponiamo che il piano π passi per l'origine O. Osserviamo che se $P \in \pi$ il vettore $\overrightarrow{OP}$ risulta ortogonale a $\mathbf{n}$, viceversa se $\overrightarrow{OP} \perp \mathbf{n}$ allora il punto $P \in \pi$ (Figura 1.18); abbiamo, quindi, dimostrato che

$$\pi = \{ P \in \mathcal{E} \mid \overrightarrow{OP} \perp \boldsymbol{n} \} = \{ P \in \mathcal{E} \mid \langle \overrightarrow{OP}, \boldsymbol{n} \rangle = 0 \}. \tag{1.14}$$

Posto $\boldsymbol{n} = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$ e $P = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$, allora abbiamo

$$\langle \overrightarrow{OP}, \boldsymbol{n} \rangle = ax + by + cz,$$

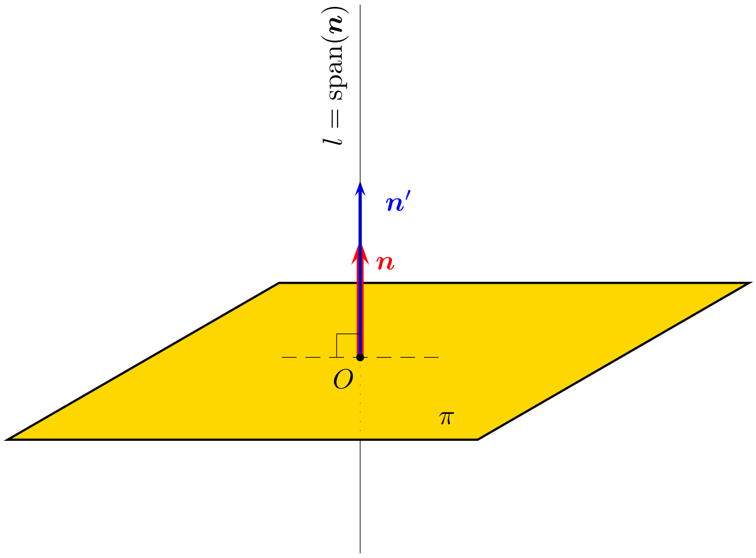

FIGURA 1.17. Il vettore $\boldsymbol{n}$ individua la retta $l = \mathrm{Span}(\boldsymbol{n})$ passante per O ed ortogonale al piano π. Ogni vettore $\boldsymbol{n}' \in \mathrm{Span}(\boldsymbol{n})$ è normale al piano π.

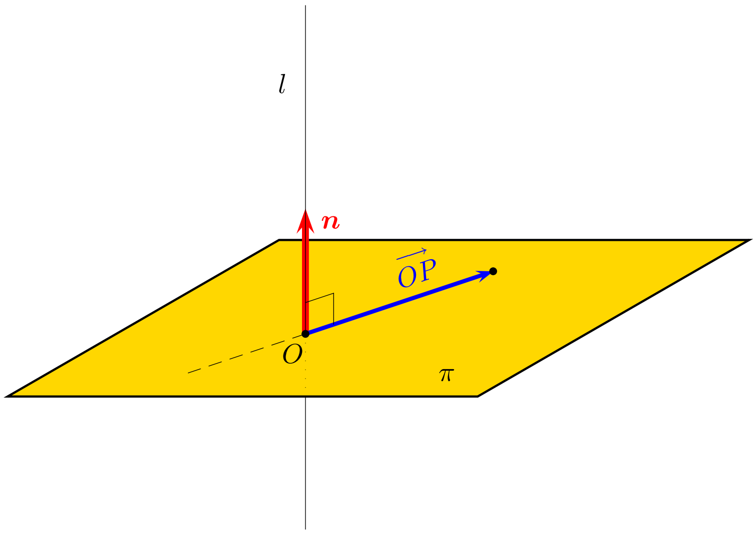

FIGURA 1.18. Se il vettore $\boldsymbol{n}$ è ortogonale al piano π passante per l'origine, preso un qualunque punto P del piano π, il vettore $\overrightarrow{OP}$ è ortogonale a $\boldsymbol{n}$, ossia alla retta $l = \mathrm{Span}(\boldsymbol{n})$ individuata da $\boldsymbol{n}$, e viceversa. Il piano π è dunque il luogo dei punti che individuano vettori $\overrightarrow{OP}$ ortogonali ad $\boldsymbol{n}$.

dunque otteniamo l'equazione cartesiana del piano π per O e di direzione normale $\boldsymbol{n}$:

$$ax + by + cz = 0, \tag{1.15}$$

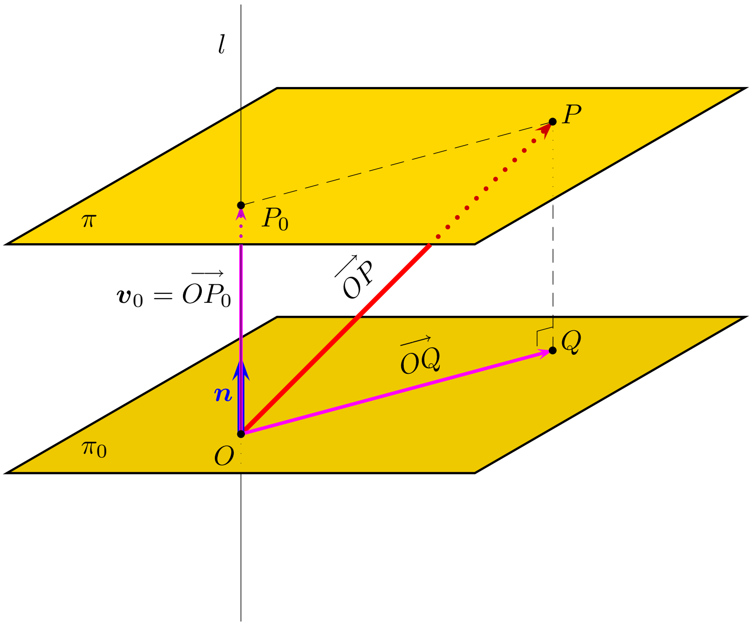

FIGURA 1.19. Il piano π si ottiene traslando il piano π_0 passante per O e parallelo ad esso del vettore $\boldsymbol{v} = \overrightarrow{OP_0}$, in cui P_0 è l'intersezione con π della retta l per O, normale ai piani.

ovvero abbiamo espresso π come luogo dei punti le cui coordinate sono le soluzioni di un'equazione lineare omogenea. I coefficienti di tale equazione sono le coordinate di $\boldsymbol{n}$.

Supponiamo ora che il piano π non passi per l'origine O. Consideriamo il piano π_0 parallelo a π passante per O e sia l la retta normale ai piani per O, tale retta interseca il piano π in un punto P_0: poniamo $\boldsymbol{v}_0 = \overrightarrow{OP_0}$. Osserviamo che risulta:
- per ogni punto $Q \in \pi_0$, il punto $P = Q + \boldsymbol{v}_0$ appartiene al piano π,
- per ogni punto $P \in \pi$, il punto $Q = P - \boldsymbol{v}_0$ appartiene al piano π_0,

possiamo cioè dire che il piano π si ottiene traslando π_0 del vettore $\boldsymbol{v}_0$ (Figura 1.19):

$$\pi = \{P \in \mathcal{E} |\ \ P = Q + \boldsymbol{v}_0,\ Q \in \pi_0\}. \tag{1.16}$$

Concludendo, essendo $\boldsymbol{v}_0 = \overrightarrow{OP_0}$ si ha:

$$\pi = \{P \in \mathcal{E} |\ \langle \overrightarrow{OP} - \overrightarrow{OP_0}, \boldsymbol{n} \rangle = 0\ \}. \tag{1.17}$$

Infine posto $\boldsymbol{n} = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$, $P_0 = \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix}$ e $P = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$, otteniamo l'equazione:

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0, \tag{1.18}$$

ovvero posto $(ax_0 + by_0 + cz_0) = d$, otteniamo l'equazione del piano π per P_0 e di direzione normale $\boldsymbol{n}$:

$$ax + by + cz = d. \tag{1.19}$$

DEFINIZIONE 1.14. L'equazione $ax+by+cz = d$ è detta equazione cartesiana del piano π.

OSSERVAZIONE 1.21. Dato un piano π esistono infiniti vettori normali al piano: infatti $\boldsymbol{n'} \in \mathbb{E}^3_O$ è normale al piano se e solo $\boldsymbol{n'} \in \mathrm{Span}(\boldsymbol{n})$, cioè esiste $k \in \mathbb{R}$, $k \neq 0$ tale che $\boldsymbol{n'} = k\boldsymbol{n}$. Al vettore $\boldsymbol{n'}$ corrisponde l'equazione cartesiana

$$kax + kby + kcz = kd,$$

che rappresenta ancora il piano π. Quindi il piano π ammette infinite equazioni cartesiane ed esse sono tutte multiple l'una dell'altra.
Infine, ogni equazione lineare

$$ax + by + cz = d,$$

con $\begin{pmatrix} a \\ b \\ c \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$, rappresenta un piano di direzione normale $\boldsymbol{n}$.

ESEMPIO 1.22. Fissiamo nello spazio un sistema di riferimento cartesiano ortonormale $\mathcal{R}(O, \hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}, \hat{\boldsymbol{k}})$.

(1) Scriviamo l'equazione cartesiana del piano π passante per O e avente direzione normale $\boldsymbol{n} = \hat{\boldsymbol{\imath}} - \hat{\boldsymbol{\jmath}} + 2\hat{\boldsymbol{k}}$.

Abbiamo $\boldsymbol{n} = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}$, da cui otteniamo l'equazione cartesiana del piano π:

$$1x - 1y + 2z = x - y + 2z = 0.$$

(2) Scriviamo l'equazione cartesiana del piano π passante per $A = \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}$ e avente direzione normale $\boldsymbol{n} = \hat{\boldsymbol{\imath}} - 2\hat{\boldsymbol{k}}$.

Abbiamo $\boldsymbol{n} = \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix}$, l'equazione cartesiana di π è quindi :

$$1(x - x_A) + 0(y - y_A) - 2(z - z_A) = 0,$$

sostituendo le coordinate di A abbiamo :

$$(x - 1) - 2(z + 1) = 0,$$

da cui ricaviamo l'equazione del piano.

$$x - 2z = 3.$$

(3) Vogliamo stabilire se il punto $A = \begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix}$ appartiene al piano π di equazione $2x - y + 2z = 4$.
Basta sostituire le coordinate del punto A nell'equazione e vedere se è soddisfatta oppure no:
$$2(1) - (2) + 2(-3) = -6 \neq 4,$$
per cui concludiamo che il punto A *non* appartiene al piano.

Fissato un sistema di riferimento cartesiano ortonormale $\mathcal{R}(O, \hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}, \hat{\boldsymbol{k}})$, abbiamo visto che esistono due modi per descrivere un piano nello spazio.
• Attraverso l'equazione cartesiana: ovvero descrivendo il piano come il luogo dei punti le cui coordinate sono le soluzioni di un'equazione lineare;
• Attraverso un'equazione parametrica: ovvero descrivendo i punti del piano in funzione di due parametri reali.

Ci si potrebbe domandare quali tra queste due descrizioni conviene usare nella pratica. In realtà la risposta a tale domanda è che *a seconda del problema* a cui si è interessati conviene usare una o l'altra descrizione. Ad esempio, se vogliamo stabilire se un punto A, di coordinate note, appartiene ad un piano π, (vedi esempio (3)): è più utile avere l'equazione cartesiana. Vedremo in seguito esempi di problema i cui è più utile avere l'equazione parametrica. È naturale quindi chiedersi:
Domanda: Dato un piano π, come passare da una descrizione parametrica ad una cartesiana e viceversa?

Supponiamo di volere trovare una rappresentazione parametrica per il piano π di equazione cartesiana
$$ax + by + cz + d = 0,$$
poiché il piano è il luogo dei punti le cui coordinate sono le soluzioni dell'equazione data, trattandosi di un'equazione lineare in 3 incognite, è chiaro che se fissiamo due delle variabili arbitrariamente, possiamo ricavare la terza variabile, e in questo modo ottenere tutte le soluzioni. Vediamo alcuni esempi.

Esempio 1.23. Fissiamo nello spazio un sistema di riferimento cartesiano ortonormale $\mathcal{R}(O, \hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}, \hat{\boldsymbol{k}})$.

(1) Troviamo una rappresentazione parametrica per il piano π di equazione cartesiana: $x + 2y - 3z = 2$.
Poniamo $y = t$ e $z = s$, sostituiamo nell'equazione:
$$x + 2t - 3s = 2,$$
da cui ricaviamo $x = 2 - 2t + 3s$. Abbiamo quindi trovato le seguenti equazioni parametriche del piano:
$$\begin{cases} x = 2 - 2t + 3s, \\ y = t, \\ z = s, \quad t, s \in \mathbb{R}. \end{cases}$$

(2) Troviamo una rappresentazione parametrica per il piano π di equazione $y - 2z = 3$.

Osserviamo che non abbiamo condizioni su x, mentre se fissiamo z allora y è determinato. Dunque poniamo $x = t$ e $z = s$, sostituendo nell'equazione

$$y - 2s = 3,$$

abbiamo $y = 3 + 2s$. Abbiamo quindi trovato le seguenti equazioni parametriche del piano:

$$\begin{cases} x = t, \\ y = 3 + 2s, \\ z = s, \quad t, s \in \mathbb{R}. \end{cases}$$

Supponiamo ora di volere trovare l'equazione cartesiana del piano π, assegnato attraverso equazioni parametriche. Ciò significa che le coordinate dei punti del piano sono funzioni lineari di due parametri reali t e s: $P = P(t, s)$. L'equazione di π è l'unica equazione lineare, a meno di un fattore moltiplicativo, del tipo

$$ax + by + cz = d,$$

che è soddisfatta dalle coordinate dei punti $P = P(s, t)$, $\forall t, s \in \mathbb{R}$. Un modo per ottenerla è quello di procedere con l'*eliminazione dei parametri* nelle equazioni parametriche, come è mostrato negli esempi seguenti.

Esempio 1.24. Fissiamo nello spazio un sistema di riferimento cartesiano ortonormale $\mathcal{R}(O, \hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}, \hat{\boldsymbol{k}})$.

(1) Troviamo l'equazione cartesiana del piano π di equazione parametrica

$$P = A + t\boldsymbol{u} + s\boldsymbol{v}, \quad t, s \in \mathbb{R},$$

dove $A = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}$, $\boldsymbol{u} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$ e $\boldsymbol{v} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$. Posto $P = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$, abbiamo:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} + t \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + s \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \quad t, s \in \mathbb{R},$$

passando alle coordinate otteniamo

$$\begin{cases} x = 2 + t + s, \\ y = 1 + t, \\ z = -s, \quad t, s \in \mathbb{R}. \end{cases}$$

Osserviamo che possiamo ricavare i parametri t e s rispettivamente dalla seconda e terza equazione:

$$t = y - 1 \quad s = -z,$$

sostituendo queste espressioni nella prima equazione otteniamo:

$$x = 2 + (y - 1) + (-z),$$

svolgendo abbiamo un'equazione lineare in x, y e z:

$$x - y + z = 1,$$

che è soddisfatta da tutti e soli i punti le cui coordinate sono assegnate dalle equazioni parametriche scritte sopra. Questa è l'equazione di π.

(2) Troviamo l'equazione cartesiana del piano π di equazione parametrica

$$P = A + t\boldsymbol{u} + s\boldsymbol{v}, \quad t, s \in \mathbb{R},$$

dove $A = \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}$, $\boldsymbol{u} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$ e $\boldsymbol{v} = \begin{pmatrix} 0 \\ 2 \\ -1 \end{pmatrix}$. Posto $P = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$, abbiamo:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} + t \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} + s \begin{pmatrix} 0 \\ 2 \\ -1 \end{pmatrix}, \quad t, s \in \mathbb{R},$$

passando alle coordinate abbiamo:

$$\begin{cases} x = 2, \\ y = t + 2s, \\ z = t - s, \quad t, s \in \mathbb{R}. \end{cases}$$

Osserviamo che la prima equazione nella rappresentazione parametrica non contiene i parametri t e s:

$$x - 2 = 0,$$

è un'equazione lineare soddisfatta da tutti i punti del piano π, è quindi l'equazione del piano π.

Fissato nello spazio un riferimento cartesiano ortonormale $\mathcal{R}(O, \hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}, \hat{\boldsymbol{k}})$, ci proponiamo ora di rappresentare una retta nello spazio con equazioni cartesiane.

Sia r una retta e siano π_1 e π_2 due piani distinti contenenti la retta r, la retta r risulta essere l'intersezione dei piani π_1 e π_2:

$$r = \pi_1 \cap \pi_2. \tag{1.20}$$

I punti della retta risultano essere i punti le cui coordinate sodisfano entrambe le equazioni lineari dei piani, cioè soddisfano il *sistema lineare* dato rispettivamente dalle equazioni di π_1 e π_2:

$$\begin{cases} a_1 x + b_1 y + c_1 z = d_1 \\ a_2 x + b_2 y + c_2 z = d_2. \end{cases} \tag{1.21}$$

Queste equazioni sono dette *equazioni cartesiane* della retta r.

Osservazione 1.25. Poiché esistono infiniti piani distinti contenenti la retta r, esistono infinite rappresentazioni distinte della retta r con equazioni cartesiane.

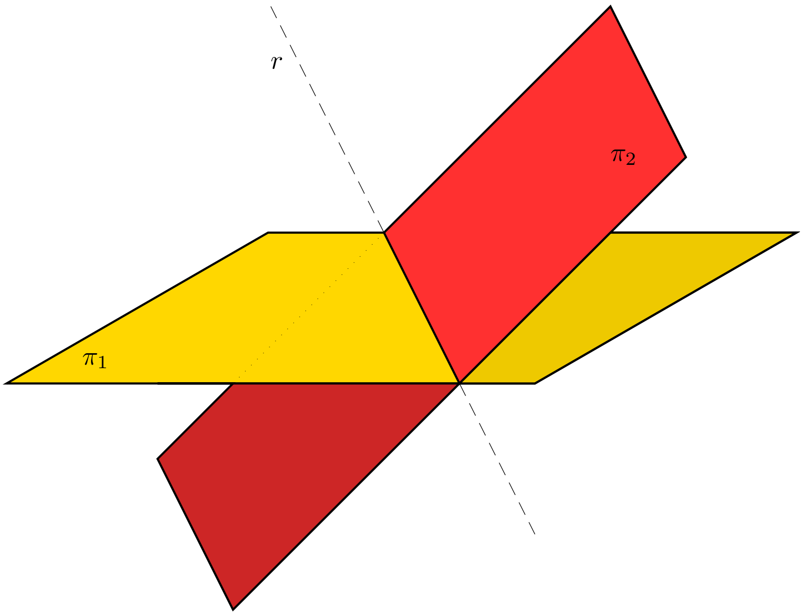

FIGURA 1.20. La retta r è l'intersezione dei piani π_1 e π_2 non paralleli (i piani sono detti *incidenti*).

Fissato un sistema di riferimento cartesiano ortonormale $\mathcal{R}(O, \hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}, \hat{\boldsymbol{k}})$, abbiamo visto che esistono due modi per descrivere una retta nello spazio.

• Attraverso le equazioni cartesiane: ovvero descrivendo la retta come il luogo dei punti le cui coordinate sono le soluzioni di un sistema di due equazioni lineari;

• Attraverso un'equazione parametrica: ovvero descrivendo i punti della retta in funzione di un parametro reale.

Come nel caso del piano, a seconda del problema a cui si è interessati conviene usare una o l'altra descrizione. Ad esempio, se vogliamo stabilire se due rette r e s sono parallele, è più utile avere l'equazione parametrica, da cui leggo facilmente la direzione. È naturale quindi chiedersi:

Domanda: Dato una retta r, come passare da una descrizione parametrica ad una cartesiana e viceversa?

Supponiamo di volere trovare una rappresentazione parametrica per la retta r di equazioni cartesiane

$$\begin{cases} a_1x + b_1y + c_1z = d_1 \\ a_2x + b_2y + c_2z = d_2, \end{cases}$$

poiché la retta è il luogo dei punti le cui coordinate sono le soluzioni del sistema dato, trattandosi di due equazioni lineari in 3 incognite, è chiaro che, in generale, se fissiamo una delle variabili arbitrariamente, possiamo ricavare le altre due variabili, e in questo modo ottenere tutte le soluzioni. Vediamo alcuni esempi.

Esempio 1.26. Fissiamo nello spazio un sistema di riferimento cartesiano ortonormale $\mathcal{R}(O, \hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}, \hat{\boldsymbol{k}})$.

(1) Troviamo una rappresentazione parametrica per la retta r di equazioni cartesiane:

$$\begin{cases} x - y + z = 0 \\ 2x + z = 1. \end{cases}$$

Poniamo $z = t$, sostituiamo nelle equazioni, spostando i termini contenenti t a secondo membro:

$$\begin{cases} x - y = -t \\ 2x = 1 - t. \end{cases}$$

Abbiamo ora un sistema lineare di due equazioni nelle incognite x e y, risolviamo:

$$\begin{cases} y = x + t = \frac{1}{2}(1-t) + t = \frac{1}{2}(1+t) \\ x = \frac{1}{2}(1-t), \end{cases}$$

otteniamo quindi la seguente rappresentazione parametrica di r:

$$\begin{cases} x = \frac{1}{2}(1-t) \\ y = \frac{1}{2}(1+t) \\ z = t, \quad t \in \mathbb{R}. \end{cases}$$

Dalla rappresentazione trovata, osserviamo che un vettore direttore di r è il vettore $\boldsymbol{n} = \begin{pmatrix} -1 \\ 1 \\ 2 \end{pmatrix}$.

(2) Troviamo una rappresentazione parametrica per la retta r di equazioni cartesiane:

$$\begin{cases} x - z = 0 \\ z + 1 = 0. \end{cases}$$

Osserviamo che la variabile y non compare nelle equazioni della retta, significa che tale coordinata può assumere qualsiasi valore reale: poniamo quindi $y = t$. Il sistema dato è un sistema lineare di due equazioni nelle incognite x e z, proviamo a risolverlo:

$$\begin{cases} x = z = -1 \\ z = -1, \end{cases}$$

otteniamo quindi la seguente rappresentazione parametrica di r:

$$\begin{cases} x = -1, \\ y = t, \\ z = -1, \quad t \in \mathbb{R}. \end{cases}$$

Dalla rappresentazione trovata, osserviamo che un vettore direttore di r è il vettore $\boldsymbol{n} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$, quindi la retta r è parallela all'asse y.

Supponiamo ora di volere trovare le equazioni cartesiane della retta r, assegnata attraverso equazioni parametriche. Ciò significa che le coordinate dei punti della retta sono funzioni lineari di un parametro reale t: $P = P(t)$. Le equazioni della retta si ottengono mettendo a sistema le equazioni cartesiane di due qualsiasi piani distinti contenenti la retta r. Un modo per ottenere queste equazioni è quello di procedere con l'*eliminazione del parametro* nelle equazioni parametriche, come è mostrato negli esempi seguenti.

Esempio 1.27. Fissiamo nello spazio un sistema di riferimento cartesiano ortonormale $\mathcal{R}(O, \hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}, \hat{\boldsymbol{k}})$.

(1) Troviamo una rappresentazione cartesiana della retta r di equazione parametrica
$$P = A + t\boldsymbol{v}, \quad t \in \mathbb{R},$$
dove $A = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}$, e $\boldsymbol{v} = \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}$. Posto $P = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$, abbiamo:
$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} + t \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}, \quad t \in \mathbb{R},$$
passando alle coordinate abbiamo:
$$\begin{cases} x = 2 + t, \\ y = 1 + 2t, \\ z = -t \quad t \in \mathbb{R}. \end{cases}$$
Ricaviamo il parametro t da ciascuna equazione:
$$\begin{cases} t = x - 2, \\ t = \frac{1}{2}(y - 1), \\ t = -z, \end{cases}$$
otteniamo quindi le seguenti uguaglianze soddisfatte da tutti e soli i punti della retta r:
$$x - 2 = \frac{1}{2}(y - 1) = -z.$$
Osserviamo che queste uguaglianze definiscono le seguenti equazioni lineari, soddisfatte dai punti della retta r:
$$x - 2 - \frac{1}{2}(y - 1) = 0, \quad x - 2 + z = 0, \quad \frac{1}{2}(y - 1) + z = 0,$$
che rappresentano piani contenenti la retta r. Per rappresentare la retta con equazioni cartesiane basta scegliere due tra queste equazioni,

ad esempio:

$$\begin{cases} x + z = 2, \\ y + 2z = 1. \end{cases}$$

(2) Troviamo una rappresentazione cartesiana della retta r di equazione parametrica

$$P = A + t\boldsymbol{v}, \quad t \in \mathbb{R},$$

dove $A = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$, e $\boldsymbol{v} = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$. Posto $P = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$, abbiamo:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} + t \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \quad t \in \mathbb{R},$$

passando alle coordinate abbiamo:

$$\begin{cases} x = 1 + t, \\ y = 1 - t, \\ z = -1, \quad t \in \mathbb{R}. \end{cases}$$

Osserviamo che nella terza equazione non compare il parametro t, essa è quindi un'equazione lineare

$$z + 1 = 0,$$

che è soddisfatta dai punti della retta r: quindi definisce un piano contenente la retta. Ricaviamo ora il parametro t da ciascuna delle rimanenti equazioni:

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

otteniamo quindi la seguente uguaglianza soddisfatte dai punti della retta r:

$$x - 1 = 1 - y.$$

Una rappresentazione della retta r con equazioni cartesiane è la seguente:

$$\begin{cases} x + y = 2, \\ z + 1 = 0. \end{cases}$$

6. Posizioni reciproche tra rette e piani.

Abbiamo visto come rappresentare le rette e i piani nello spazio, passiamo ora a discutere le posizioni reciproche in cui si possono trovare.

Applicazione 1.28. *Posizione reciproca di due piani*
Fissato nello spazio un sistema di riferimento $\mathcal{R}(O, \hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}, \hat{\boldsymbol{k}})$ cartesiano ortonormale, consideriamo i seguenti piani:

$$\pi\colon ax + by + cz = d \quad \pi'\colon a'x + b'y + c'z = d',$$

siano $\boldsymbol{n}, \boldsymbol{n}' \in \mathbb{E}^3_O$ due vettori normali rispettivamente ai piani π e π'.
• I piani π e π' sono *paralleli* se e solo se hanno la stessa retta normale per O, se e solo se i vettori $\boldsymbol{n}$ e $\boldsymbol{n}'$ generano la stessa retta, cioè $\boldsymbol{n}' \in \mathrm{Span}(\boldsymbol{n})$. Ciò significa che esiste un numero reale non nullo k take che

$$a' = ka \quad b' = kb \quad c' = kc.$$

Se inoltre risulta $d' = kd$, allora l'equazione di π' è:

$$k(ax + by + cz) = kd,$$

e quindi rappresenta il piano π.
• I piani π e π' sono *coincidenti* se esiste $k \in \mathbb{R}$, $k \neq 0$, tale che

$$a' = ka \quad b' = kb \quad c' = kc \quad d' = kd.$$

• Se i piani π e π' non sono paralleli, non sono coincidenti, allora la loro intersezione è una retta:

$$r = \pi \cap \pi'.$$

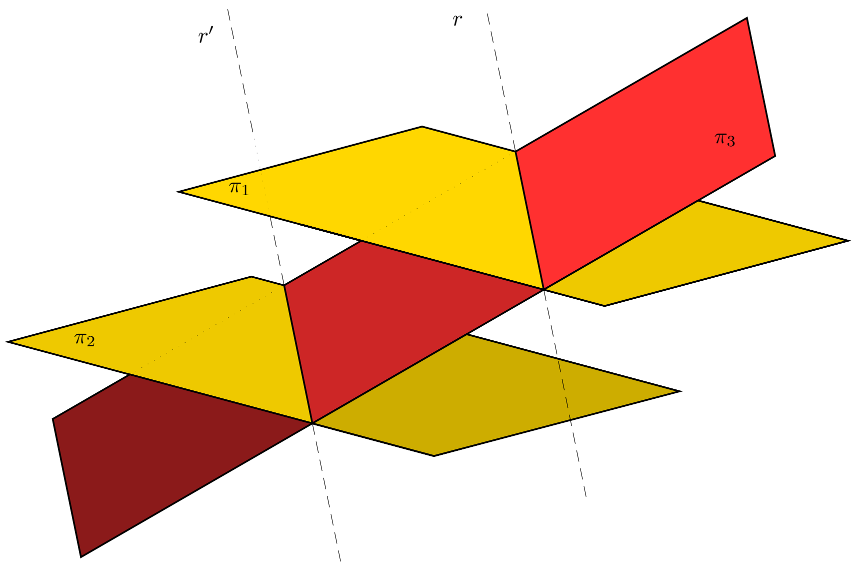

Figura 1.21. Esempi di posizione reciproca di due piani. I piani π_1 e π_2 sono paralleli; i piani π_1 e π_3 sono incidenti, e si intersecano secondo la retta r. Conseguentemente, anche π_2 e π_3 sono incidenti, e retta la $r' = \pi_2 \cap \pi_3$ è parallela ad r.

Esempio 1.29. Fissiamo nello spazio un sistema di riferimento $\mathcal{R}(O, \hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}, \hat{\boldsymbol{k}})$ cartesiano ortonormale.

(1) Verifichiamo che i piani di equazione $z = a$, $a \in \mathbb{R}$, sono tutti paralleli al piano x, y.
Osserviamo che il piano x, y è il piano $\mathrm{Span}(\hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}})$, ovvero il piano per O perpendicolare al vettore $\hat{\boldsymbol{k}}$. Un vettore normale per il piano di equazione $z = a$ è il vettore $\boldsymbol{n} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \hat{\boldsymbol{k}}$, per cui concludiamo che i piani sono paralleli.

(2) Scriviamo l'equazione cartesiana del piano π parallelo al piano α di equazione $2x - y + 3z = 5$ e passante per il punto $A = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$.
Poiché il piano π è parallelo ad α, possiamo prendere come vettore normale al piano il vettore $\boldsymbol{n} = \begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix}$. L'equazione di π risulta quindi del tipo:
$$2x - y + 3z = d,$$
con $d \in \mathbb{R}$. Determiniamo ora per quale valore d il piano contiene il punto A, sostituendo le coordinate di A nell'equazione di π otteniamo
$$2(1) - (-2) + 3(1) = d,$$
da cui ricaviamo $d = 7$. L'equazione del piano è la seguente:
$$2x - y + 3z = 7.$$

Applicazione 1.30. *Posizione reciproca di due rette.*
Ricordiamo che nel piano due rette distinte o sono parallele oppure si intersecano in un solo punto. Questa proprietà non vale per due rette nello spazio. Fissiamo nello spazio un sistema di riferimento $\mathcal{R}(O, \hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}, \hat{\boldsymbol{k}})$ cartesiano ortonormale. Consideriamo due rette distinte r_1 e r_2 e siano $\boldsymbol{v}_1, \boldsymbol{v}_2 \in \mathbb{E}^3_O$ rispettivamente vettori direttori di r_1 e r_2.
• Le rette r_1 e r_2 sono *parallele* se e solo se hanno la stessa direzione, se e solo se i loro vettori direttori generano la stessa retta, cioè $\boldsymbol{v}_2 \in \mathrm{Span}(\boldsymbol{v}_1)$.
• Le rette r_1 e r_2 sono *incidenti* se e solo se si intersecano in un unico punto:
$$r_1 \cap r_2 = \{P\}.$$
• Le rette r_1 e r_2 sono *sghembe* se non sono parallele e non sono incidenti.
• Le rette r_1 e r_2 sono *complanari* se e solo se esiste un piano π che le contiene.

Osservazione 1.31.
- Consideriamo due piani paralleli, π_1 e π_2. Siano r_1 e r_4 due rette di direzioni diverse contenute rispettivamente nei piani π_1 e π_2. Osserviamo che le rette r_1 e r_4 sono sghembe: infatti, non sono parallele, perché hanno direzioni diverse, e non sono incidenti perché i due piani sono paralleli, quindi non hanno punti di intersezione: $\pi_1 \cap \pi_2 = \emptyset$.

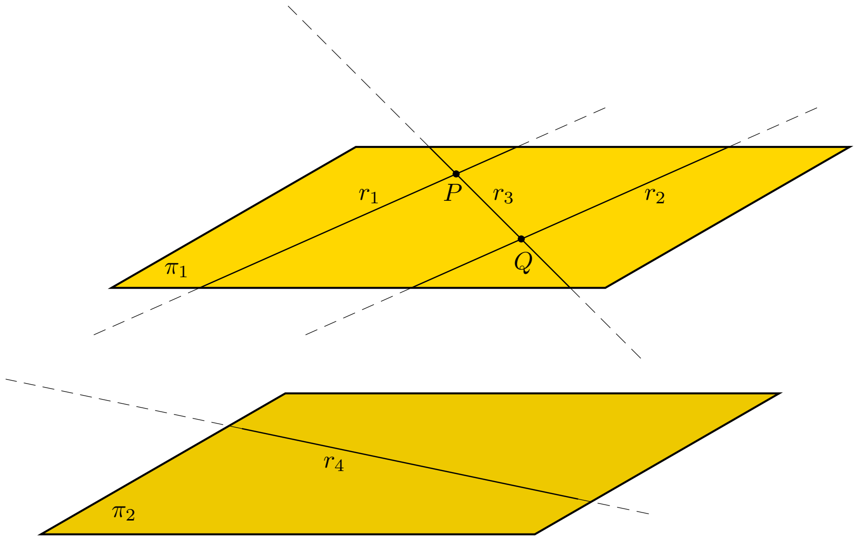

FIGURA 1.22. Esempi di posizione reciproche fra rette. I piani π_1 e π_2 sono paralleli; $r_1, r_2, r_3 \in \pi_1$; $r_4 \in \pi_2$. Le rette r_1 ed r_2 non hanno punti in comune; sono *parallele*, essendo nello stesso piano. Le rette r_1 ed r_3 si intersecano in P: sono pertanto incidenti; conseguentemente, r_2 ed r_3 sono *incidenti*; la loro intersezione è il punto Q. Le rette r_1 ed r_4 non hanno punti in comune e no esiste un piano le le contiene entrambe, sono, pertanto, *sghembe*.

- Due rette distinte r_1 e r_2 sono complanari se e solo se o sono parallele o sono incidenti.

ESEMPIO 1.32. Fissiamo nello spazio un sistema di riferimento $\mathcal{R}(O, \hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}, \hat{\boldsymbol{k}})$ cartesiano ortonormale.

(1) Scriviamo una rappresentazione cartesiana per la retta r passante per O parallela alla retta s di equazioni cartesiane $\begin{cases} x + 2y = 0 \\ x + z = 2 \end{cases}$.
Troviamo una rappresentazione parametrica per la retta s, poniamo $x = t$ e otteniamo:
$$\begin{cases} 2y = -t \\ z = 2 - t, \end{cases}$$
abbiamo quindi trovato una rappresentazione parametrica per la retta s:
$$\begin{cases} x = t \\ y = \frac{1}{2}t \\ z = 2 - t, \quad t \in \mathbb{R}. \end{cases}$$

La direzione della retta r è data dal vettore $\boldsymbol{v} = \begin{pmatrix} 2 \\ 1 \\ -2 \end{pmatrix}$, per cui la retta r è la retta per O di direzione $\boldsymbol{v}$, di equazione parametrica:

$$P = O + \alpha \boldsymbol{v}, \quad \alpha \in \mathbb{R},$$

da cui ricaviamo:

$$\begin{cases} x = 2\alpha \\ y = \alpha \\ z = -2\alpha, \quad \alpha \in \mathbb{R}. \end{cases}$$

Per trovare due equazioni cartesiane per la retta r, ricaviamo il parametro α dalla seconda equazione $\alpha = y$ e sostituiamo nelle rimanenti equazioni:

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

(2) Siano r la retta per il punto $A = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}$ di direzione $\boldsymbol{v} = \hat{\boldsymbol{\imath}} - \hat{\boldsymbol{\jmath}} + 2\hat{\boldsymbol{k}}$ e s la retta di equazioni cartesiane: $\begin{cases} x + y + z = 0 \\ x - z = 1 \end{cases}$. Ci proponiamo di stabilire la posizione reciproca delle rette r e s.
Troviamo una rapresentazione parametrica per la retta s, poniamo $z = t$ e otteniamo:

$$\begin{cases} x + y = -t \\ x = 1 + t, \end{cases}$$

sostituendo nella prima equazione al posto di x l'espressione trovata, abbiamo:

$$\begin{cases} y = -1 - 2t \\ x = 1 + t, \end{cases}$$

abbiamo quindi trovato una rappresentazione parametrica per la retta s:

$$\begin{cases} x = 1 + t \\ y = -1 - 2t \\ z = t \quad t \in \mathbb{R}. \end{cases}$$

Ricaviamo immediatamente che la direzione di s è data dal vettore

$$\boldsymbol{w} = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix},$$

inoltre si ha: $\operatorname{Span}(\boldsymbol{w}) = \{\begin{pmatrix} k \\ -2k \\ k \end{pmatrix}, k \in \mathbb{R}\}$. Osserviamo che $\boldsymbol{v} \in \operatorname{Span}(\boldsymbol{w})$ se e solo se esiste $k \neq 0$ tale che

$$\begin{cases} k = 1 \\ -2k = -1 \\ k = 2. \end{cases}$$

Poiché il sistema non ha soluzioni, possiamo concludere che le due rette non sono parallele.
Per stabilire se sono incidenti, troviamo una rappresentazione cartesiana di r. La retta r ha equazione parametrica:

$$P = A + t\boldsymbol{v}, \quad t \in \mathbb{R},$$

posto $P = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$, abbiamo:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix} + t \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}, \quad t \in \mathbb{R},$$

passando alle coordinate abbiamo:

$$\begin{cases} x = 1 + t, \\ y = 2 - t, \\ z = 2 + 2t, \quad t \in \mathbb{R}. \end{cases}$$

Per stabilire se siste un punto di intersezione tra le rette, vediamo se esiste un punto sulla retta r che verifica le equazioni di s. Sostituendo le coordinate del punto $P = P(t)$ nelle equazioni cartesiane di s otteniamo il sistema nell'incognita t:

$$\begin{cases} (1 + t) + (2 - t) + (2 + 2t) = 0 \\ (1 + t) - (2 + 2t) = 1, \end{cases}$$

semplificando ottieniamo

$$\begin{cases} 2t + 5 = 0 \\ t + 2 = 0, \end{cases}$$

che non ammette soluzioni. Le rette sono quindi sghembe.
In alternativa, se la retta r è assegnata con equazioni cartesiane si può precedere nel seguente modo. Sia r la retta di equazioni:

$$\begin{cases} x + y = 3 \\ 2x - z = 0, \end{cases}$$

le rette r e s sono incidenti se e solo se esiste un punto che verifica le equazioni di entrambe le rette. Ciò equivale a richiedere che il seguente sistema lineare ammette una soluzione:

$$\begin{cases} x+y+z=0 \\ x-z=1 \\ x+y=3 \\ 2x-z=0. \end{cases}$$

Facendo la differenza tra la seconda e la quarta equazione , facendo la differenza tra la prima equazione e la terza otteniamo:

$$\begin{cases} z=-3 \\ x=-1 \\ x+y=3 \\ 2x-z=0. \end{cases}$$

Sostituendo nell'ultima equazione otteniamo $1=0$, che è impossibile. Possiamo quindi concludere che le rette sono sghembe.

Applicazione 1.33. *Posizione reciproca retta-piano.*
Fissiamo nello spazio un sistema di riferimento $\mathcal{R}(O,\hat{\boldsymbol{\imath}},\hat{\boldsymbol{\jmath}},\hat{\boldsymbol{k}})$ cartesiano ortonormale. Consideriamo una retta r ed un piano π, siano $\boldsymbol{v}\in\mathbb{E}^3_O$ un vettore direttore di r ed $\boldsymbol{n}\in\mathbb{E}^3_O$ un vettore normale a π, infine sia $\{\boldsymbol{u}_1,\boldsymbol{u}_2\}$ una coppia di vettori linearmente indipendenti che danno la giacitura di π.

- La retta r ed il piano π sono *incidenti* se e solo se si intersecano in un unico punto:
$$r\cap\pi=\{P\}.$$
Ciò avviene se solo se $\boldsymbol{v}\notin\mathrm{Span}(\boldsymbol{u}_1,\boldsymbol{u}_2)$, se e solo se $\langle\boldsymbol{v},\boldsymbol{n}\rangle\neq 0$.
- La retta r è *perpendicolare* al piano π se e solo se la direzione di r coincide con la direzione normale al piano, cioè $\boldsymbol{v}\in\mathrm{Span}(\boldsymbol{n})$.
- La retta r e il piano π sono *paralleli* se e solo $\boldsymbol{v}\in\mathrm{Span}(\boldsymbol{u}_1,\boldsymbol{u}_2)$ e $r\cap\pi=\emptyset$. La prima condizione equivale a $\boldsymbol{v}\perp\boldsymbol{n}$, cioè $\langle\boldsymbol{v},\boldsymbol{n}\rangle=0$.
- La retta r è *contenuta* nel piano π se e solo se $\boldsymbol{v}\in\mathrm{Span}(\boldsymbol{u}_1,\boldsymbol{u}_2)$ e $r\cap\pi\neq\emptyset$.

Esempio 1.34. Fissiamo nello spazio un sistema di riferimento $\mathcal{R}(O,\hat{\boldsymbol{\imath}},\hat{\boldsymbol{\jmath}},\hat{\boldsymbol{k}})$ cartesiano ortonormale.

(1) Consideriamo il piano π di equazione $2x-y+2z=1$ e la retta r passante per O di direzione $\boldsymbol{v}=\hat{\boldsymbol{\imath}}-\hat{\boldsymbol{\jmath}}$. Verifichiamo che la retta r ed il piano π sono incidenti, determiniamo infine il punto di intersezione.
Un vettore direttore di r è il vettore $\boldsymbol{v}=\begin{pmatrix}-1\\1\\0\end{pmatrix}$, un vettore normale al piano è il vettore $\boldsymbol{n}=\begin{pmatrix}2\\-1\\2\end{pmatrix}$. Calcoliamo il prodotto scalare
$$\langle\boldsymbol{v},\boldsymbol{n}\rangle=-1(2)+1(-1)+0(2)=-3\neq 0,$$

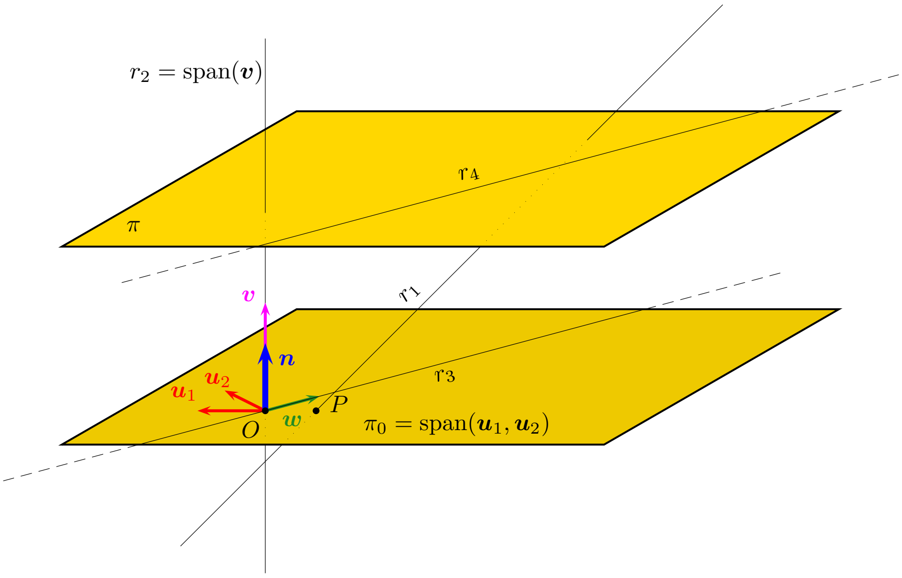

FIGURA 1.23. Esempi di posizione reciproca fra rette e piani. Il piano π_0 passante per l'origine è lo span dei vettori $(\boldsymbol{u}_1, \boldsymbol{u}_2)$; il piano π è parallelo a questo, ma non passante per l'origine. La retta r_1 è incidente al piano π_0: lo interseca nel solo punto P (conseguentmente, è anche incidente al piano π). La retta $r_2 = \mathrm{Span}(\boldsymbol{v})$ è ortogonale ai due piani: il suo direttore e la normale ai piani $\boldsymbol{n}$ sono paralleli (ossia, $\mathrm{Span}(\boldsymbol{v}) = \mathrm{Span}(\boldsymbol{n})$). La retta r_3 è contenuta in π_0, poiché il suo direttore $\boldsymbol{w} \in \mathrm{Span}(\boldsymbol{u}_1, \boldsymbol{u}_2)$ ed ha almeno un punto in comune con il piano. La retta r_4 (parallela ad r_3) è parallela a π_0: il suo direttore è ancora $\boldsymbol{w} \in \mathrm{Span}(\boldsymbol{u}_1, \boldsymbol{u}_2)$, ma non ha alcun punto in comune con il piano (essa è, infatti, contenuta in π).

per cui concludiamo che la retta ed il piano sono incidenti.
In questo caso è più utile rappresentare la retta r con equazione parametrica:

$$P = O + t\boldsymbol{v}, \quad t \in \mathbb{R},$$

passando alle coordinate otteniamo:

$$\begin{cases} x = -t \\ y = t \\ z = 0, \quad t \in \mathbb{R}. \end{cases}$$

Determinare il punto di intersezione di r e π equivale a cercare per quale valore di t il corrispondente punto della retta r appartiene al piano π. Sostituiamo nell'equazione del piano π le rappresentazioni parametriche trovate in funzione di t:

$$2(-t) - (t) + 2(0) = 1,$$

otteniamo un'equazione lineare in t:

$$-3t = 1,$$

da cui ricaviamo $t = -\frac{1}{3}$. Il punto di intersezione è quindi il punto della retta corrispondente al parametro $t = -\frac{1}{3}$:

$$P = \begin{pmatrix} \frac{1}{3} \\ -\frac{1}{3} \\ 0 \end{pmatrix}.$$

In alternativa, si può procedere nel seguente modo: determinare le equazioni cartesiane di r, risolvere il sistema lineare di 3 equazioni e 3 incognite dato dalle equazioni di r e π.

(2) Determiniamo la retta r passante per il punto $A = \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$ e ortogonale al piano π di equazione $2x - y + 3z = 8$.
Un vettore normale al piano è $\boldsymbol{n} = \begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix}$. La retta r è la retta per A di direzione n, perciò ha equazione parametrica:

$$P = A + t\boldsymbol{n}, \quad t \in \mathbb{R}.$$

Passando alle coordinate abbiamo:

$$\begin{cases} x = -1 + 2t \\ y = 1 - t \\ z = 1 + 3t, \quad t \in \mathbb{R}. \end{cases}$$

Per trovare due equazioni cartesiane della retta r, ricaviamo il parametro t dalla seconda equazione $t = 1 - y$ e sostituiamo nelle rimanenti:

$$\begin{cases} x + 2y = 1 \\ 3y + z = 4. \end{cases}$$

(3) Determiniamo il piano α passante per il punto $A = \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix}$ e ortogonale alla retta s equazioni $\begin{cases} x + y = 0 \\ x - y + z = 1 \end{cases}$.
Innanzittutto troviamo una rappresentazione parametrica per la retta s, poniamo $y = t$ e sostituiamo nelle due equazioni:

$$\begin{cases} x = -t \\ z = y - x + 1 = t - (-t) + 1 = 2t + 1, \end{cases}$$

da cui ricaviamo la rappresentazione seguente:

$$\begin{cases} x = -t \\ y = t \\ z = 1 + 2t, \quad t \in \mathbb{R}. \end{cases}$$

Osserviamo che la direzione della retta s è data dal vettore $v = \begin{pmatrix} -1 \\ 1 \\ 2 \end{pmatrix}$, quindi il piano α è il piano per A con direzione normale $\boldsymbol{v}$:

$$-1(x + 1) + 1(y - 2) + 2(z - 1) = 0,$$

cioè il piano di equazione cartesiana $x - y - 2z = -5$.

7. Fasci di piani.

Fissiamo nello spazio un sistema di riferimento $\mathcal{R}(O, \hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}, \hat{\boldsymbol{k}})$ cartesiano ortonormale.

Definizione 1.15. L'insieme $\mathcal{F}_r$ dei piani contenenti una retta r è detto *fascio proprio di piani di sostegno* r (Figura 1.24); l'insieme $\mathcal{F}_n$ dei piani aventi la stessa direzione normale $\boldsymbol{n}$ è detto *fascio improprio di piani* (Figura 1.25).

Proposizione 1.13. *Sia $\mathcal{F}_r$ il fascio proprio con sostegno la retta r, i piani del fascio sono tutti e soli quelli la cui equazione si può scrivere*

$$\lambda(a_1x + b_1y + c_1z + d_1) + \mu(a_2x + b_2y + c_2z + d_2) = 0 \quad (\lambda, \mu) \neq (0, 0), \tag{1.22}$$

dove $a_1x + b_1y + c_1z + d_1 = 0$ e $a_2x + b_2y + c_2z + d_2 = 0$ rappresentano due piani distinti π_1 e π_2 appartenenti al fascio.

Osserviamo che $\forall(\lambda, \mu) \neq (0, 0)$, l'equazione scritta rappresenta un piano che contiene la retta r, poiché entrambe le due equazioni sono soddisfatte dai punti della retta r. Mostriamo infine che ogni piano del fascio può essere scritto in questo modo. Sia $\alpha \in \mathcal{F}_r$, scegliamo un punto $A \in \alpha$ tale che $A \notin r$, osserviamo che α è l'unico piano del fascio contenente A. Sostituiamo le coordinate di A nell'equazione (1.22), poiché $A \notin r$, abbiamo un'equazione lineare non identicamente nulla:

$$\lambda(a_1x_A + b_1y_A + c_1z_A + d_1) + \mu(a_2x_A + b_2y_A + c_2z_A + d_2) = 0.$$

Questa equazione ammette infinite soluzioni $(\lambda, \mu) = \rho(\lambda_0, \mu_0)$, che danno infinite equazioni tutte proporzionali, quindi definscono un unico piano del fascio contenente il punto A: il piano α.

Proposizione 1.14. *Sia $\mathcal{F}_n$ il fascio improprio con direzione normale $\boldsymbol{n}$, i piani del fascio sono tutti e soli quelli la cui equazione si può scrivere*

$$ax + by + cz + d = 0, \tag{1.23}$$

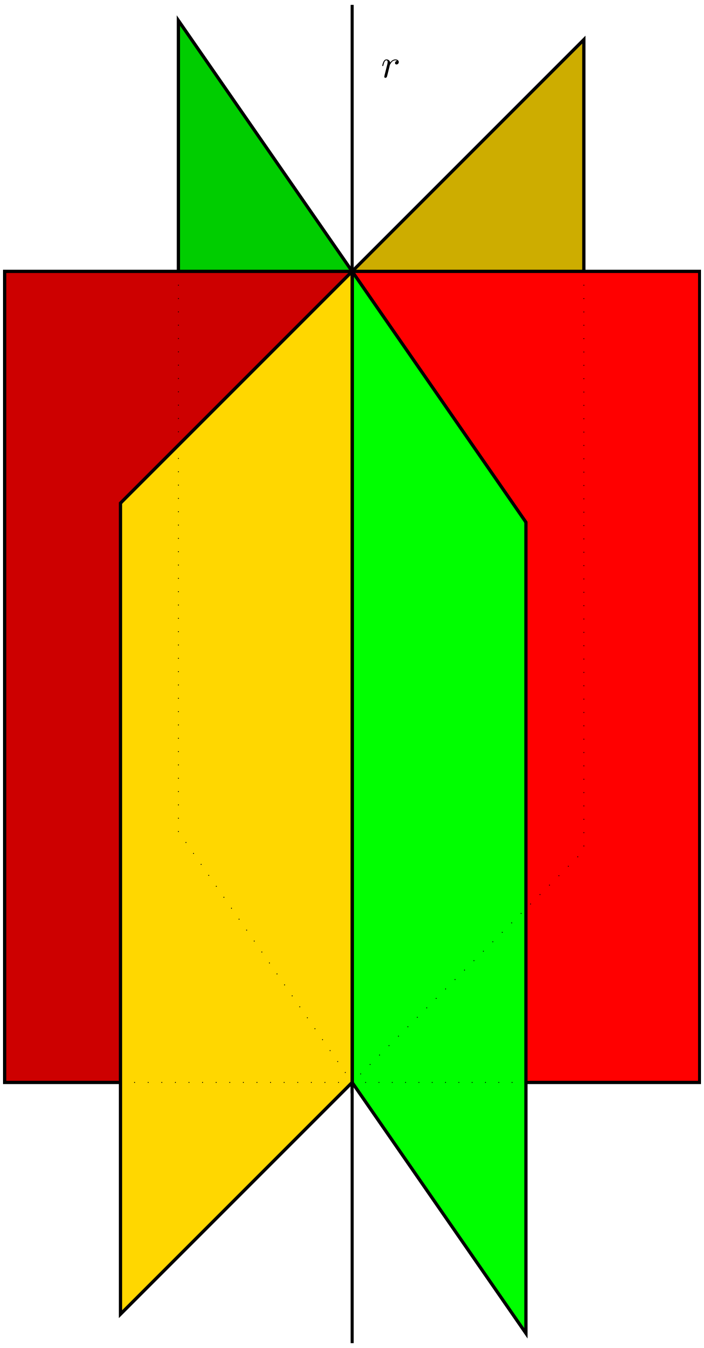

FIGURA 1.24. Rappresentazione schematica di un fascio proprio di piani avente per sostegno la retta r; sono raffigurati alcuni dei piani del fascio, tutti contenenti la retta r.

dove $\boldsymbol{n} = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$, $d \in \mathbb{R}$.

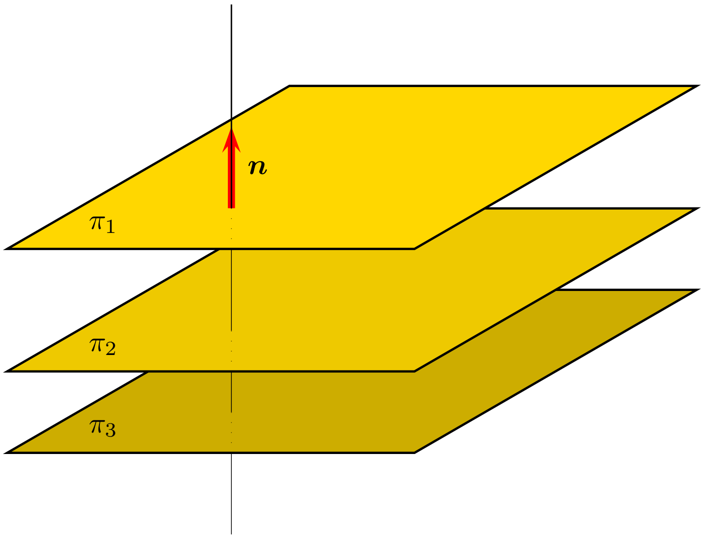

FIGURA 1.25. Rappresentazione schematica di un fascio improprio di piani, caratterizzati dalla medesima normale $\boldsymbol{n}$; la retta $r' = \mathrm{Span}(\boldsymbol{n})$ è ortogonale a tutti i piani del fascio.

ESEMPIO 1.35. Fissiamo nello spazio un sistema di riferimento cartesiano ortonormale $\mathcal{R}(O, \hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}, \hat{\boldsymbol{k}})$.

(1) Consideriamo la retta r passante per $A = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ e di direzione $\boldsymbol{v} = \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}$. Determiniamo il piano π contenente la retta r ed il punto $B = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}$.

Troviamo una rappresentazione parametrica per la retta r:

$$P = A + t\boldsymbol{v}, \quad t \in \mathbb{R},$$

passando alle coordinate abbiamo:

$$\begin{cases} x = 1 + t \\ y = t \\ z = -2t, \quad t \in \mathbb{R}. \end{cases}$$

Sostituiamo $t = y$ nella prima e terza equazione, otteniamo le equazioni di due piani contenenti la retta r:

$$x - y = 1 \quad z + 2y = 0.$$

Consideriamo ora il fascio $\mathcal{F}_r$ di piani contenenti la retta r, un piano del fascio ha equazione:

$$\lambda(x-y-1)+\mu(z+2y)=0, \quad (\lambda,\mu)\neq(0,0).$$

Cerchiamo il piano del fascio che contiene il punto B, sostituiamo le coordinate del punto B:

$$\lambda(1-2-1)+\mu(2+4)=0,$$

otteniamo l'equazione lineare omogenea:

$$-2\lambda+6\mu=0,$$

che esplicitiamo nella seguente:

$$\lambda=3\mu,$$

le cui soluzioni sono:

$$(\lambda,\mu)=\rho(3,1).$$

Il piano π ha quindi equazione cartesiana:

$$3(x-y-1)+1(z+2y)=0,$$

cioè $3x-y+z=3$.

(2) Determiniamo l'equazione del piano π contenente la retta s di equazioni $x-y=y-2z=0$ e perpendicolare al piano α di equazione $x-y+3z=7$.
Consideriamo il fascio $\mathcal{F}_s$ di piani contenenti la retta s, un piano del fascio ha equazione:

$$\lambda(x-y)+\mu(y-2z)=0, \quad (\lambda,\mu)\neq(0,0).$$

Cerchiamo il piano del fascio perpendicolare al piano α: osseviamo che un piano $\pi\in\mathcal{F}_s$ è perpendicolare ad α se e solo se le loro direzioni normali sono perpendicolari. Indicati con $\boldsymbol{n}_\pi$ e $\boldsymbol{n}_\alpha$ rispettivamente due vettori normali ai piani π e α, risulta $\pi\perp\alpha$ se e solo se $<\boldsymbol{n}_\pi,\boldsymbol{n}_\alpha>=0$.
L'equazione di π è:

$$(\lambda)x+(\mu-\lambda)y+(-2\mu)z=0,$$

da cui otteniamo $\boldsymbol{n}_\pi=\begin{pmatrix}\lambda\\ \mu-\lambda\\ -2\mu\end{pmatrix}$, infine abbiamo $\boldsymbol{n}_\alpha=\begin{pmatrix}1\\ -1\\ 3\end{pmatrix}$. Cerchiamo ora per quali valori dei parametri (λ,μ) il prodotto scalare è zero:

$$<\boldsymbol{n}_\pi,\boldsymbol{n}_\alpha>=1(\lambda)-1(\mu-\lambda)+3(-2\mu)=0,$$

otteniamo l'equazione lineare omogenea:

$$2\lambda-7\mu=0,$$

le cui soluzioni sono: $(\lambda,\mu)=\rho(7,2)$. Il piano π ha quindi equazione cartesiana:

$$7(x-y)+2(y-2z)=0,$$

cioè $7x-5y-4z=0$.

(3) Stabiliamo se le seguenti rette sono complanari:

$$r\colon \begin{cases} x+y=0 \\ z-1=0 \end{cases} \qquad s\colon \begin{cases} 2x-z=0 \\ y+1=0. \end{cases}$$

Le rette r e s sono complanari se e solo se esiste un piano che le contiene. Consideriamo il fascio $\mathcal{F}_r$ di piani contenenti la retta r, un piano del fascio ha equazione:

$$\lambda(x+y)+\mu(z-1)=0, \quad (\lambda,\mu)\neq(0,0).$$

Scegliamo due punti distinti sulla retta s, troviamo dapprima una rappresentazione parametrica per la retta s:

$$\begin{cases} x=t \\ y=-1 \\ z=2t, t\in\mathbb{R}. \end{cases}$$

Ponendo $t=0$ e $t=1$ otteniamo i punti corrispondenti su s:

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

Osserviamo che esiste un piano $\pi\in\mathcal{F}_r$ contenente la retta s se e solo se esiste un piano $\pi\in\mathcal{F}_r$ passante per i punti A e B. Imponendo il passaggio per i punti A e B, otteniamo il seguente sistema:

$$\begin{cases} \lambda(-1)+\mu(-1)=0 \\ \mu=0 \end{cases}$$

che ammette l'unica soluzione: $(\lambda,\mu)=(0,0)$, che non rappresenta alcun piano del fascio. Possiamo concludere che non esiste un piano π del fascio contenente la retta s, quindi le rette r e s non sono complanari.

8. Distanza punto-retta, punto-piano.

Consideriamo un piano π ed un punto A. La distanza del punto A dal piano π è definita nel seguente modo:

$$d(A,\pi)=\inf\{|\overrightarrow{AP}|,\ P\in\pi\}.$$

Si verifica che risulta:

$$d(A,\pi)=|\overrightarrow{AH}|,$$

dove H è il piede della retta perpendicolare al piano π passante per A (Figura 1.26).

Osserviamo che $d(A,\pi)=0$ se e solo se $A\in\pi$.

Consideriamo ora una retta r ed un punto A. La distanza del punto A dalla retta r è definita nel seguente modo:

$$d(A,r)=\inf\{|\overrightarrow{AP}|,\ P\in r\}.$$

Si verifica che risulta:

$$d(A,r)=|\overrightarrow{AH}|,$$

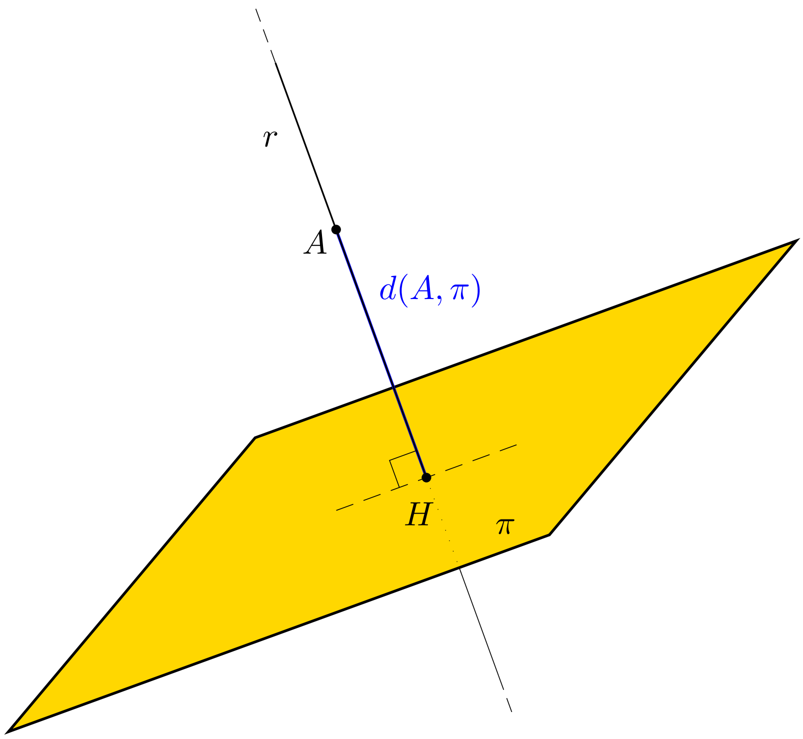

FIGURA 1.26. La distanza $d(A, \pi)$ fra il punto A ed il piano π è la lunghezza del segmento aventi per estremi A ed il piede H della retta r passante per A e perpendicolare a π.

dove H è il punto di intersezione tra la retta r ed il piano π passante per A perpendicolare alla retta r (Figura 1.27).

Ricordiamo che nella geometria analitica del piano, esiste una formula che consente di calcolare la distanza di un punto da una retta, tale formula può essere generalizzata anche alla geometria dello spazio per la distanza punto-piano.

Fissiamo nello spazio un sistema di riferimento $\mathcal{R}(O, \hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}, \hat{\boldsymbol{k}})$ cartesiano ortonormale, siano π il piano di equazione $ax + by + cz + d = 0$ e $A = \begin{pmatrix} x_A \\ y_A \\ z_A \end{pmatrix}$, risulta:

$$d(A, \pi) = \frac{|ax_A + by_A + cz_A + d|}{\sqrt{a^2 + b^2 + c^2}}.$$

Verifichiamo la formula. Sia r la retta per A perpendicolare al piano π, r ha la direzione del vettore $\boldsymbol{n}$ normale al piano, per cui r ha rappresentazione parametrica:

$$\begin{cases} x = x_A + ta, \\ y = y_A + tb, \\ z = z_A + tc, \quad t \in \mathbb{R}. \end{cases}$$

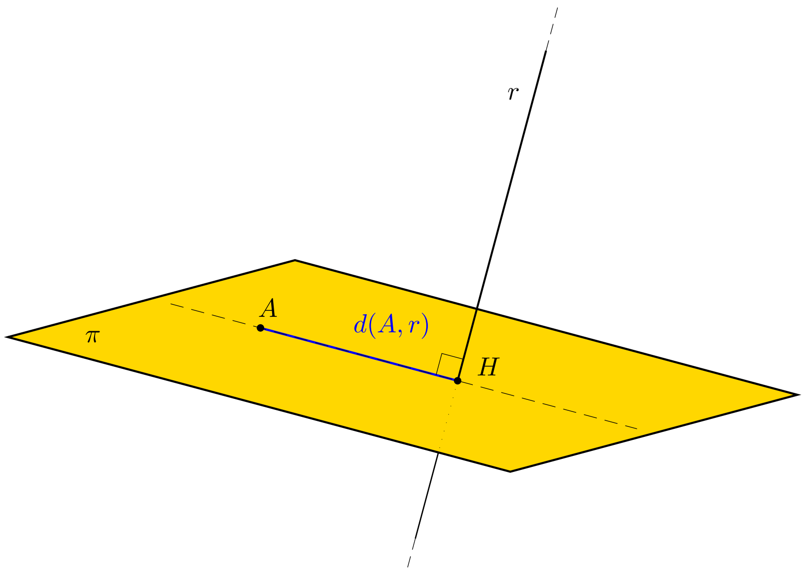

FIGURA 1.27. La distanza $d(A,r)$ fra il punto A ed la retta r è la lunghezza del segmento aventi per estremi A ed il punto H di intersezione fra la retta r ed il piano π passante per A e perpendicolare a r.

Il punto H è il punto di intersezione di r e π, sostitendo nell'equazione di π le coordinate di un punto generico di r, abbiamo:

$$a(x_A + ta) + b(y_A + tb) + c(z_A + tc) + d = 0,$$
$$(a^2 + b^2 + c^2)t + ax_A + by_A + cz_A + d = 0,$$

otteniamo quindi un'equazione lineare in t, la cui soluzione è:

$$t_H = \frac{-(ax_A + by_A + cz_A + d)}{(a^2 + b^2 + c^2)},$$

valore del parametro che corrisponde al punto H. Si ha quindi:

$$H = \begin{pmatrix} x_A + at_H \\ y_A + bt_H \\ z_A + ct_H \end{pmatrix},$$

$$\begin{aligned} |\overrightarrow{AH}| &= \sqrt{(x_A - x_H)^2 + (y_A - y_H)^2 + (z_A - z_H)^2} \\ &= |t_H|\sqrt{a^2 + b^2 + c^2} \\ &= \frac{|ax_A + by_A + cz_A + d|}{\sqrt{a^2 + b^2 + c^2}}. \end{aligned}$$

Per il calcolo della distanza punto-retta, non esiste una formula analoga che consente di evitare il procedimento, si procede quindi ogni volta determinando analiticamente il punto H.

ESEMPIO 1.36. Fissiamo nello spazio un sistema di riferimento cartesiano ortonormale $\mathcal{R}(O, \hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}, \hat{\boldsymbol{k}})$.

(1) Calcoliamo la distanza del punto $A = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ dal piano π di equazione: $x + y - z = 4$.

Applicando la formula si ha:

$$d(A, \pi) = \frac{|ax_A + by_A + cz_A + d|}{\sqrt{a^2 + b^2 + c^2}} = \frac{|1 + 1 - 1 - 4|}{\sqrt{1 + 1 + 1}} = \sqrt{3}.$$

(2) Calcoliamo la distanza del punto $A = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}$ dalla retta r di equazioni: $x - y = z = 0$.

La retta r ha equazioni parametriche:

$$\begin{cases} x = t \\ y = t, \\ z = 0, \quad t \in \mathbb{R}. \end{cases}$$

La direzione di r è data dal vettore $\boldsymbol{v} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$. Consideriamo il piano π passante per il punto A e perpendicolare alla retta r, l'equazione di π è:

$$1(x - x_A) + 1(y - y_A) + 0(z - z_A) = 0,$$

sostituendo le coordinate di A otteniamo

$$x + y = 3.$$

Il punto H è l'intersezione di r con il piano π, sostituendo nell'equazione di π le coordinate di un punto generico di r abbiamo:

$$t + t = 3,$$

risolvendo l'equazione otteniamo $t_H = \frac{3}{2}$, valore dl parametro che corrsiponde al punto H:

$$H = \begin{pmatrix} \frac{3}{2} \\ \frac{3}{2} \\ 0 \end{pmatrix},$$

e quindi

$$\begin{aligned} d(A, r) = |\overrightarrow{AH}| &= \sqrt{(x_A - x_H)^2 + (y_A - y_H)^2 + (z_A - z_H)^2} \\ &= \sqrt{\frac{1}{4} + \frac{1}{4} + 1} = \frac{\sqrt{6}}{2}. \end{aligned}$$

9. Superficie sferica.

Ricordiamo che nel piano una circonferenza di centro C e raggio R è il luogo geometrico dei punti del piano che hanno distanza uguale a R dal punto C: cioè è l'insieme di tutti e soli i punti P del piano tali che $|\overrightarrow{PC}| = R$. Questo luogo nello spazio è una nota superficie:

DEFINIZIONE 1.16. Fissati un punto C nello spazio ed un numero reale $R > 0$, il luogo geometrico dei punti P la cui distanza dal punto C è R:

$$\Sigma = \{P|\ d(P, C) = R\ \},$$

è detto *superficie sferica* o *sfera* di centro C e raggio R (figura 1.28).

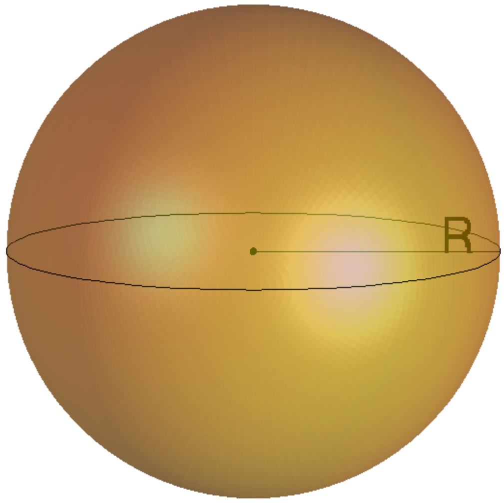

FIGURA 1.28. Rappresentazione schematica di una sfera di centro C e raggio R.

Fissiamo nello spazio un sistema di riferimento $\mathcal{R}(O, \hat{\imath}, \hat{\jmath}, \hat{k})$ cartesiano ortonormale, siano $C = \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix}$ e $P = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$, ponendo

$$d(P, C)^2 = R^2$$

otteniamo la seguente equazione

$$(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2 = R^2, \tag{1.24}$$

equazione sodisfatta da tutti e soli i punti appartenenti alla sfera Σ di centro C e raggio R. Sviluppando i quadrati otteniamo:

$$x^2 + y^2 + z^2 - 2x_0x - 2y_0y - 2z_0z + (x_0^2 + y_0^2 + z_0^2 - R^2) = 0,$$

ponendo $a = -2x_0$, $b = -2y_0$, $c = -2z_0$, infine $d = (x_0^2 + y_0^2 + z_0^2 - R^2)$, possiamo scrivere l'equazione della sfera nella forma seguente:

$$x^2 + y^2 + z^2 + ax + by + cz + d = 0, \tag{1.25}$$

detta *equazione cartesiana della sfera* Σ.

Ci chiediamo ora se una qualsiasi equazione di questo tipo rappresenta una sfera: proviamo a vedere se è possibile determinare centro e raggio a partire da a, b, c, d. Innanzitutto completiamo i quadrati:

$$(x^2 + ax + \frac{a^2}{4}) + (y^2 + by + \frac{b^2}{4}) + (z^2 + cz + \frac{c^2}{4}) = (\frac{a}{2})^2 + (\frac{b}{2})^2 + (\frac{c}{2})^2 - d,$$

equivalentemente:

$$(x + \frac{a}{2})^2 + (y + \frac{b}{2})^2 + (z + \frac{c}{2})^2 = (\frac{a}{2})^2 + (\frac{b}{2})^2 + (\frac{c}{2})^2 - d.$$

Osserviamo che se questa equazione rappresenta una sfera necessariamnete il suo centro è il punto

$$C = \begin{pmatrix} -\frac{a}{2} \\ -\frac{b}{2} \\ -\frac{c}{2} \end{pmatrix}.$$

Ora l'equazione scritta è l'equazione di una sfera di centro C se e solo se il numero reale che compare a secondo membro è strettamante positivo:

$$\frac{a^2}{4} + \frac{b^2}{4} + \frac{c^2}{4} - d > 0. \tag{1.26}$$

Se questa condizione è soddisfatta poniamo

$$R = \sqrt{\frac{a^2}{4} + \frac{b^2}{4} + \frac{c^2}{4} - d},$$

ed otteniamo l'equazione della sfera di centro C e raggio R. Se la condizione non è soddisfatta, l'equazione non ha soluzioni reali.

Esempio 1.37. Fissiamo nello spazio un sistema di riferimento cartesiano ortonormale $\mathcal{R}(O, \hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}, \hat{\boldsymbol{k}})$.

(1) L'equazione della sfera avente come centro l'origine O e raggio $R > 0$ è la seguente:
$$x^2 + y^2 + z^2 = R^2.$$
(2) L'equazione $x^2 + y^2 + z^2 - 2x + 4y + 6 = 0$ rappresenta una sfera? Completiamo i quadrati:
$$(x - 1)^2 + (y + 2)^2 + z^2 = 1 + 4 - 6 = -1,$$
osserviamo che il numero che compare a secondo membro è negativo, l'equazione non ha soluzioni reali e pertanto non rappresenta una sfera.
(3) Determiniamo centro e raggio della sfera Σ di equazione:
$$x^2 + y^2 + z^2 + 2x - y + 3z = 1.$$
Completiamo i quadrati:
$$(x + 1)^2 + (y - \frac{1}{2})^2 + (z + \frac{3}{2})^2 = 1 + 1 + \frac{1}{4} + \frac{9}{4} = \frac{9}{2},$$

da cui ricaviamo le coordinate del centro ed il raggio:

$$C = \begin{pmatrix} -1 \\ \frac{1}{2} \\ -\frac{3}{2} \end{pmatrix} \qquad R = \sqrt{\frac{9}{2}}.$$

Ci proponiamo ora di studiare le posizioni reciproche retta-sfera e piano-sfera.

Date una sfera Σ ed una retta r, per determinare la loro posizione reciproca cerchiamo gli eventuali punti di intersezione. Fissiamo un sistema di riferimento cartesiano ortonormale $\mathcal{R}(O, \hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}, \hat{\boldsymbol{k}})$, rappresentiamo la retta r con equazione parametrica

$$P = P_0 + t\boldsymbol{v}, \quad t \in \mathbb{R},$$

dove $P_0 = \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix}$ e $\boldsymbol{v} = \begin{pmatrix} l \\ m \\ n \end{pmatrix}$, sia infine

$$x^2 + y^2 + z^2 + ax + by + cz + d = 0$$

l'equazione di Σ. Sostituendo le coordinate di P nell'equazione della sfera Σ otteniamo la seguente equazione di secondo grado nell'incognita t:

$$\begin{aligned}(l^2 + m^2 + n^2)t^2 + t((a + 2x_0)l + (b + 2y_0)m + (c + 2z_0)n) \\ + (x_0^2 + y_0^2 + z_0^2 + ax_0 + by_0 + cz_0 + d) = 0.\end{aligned} \tag{1.27}$$

Dalla discussione dell'equazione trovata deduciamo che si possono verificare i seguenti casi (Figura 1.29):

(1) la retta r è *esterna*: l'equazione non ha soluzioni reali, la retta e la sfera non hanno punti comuni;
(2) la retta r è *secante*: la retta e la sfera si intersecano in due punti distinti P_1 e P_2 corrispondenti alle due soluzioni reali distinte $t_1 \neq t_2$ dell'equazione;
(3) la retta r è *tangente*: la retta e la sfera si intersecano in due punti coincidenti $P_1 = P_2$, corrispondenti alle due soluzioni reali coincidenti $t_1 = t_2$ dell'equazione.

Date una sfera Σ ed un piano π, per determinare la loro posizione reciproca confrontiamo il raggio R di Σ con la distanza del piano π dal centro C di Σ: $d(C, \pi)$. Si hanno i seguenti casi (Figura 1.30):

(1) $d(C, \pi) > R$: allora per ogni punto P del piano si ha $|\overrightarrow{PC}| \geq d(C, \pi) > R$, il piano e la sfera non hanno punti comuni, diciamo che il piano π è *esterno*;
(2) $d(C, \pi) < R$: il piano e la sfera si intersecano, la loro interezione è la circonferenza γ del piano π avente centro nel punto C', intersezione di π con la retta per C normale al piano π, e raggio $R' = \sqrt{R^2 - d(C, \pi)^2}$, diciamo che il piano è *secante*;
(3) $d(C, \pi) = R$: il piano e la sfera si intersecano in un solo punto P, il raggio per P risulta essere perpendicolare al piano π, diciamo che il piano è *tangente* alla sfera in P.

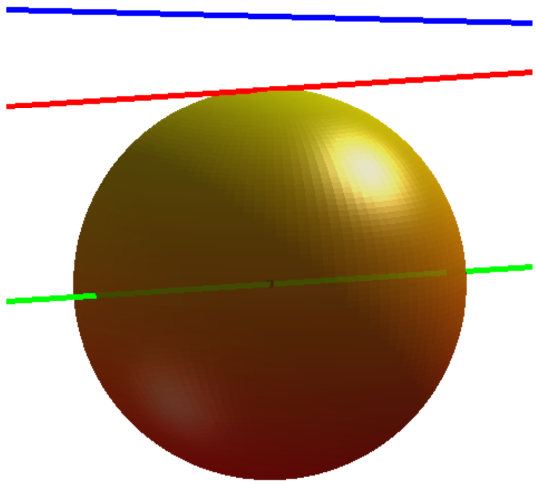

FIGURA 1.29. Rappresentazione schematica delle possibili posizioni reciproche fra una retta ed una sfera: retta secante (verde), retta tangente (rossa), retta esterna (blu).

Osserviamo che se $d(C, \pi) = R$, allora il piano π contiene il centro della sfera, quindi l'intersezione tra la sfera ed il piano π è la circonferenza γ del piano π di centro C e raggio R, cioè il piano interseca la sfera in un circonfernza di raggio massimo.
Osserviamo che in ogni punto P della sfera Σ esiste un unico piano tangente π alla sfera in P: è il piano per P perpendicolare al raggio per P. Infine se una retta r è tangente ad una sfera in un punto P, la retta è contenuta nel piano tangente alla sfera nel punto P.

ESEMPIO 1.38. Fissiamo nello spazio un sistema di riferimento $\mathcal{R}(O, \hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}, \hat{\boldsymbol{k}})$ cartesiano ortonormale.

(1) Data la sfera Σ di equazione $x^2 + y^2 + z^2 - 4x + 2z = 0$, verifichiamo che la retta r di equazioni $x - y = z + 2y = 0$ è secante e determiniamo i punti di intersezione.

Troviamo innanzittutto una rappresentazione parametrica per la retta r:

$$\begin{cases} x = t, \\ y = t, \\ z = -2t, \quad t \in \mathbb{R}. \end{cases}$$

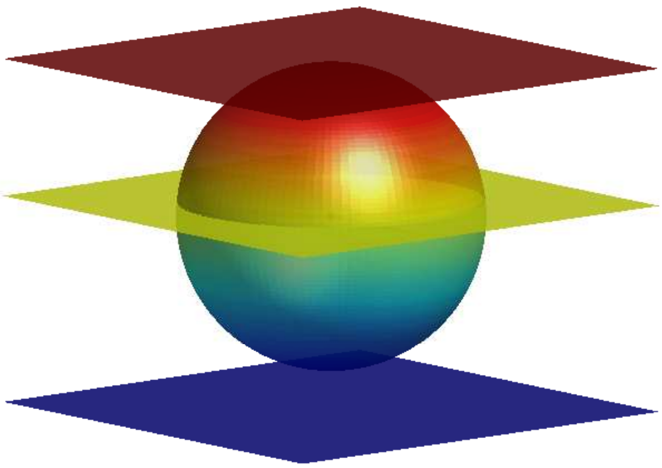

FIGURA 1.30. Rappresentazione schematica delle possibili posizioni reciproche fra un piano ed una sfera: piano secante (giallo-verde), piano tangente (rosso), piano esterno (blu).

Sostituendo le coordinate di un punto $P \in r$ nell'equazione della sfera Σ abbiamo la seguente equazione di secondo grado nell'incognita t:

$$6t^2 - 8t = 0,$$

che ammette le soluzioni reali distinte: $t_1 = 0$ e $t_2 = \frac{4}{3}$. La retta r è quindi secante, ai valori dei parametri trovati corrispondono due punti distinti della retta r:

$$P_1 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad P_2 = \begin{pmatrix} \frac{4}{3} \\ \frac{4}{3} \\ -\frac{8}{3} \end{pmatrix}.$$

(2) Data la sfera Σ di equazione $x^2+y^2+z^2+2x-2y-4z+1=0$, scriviamo l'equazione del piano tangente alla sfera nel punto $A = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$.

Per trovare il centro ed il raggio di Σ completiamo i quadrati:

$$(x+1)^2 + (y-1)^2 + (z-2)^2 = 1 + 1 + 4 - 1 = 5,$$

da cui ricaviamo

$$C = \begin{pmatrix} -1 \\ 1 \\ 2 \end{pmatrix} \quad R = \sqrt{5}.$$

Verifichiamo che il punto A appartiene alla sfera, sostituendo le coordinate di A nell'equazione di Σ:

$$2^2 + 0^2 + (-1)^2 = 5.$$

Il piano π tangente alla sfera in A è il piano per A perpendicolare al raggio per A. La direzione di tale raggio è data dal vettore:

$$\overrightarrow{AC} = \begin{pmatrix} x_C - x_A \\ y_C - y_A \\ z_C - z_A \end{pmatrix} = \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix},$$

il piano π è quindi il piano per A con direzione normale $\overrightarrow{AC}$:

$$-2(x - x_A) + 0(y - y_A) + 1(z - z_A) = 0,$$

sostituendo le coordinate di A otteniamo l'equazione del piano tangente:

$$2x - z = 1.$$

(3) Data la sfera Σ di equazione $x^2+y^2+z^2-2x+4y+2z = 0$, verifichiamo che il piano π di equazione $x - y = 0$ è secante, determiniamo centro e raggio della circonferenza γ intersezione di Σ e π.
Per trovare il centro ed il raggio di Σ completiamo i quadrati:

$$(x - 1)^2 + (y + 2)^2 + (z + 1)^2 = 1 + 4 + 1 = 6,$$

da cui ricaviamo

$$C = \begin{pmatrix} 1 \\ -2 \\ -1 \end{pmatrix} \qquad R = \sqrt{6}.$$

Calcoliamo la distanza del piano π dal centro C:

$$d(C, \pi) = \frac{|x_C - y_C|}{\sqrt{1^2 + 1^2}} = \frac{3}{\sqrt{2}}.$$

Osserviamo che risulta $d(C, \pi) < R$, per cui il piano è secante, cioè taglia sulla sfera una circonferenza γ. Il centro di γ è il punto C' che è intersezione di π con la retta r passante per C perpendicolare al piano π. La retta r ha equazioni parametriche:

$$\begin{cases} x = 1 + t, \\ y = -2 - t, \\ z = -1, \quad t \in \mathbb{R}. \end{cases}$$

Per determinare il punto C' sostituiamo le coordinate di un punto di r nell'equazione di π, otteniamo l'equazione di primo grado in t

$$2t + 3 = 0,$$

la cui soluzione è $t = -\frac{3}{2}$, che corrisponde al punto $C' = \begin{pmatrix} -\frac{1}{2} \\ -\frac{1}{2} \\ -1 \end{pmatrix}$.

Il raggio della circonferenza γ si ottiene appplicando il teorema di

Pitagora al triangolo rettangolo C, C', P, dove P è un punto di γ:

$$r = \sqrt{R^2 - d(C,\pi)^2} = \sqrt{6 - \frac{9}{2}} = \sqrt{\frac{3}{2}}.$$

CAPITOLO 2

Spazi vettoriali

Nel capitolo precedente (1) abbiamo visto come, dette x, y, z le coordinate dei punti nello spazio rispetto ad un sistema cartesiano ortogonale, l'insieme delle soluzioni di un'equazione lineare in tre variabili della forma

$$ax + by + cz = d$$

rappresenta geometricamente un piano (una volta che sia stato fissato un sistema di riferimento ortogonale). Di conseguenza lo studio di rette e piani nello spazio risulta essere equivalente allo studio di sistemi lineari in tre incognite.

Spesso nella pratica si ha a che fare con problemi in cui il numero delle incognite è maggiore di 3.

In generale le grandezze incognite $x_1, x_2, \ldots, x_n$ sono legate da una serie di relazioni esprimibili matematicamente attraverso equazioni. Trovare una soluzione del problema vuol dire attribuire un valore numerico a ciascuna incognita in modo che **tutte** le equazioni siano soddisfatte. La complessità del problema è legata a due fattori

(1) La complessità di ciascuna equazione.

Ad esempio trovare soluzioni dell'equazione

$$(e^{x_1} - x_2) + \sin(x_1^2\sqrt{x_1x_2^2 + x_1^6 + 1}) = 35 \tag{2.1}$$

è sicuramente più complicato che trovare soluzioni dell'equazione

$$3x_1 + 2x_2 = 7\,. \tag{2.2}$$

Tra le equazioni possibili, le più semplici sono quelle **lineari**, ovvero quelle della forma

$$a_1x_1 + a_2x_2 + \ldots + a_nx_n = b$$

dove $a_1, \ldots, a_n$ sono i coefficienti dell'equazione (e dunque sono numeri assegnati).

Ad esempio l'equazione (2.2) è lineare mentre l'equazione (2.1) non lo è.

(2) Il secondo fattore da cui dipende la complessità del problema è il numero di incognite. Infatti mentre il sistema lineare di due equazioni in due incognite

$$\begin{cases} 3x_1 + 2x_2 = 3 \\ 5x_1 + 7x_2 = -16 \end{cases}$$

può essere risolto in modo elementare, il sistema di 9 equazioni in 9 incognite

$$
\left\{
\begin{array}{l}
8x_1 + 32x_2 + 4x_4 + 7x_5 + 22x_6 - 12x_7 - 18x_9 = 0 \\
-12x_1 + 3x_2 + 53x_4 + 3x_5 + 71x_6 - 2x_7 - 9x_9 = -1 \\
-10x_1 + 2x_3 + 25x_4 - 7x_6 + 2x_7 - 12x_8 - 11x_9 = 3 \\
7x_1 + 9x_2 + 5x_3 + 18x_5 + 14x_6 - x_7 - 21x_9 = 27 \\
16x_1 + 17x_3 + 7x_4 + 76x_5 + 67x_7 + 11x_7 - x_9 = 12 \\
23x_1 + 15x_2 + 18x_4 + 64x_5 + 2x_6 + 15x_7 + 5x_9 = 4 \\
31x_1 + 23x_3 + 3x_5 + 4x_6 + 16x_6 - x_7 - 23x_9 = 5 \\
-58x_1 + 33x_2 + 6x_3 + 21x_4 + 41x_5 - 13x_7 - 8x_9 = 6 \\
17x_1 + 57x_3 + 18x_4 + 62x_5 + 37x_6 - 27x_7 - 4x_9 = 8
\end{array}
\right.
\tag{2.3}
$$

rappresenta da un punto di vista teorico e computazionale un problema di non immediata soluzione.

In questo corso affronteremo problemi in cui il numero delle incognite può essere arbitrariamente grande, ma le equazioni che legano le varie incognite sono le più semplici possibili, ovvero sono *lineari*. In altre parole l'argomento centrale del corso sarà lo studio dei sistemi di equazioni lineari in molte incognite.

Questo tipo di sistemi ricorre nella risoluzione di molti problemi, in campi differenti quali, ad esempio, in statica, la determinazione delle incognite vincolari di un sistema di corpi rigidi. Si consideri il seguente

Esempio 2.1. La struttura rigida piana in Figura 2.1 è composta da 3 aste rettilinee omogenee: OB, verticale, di lunghezza $\sqrt{3}\ell$ e peso vincolata a terra da una cerniera in O; AB, di lunghezza 2ℓ e peso incernierata alla prima in B, e vincolata da un carrello a terra in A, alla stessa quota di O, a distanza ℓ da questo punto; CO', orizzontale, di lunghezza $\ell/2$ e peso incernierata a terra in O' ed all'asta AB nel suo punto medio. In B è applicata una forza Trovare le incognite vincolari.

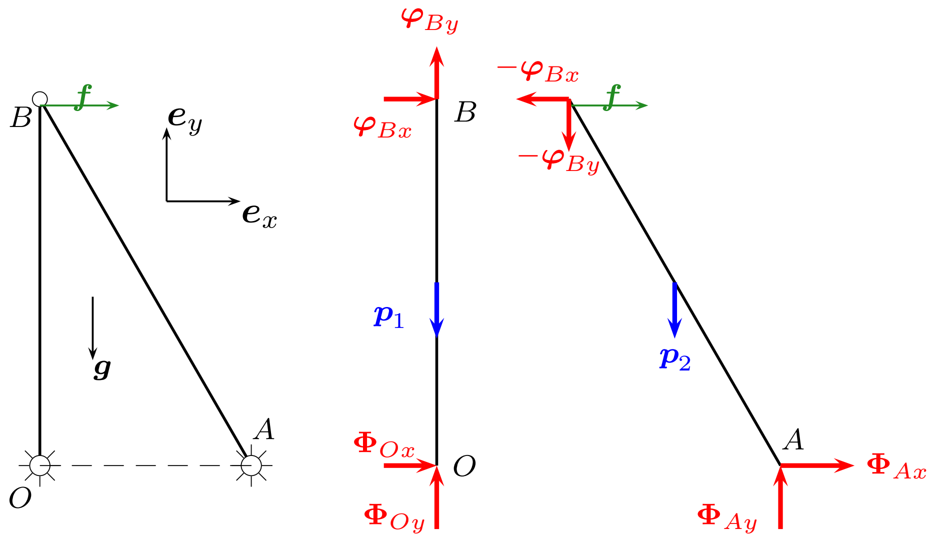

Figura 2.1. Esempio di struttura statica di cui si vogliono determinare le incognite vincolari.

Per risolvere il sistema si possono considerare separatamente le tre aste, applicando a ciscuna di esse le corrette forze interne ed esterne, ivi comprese le incognite vincolari; queste risultano (vedi Figura 2.1):

$$\begin{aligned}
\boldsymbol{\Phi}_{Ox} &= \Phi_{Ox}\boldsymbol{e}_x = x_1\boldsymbol{e}_x \\
\boldsymbol{\Phi}_{Oy} &= \Phi_{Oy}\boldsymbol{e}_y = x_2\boldsymbol{e}_y \\
\boldsymbol{\varphi}_{Bx} &= \varphi_{Bx}\boldsymbol{e}_x = x_3\boldsymbol{e}_x \\
\boldsymbol{\varphi}_{By} &= \varphi_{By}\boldsymbol{e}_y = x_4\boldsymbol{e}_y \\
\boldsymbol{\Phi}_{Ax} &= \Phi_{Ax}\boldsymbol{e}_y = x_5\boldsymbol{e}_y \\
\boldsymbol{\Phi}_{Ay} &= \Phi_{Ay}\boldsymbol{e}_y = x_6\boldsymbol{e}_y\,.
\end{aligned} \tag{2.4}$$

Le equazioni risolutive si ottengono imponendo l'equilibrio delle forze lungo le direzioni $\boldsymbol{e}_x$ ed $\boldsymbol{e}_y$ per ogni asta (prima equazione cardinale della statica), e dei momenti calcolati rispetto ai punti O per la prima e rispetto a B per la seconda asta (seconda equazione cardinale della statica).

$$\begin{cases}
\Phi_{Ox} + \varphi_{Bx} = 0 \\
\Phi_{Oy} + \varphi_{By} - p_1 = 0 \\
\varphi_{By}\sqrt{3}\,\ell = 0 \\
\Phi_{Ax} - \varphi_{Bx} + f = 0 \\
\Phi_{Ay} - \varphi_{By} - p_2 = 0 \\
\Phi_{Ay}\ell + \Phi_{Ax}\sqrt{3}\,\ell - p_2\,\frac{\ell}{2} = 0
\end{cases} \tag{2.5}$$

Usando le incognite $x_1 \dots x_6$ introdotte nelle (2.4) e semplificando, il sistema risolutivo (2.5) si riduce a:

$$\begin{cases}
x_1 + x_3 = 0 \\
x_2 + x_4 = p_1 \\
x_4 = 0 \\
x_5 - x_3 = -f \\
x_6 - x_4 = p_2 \\
x_6 + \sqrt{3}x_5 = \frac{p_2}{2}
\end{cases} \tag{2.6}$$

che è un sistema del tipo (2.3) discusso sopra. Sembra quindi utile possedere un metodo per risolvere tali sistemi.

Volendo trattare il problema per un qualsiasi numero di incognite sarà necessario sviluppare una teoria che sia il più possibile slegata da tale numero (e dunque applicabile a qualsiasi sistema lineare). Questa è la finalità che ci porta a considerare gli spazi vettoriali astratti (che verranno introdotti in questo capitolo). In maniera approsimativa potremmo affermare che ogni soluzione di un sistema di equazioni lineari con n incognite rappresenta un punto in un certo spazio vettoriale (ad esempio nel caso di 3 incognite, la soluzione come abbiamo visto rappresenta un punto dello spazio $\mathbb{E}^3_O$). Sviluppando una teoria che valga per tutti gli spazi vettoriali saremo alla fine in grado di trattare sistemi con un numero arbitrario di incognite.

1. La nozione di spazio vettoriale

1.1. Gli spazi $\mathbb{R}^n$.

Fissato un numero intero positivo n, una n-upla ordinata è un insieme ordinato di n-numeri.

Per denotare una n-upla ordinata utilizzeremo la seguente notazione

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

dove x_i sono i numeri che compongono e vengono chiamati le **componenti** della n-upla. Ad esempio $\boldsymbol{v} = \begin{pmatrix} 23 \\ 47 \\ 13 \end{pmatrix}$ è una terna ordinata, mentre $\boldsymbol{w} = \begin{pmatrix} 2 \\ 16 \\ 17 \\ 18 \end{pmatrix}$ è una quaterna ordinata. La seconda componente di $\boldsymbol{v}$ è $v_2 = 47$ mentre la terza componente di $\boldsymbol{w}$ è $w_3 = 17$.

Osservazione 2.2. Considerare n-uple ordinate significa che due n-uple sono considerate uguali se e soltanto se la prima componente della prima coincide con la prima componente della seconda, la seconda componente della prima coincide con la seconda componente della seconda e così via.

Ad esempio le terne

$$\begin{pmatrix} 23 \\ 47 \\ 13 \end{pmatrix} \qquad \begin{pmatrix} 13 \\ 47 \\ 23 \end{pmatrix}$$

verranno considerate diverse.

Definizione 2.1. Fissato un numero intero positivo n, **lo spazio $\mathbb{R}^n$ è l'insieme di tutte le n-uple ordinate di numeri reali**.

Ad esempio $\begin{pmatrix} 3 \\ 13 \\ -65 \\ 98 \end{pmatrix}$ appartiene a $\mathbb{R}^4$ mentre $\begin{pmatrix} -117 \\ 23 \\ 86 \end{pmatrix}$ appartiene a $\mathbb{R}^3$ e

$$\begin{pmatrix} 12 \\ -3 \\ 98 \\ 76 \\ 0 \\ 0 \\ 1 \\ -35 \end{pmatrix}$$

appartiene a $\mathbb{R}^8$.

Abbiamo visto che fissata una base ortonormale $\mathcal{B} = \left\{ \hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}, \hat{\boldsymbol{k}} \right\}$ dello spazio $\mathbb{E}^3_O$ è possibile associare ad ogni vettore una terna di numeri che rappresentano le coordinate del vettore rispetto alla base. Così ad esempio alla terna $\begin{pmatrix} 11 \\ -3 \\ 7 \end{pmatrix}$ corrisponde il vettore $\boldsymbol{v} = 11\hat{\boldsymbol{\imath}} - 3\hat{\boldsymbol{\jmath}} + 7\hat{\boldsymbol{k}}$ e così via.

Abbiamo già osservato nello scorso capitolo (1.5) che se le coordinate del vettore $\boldsymbol{v}$ sono $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ e le coordinate del vettore $\boldsymbol{w}$ sono $\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$ allora le coordinate del vettore $\boldsymbol{v} + \boldsymbol{w}$ sono

$$\begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ x_3 + y_3 \end{pmatrix}$$

mentre le coordinate del vettore $\lambda\boldsymbol{v}$ sono

$$\begin{pmatrix} \lambda x_1 \\ \lambda x_2 \\ \lambda x_3 \end{pmatrix}$$

Tutto ciò suggerisce che si possa introdurre un'operazione di **somma** e di **moltiplicazione per scalare** sull'insieme $\mathbb{R}^n$.

Definizione 2.2. In particolare la somma sarà definita nel seguente modo. Dato $\boldsymbol{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$ e $\boldsymbol{w} = \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix}$ il vettore $\boldsymbol{v} + \boldsymbol{w}$ è una n-upla ordinata la cui prima componente è la somma delle prime componenti di $\boldsymbol{v}$ e $\boldsymbol{w}$, la cui seconda componente è la somma delle seconde componenti di $\boldsymbol{v}$ e $\boldsymbol{w}$ e così via. In simboli

$$\boldsymbol{v} + \boldsymbol{w} = \begin{pmatrix} v_1 + w_1 \\ v_2 + w_2 \\ \vdots \\ v_n + w_n \end{pmatrix}$$

Analogamente se λ è uno scalare il vettore $\lambda\boldsymbol{v}$ si otterrà moltiplicando tutte le componenti di v per il fattore λ. In simboli

$$\lambda\boldsymbol{v} = \begin{pmatrix} \lambda v_1 \\ \lambda v_2 \\ \vdots \\ \lambda v_n \end{pmatrix}$$

Per analogia con il caso dei vettori applicati, gli elementi di $\mathbb{R}^n$ sono detti **vettori numerici**. In altre parole, un vettore numerico è una n-upla ordinata di numeri.

Esempio 2.3. Vediamo quale esempio di queste operazioni con vettori numerici:

(1) $$\begin{pmatrix} 4 \\ -3 \\ 9 \\ 10 \end{pmatrix} + \begin{pmatrix} 9 \\ 0 \\ 7 \\ 2 \end{pmatrix} = \begin{pmatrix} 13 \\ -3 \\ 16 \\ 12 \end{pmatrix}$$

(2) $$3 \begin{pmatrix} 2 \\ 3 \\ 9 \end{pmatrix} = \begin{pmatrix} 6 \\ 9 \\ 27 \end{pmatrix}$$

(3) $$\frac{1}{3} \begin{pmatrix} 6 \\ 2 \\ 12 \end{pmatrix} = \begin{pmatrix} 2 \\ \frac{2}{3} \\ 4 \end{pmatrix}$$

(4) $$5 \begin{pmatrix} 2 \\ -1 \\ 6 \\ -3 \end{pmatrix} + \frac{2}{3} \begin{pmatrix} -1 \\ 0 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 10 - \frac{2}{3} \\ -5 \\ 30 + \frac{4}{3} \\ -15 + 2 \end{pmatrix} = \begin{pmatrix} \frac{28}{3} \\ -5 \\ \frac{94}{3} \\ -13 \end{pmatrix}$$

OSSERVAZIONE 2.4. Non si può sommare un elemento di $\mathbb{R}^3$ con un elemento di $\mathbb{R}^4$ o $\mathbb{R}^5$. In generale si possono sommare solo n-uple che contengono lo stesso numero di componenti.

Osserviamo che la somma di vettori numerici gode dell e seguenti proprietà:

(S_1) **Proprietà commutativa** Infatti si ha

$$\boldsymbol{v} + \boldsymbol{w} = \begin{pmatrix} v_1 + w_1 \\ \vdots \\ v_n + w_n \end{pmatrix} \qquad \boldsymbol{w} + \boldsymbol{v} = \begin{pmatrix} w_1 + v_1 \\ \vdots \\ w_n + v_n \end{pmatrix}$$

e poiché vale che $v_1 + w_1 = w_1 + v_1$, $v_2 + w_2 = w_2 + v_2, \ldots$, si ricava che tutte le componenti di $\boldsymbol{v} + \boldsymbol{w}$ sono uguali alle corrispettive componenti di $\boldsymbol{w} + \boldsymbol{v}$ e dunque $\boldsymbol{v} + \boldsymbol{w} = \boldsymbol{w} + \boldsymbol{v}$.

(S_2) Analogamente si può mostrare che la somma gode della proprietà **associativa**: $(\boldsymbol{v} + \boldsymbol{w}) + \boldsymbol{u} = \boldsymbol{v} + (\boldsymbol{w} + \boldsymbol{u})$.

(S_3) Il **vettore nullo** di $\mathbb{R}^n$ è la n-upla le cui entrate sono tutte nulle

$$\mathbf{0}_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Si noti che $\mathbf{0}_n$ è l'elemento neutro della somma ovvero per ogni $\boldsymbol{v} \in \mathbb{R}^n$ si ha

$$\boldsymbol{v} + \mathbf{0}_n = \boldsymbol{v}\,.$$

OSSERVAZIONE 2.5. L'indice n nella notazione $\mathbf{0}_n$ sta a significare il numero di componenti da cui è formato il vettore. Così ad esempio si ha

$$\mathbf{0}_2 = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \qquad \mathbf{0}_3 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \qquad \mathbf{0}_4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \ldots$$

Si noti che $\mathbf{0}_2 \in \mathbb{R}^2$, $\mathbf{0}_3 \in \mathbb{R}^3$ e in generale $\mathbf{0}_n \in \mathbb{R}^n$.

Nel seguito laddove non ci sia ambiguità ometteremo l'indice n indicando semplicemente con $\mathbf{0}$ il vettore nullo di $\mathbb{R}^n$.

(S_4) Infine che per ogni vettore $\boldsymbol{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \in \mathbb{R}^n$ **l'opposto di $\boldsymbol{v}$** è il vettore $(-1) \cdot \boldsymbol{v} = \begin{pmatrix} -v_1 \\ \vdots \\ -v_n \end{pmatrix}$, che denoteremo semplicemente con $-\boldsymbol{v}$.

Chiaramente si ha

$$\boldsymbol{v} + (-\boldsymbol{v}) = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} + \begin{pmatrix} -v_1 \\ \vdots \\ -v_n \end{pmatrix} = \begin{pmatrix} v_1 - v_1 \\ \vdots \\ v_n - v_n \end{pmatrix} = \mathbf{0}_n$$

ovvero $-\boldsymbol{v}$ è l'inverso di $\boldsymbol{v}$ per la somma.

Enunciamo ora le principali proprietà algebriche della moltiplicazione per scalare,

(P_1) $\alpha(\beta\boldsymbol{v}) = (\alpha\beta)\boldsymbol{v}, \quad \forall \alpha, \beta \in \mathbb{R}, \forall \boldsymbol{v} \in \mathbb{R}^n$;

(P_2) $\alpha(\boldsymbol{v} + \boldsymbol{w}) = \alpha\boldsymbol{v} + \alpha\boldsymbol{w}, \quad \forall \alpha \in \mathbb{R}, \forall \boldsymbol{v}, \boldsymbol{w} \in \mathbb{R}^n$
(proprietà distributiva rispetto alla somma di vettori).

(P_3) $(\alpha + \beta)\boldsymbol{v} = \alpha\boldsymbol{v} + \beta\boldsymbol{v}, \quad \forall \alpha, \beta \in \mathbb{R}, \forall \boldsymbol{v} \in \mathbb{R}^n$
(proprietà distributiva rispetto alla somma di scalari).

(P_4) $1 \cdot \boldsymbol{v} = \boldsymbol{v}, \quad \forall \boldsymbol{v} \in \mathbb{R}^n$.

Verifichiamo la seconda di queste proprietà lasciando allo studente il compito di verificare le altre.

Osserviamo che

$$\alpha(\boldsymbol{v} + \boldsymbol{w}) = \begin{pmatrix} \alpha(v_1 + w_1) \\ \vdots \\ \alpha(v_n + w_n) \end{pmatrix}$$

mentre

$$\alpha\boldsymbol{v} + \alpha\boldsymbol{w} = \begin{pmatrix} \alpha v_1 \\ \vdots \\ \alpha v_n \end{pmatrix} + \begin{pmatrix} \alpha w_1 \\ \vdots \\ \alpha w_n \end{pmatrix} = \begin{pmatrix} \alpha v_1 + \alpha w_1 \\ \vdots \\ \alpha v_n + \alpha w_n \end{pmatrix}$$

e confrontando queste espressioni, risulta che ciascuna componente di $\alpha(\boldsymbol{v}+\boldsymbol{w})$ è uguale alla corrispondente componente di $\alpha\boldsymbol{v}+\alpha\boldsymbol{w}$, da cui segue l'uguaglianza cercata.

2. Spazi vettoriali astratti

Le considerazioni appena svolte ci portano a identificare in modo astratto e generale le strutture che si comportano come abbiamo appena visto per lo spazio $\mathbb{R}^n$ con le operazioni di addizione e moltiplicazione per uno scalare definite in (2.2).

DEFINIZIONE 2.3. Uno **spazio vettoriale reale** è un insieme V su cui siano definite due operazioni

(1) la **somma** che ad ogni coppia di elementi $\boldsymbol{v}, \boldsymbol{w}$ di V associa un terzo elemento denotato con $\boldsymbol{v} + \boldsymbol{w}$.
(2) la **moltiplicazione per scalare** che ad ogni elemento $\boldsymbol{v} \in V$ e ad ogni scalare $\lambda \in \mathbb{R}$ associa un elemento in V denotato con $\lambda\boldsymbol{v}$.

in modo tale che le seguenti proprietà siano soddisfatte:

(S_1) $\boldsymbol{v} + \boldsymbol{w} = \boldsymbol{w} + \boldsymbol{v}$ [proprietà commutativa];
(S_2) $(\boldsymbol{v} + \boldsymbol{w}) + \boldsymbol{u} = \boldsymbol{v} + (\boldsymbol{w} + \boldsymbol{u})$ [proprietà associativa della somma];
(S_3) Esiste un elemento $\mathbf{0}_V$ tale che $\boldsymbol{v} + \mathbf{0}_V = \boldsymbol{v}$ per ogni $\boldsymbol{v} \in V$ [esistenza dell'elemento neutro];
(S_4) Dato $\boldsymbol{v} \in V$ esiste un elemento in V denotato con $-\boldsymbol{v}$ tale che $\boldsymbol{v} + (-\boldsymbol{v}) = \mathbf{0}_V$ [esistenza dell'opposto];
(P_1) $\alpha(\beta\boldsymbol{v}) = (\alpha\beta)\boldsymbol{v}$;
(P_2) $\alpha(\boldsymbol{v} + \boldsymbol{w}) = \alpha\boldsymbol{v} + \alpha\boldsymbol{w}$ [proprietà distributiva della moltiplicazione rispetto alla somma di vettori];
(P_3) $(\alpha+\beta)\boldsymbol{v} = \alpha\boldsymbol{v}+\beta\boldsymbol{v}$ [proprietà distributiva della moltiplicazione rispetto alla somma di scalari];
(P_4) $1 \cdot \boldsymbol{v} = \boldsymbol{v}$ per ogni $\boldsymbol{v} \in V$.

Gli elementi di uno spazio vettoriale V sono detti **vettori**. Normalmente indicheremo i vettori di uno spazio vettoriale in grassetto per distinguerli dalle grandezze scalari.

L'elemento $\mathbf{0}_V$ è detto **vettore nullo**. A volte, laddove non ci sia ambiguità, denoteremo il vettore nullo semplicemente con il simbolo $\mathbf{0}$.

OSSERVAZIONE 2.6. Le proprietà $(S1) - (S4)$ si riassumono dicendo che V è un gruppo rispetto all'operazione di addizione fra vettori (Sezione 3 del Capitolo introduttivo 0).

In base a questa definizione, ci rendiamo conto che finora abbiamo incontrato vari esempi di spazi vettoriali:

(1) lo spazio $\mathbb{E}^3_O$ dei vettori geometrici applicati nel punto O;
(2) gli spazi $\mathbb{R}^n$ al variare di n.

In effetti esistono molti altri esempi di spazi vettoriali. Riportiamo un esempio.

ESEMPIO 2.7. Lo spazio dei polinomi $\mathbb{R}[x]$. Ricordiamo che un polinomio nella variabile x è una funzione della variabile x della forma

$$p(x) = a_0 + a_1x + a_2x^2 + \ldots + a_nx^n\,.$$

L'insieme di tutti i polinomi nella variabile x è denotato con $\mathbb{R}[x]$. Dati due polinomi p, q è chiaramente definita una somma di questi polinomi $p + q$ semplicemente ponendo

$$(p + q)(x) = p(x) + q(x)$$

e risulta che $p+q$ è ancora un polinomio. Ad esempio se $p(x) = 3+2x^2-5x^3$ e $q(x) = 2-x+x^2$ allora

$$\begin{aligned}(p+q)(x) = p(x)+q(x) = &3+2x^2-5x^3 \ + \ 2-x+x^2 = \\ &= 5-x+3x^2-5x^3\,.\end{aligned}$$

Analogamente dato un polinomio p e un numero λ è definito il prodotto λp ponendo

$$(\lambda p)(x) = \lambda p(x)$$

Ad esempio se $p(x) = 7-13x+8x^2+12x^4$ allora il polinomio $3p$ è semplicemente

$$3p(x) = 21-39x+24x^2+36x^4\,.$$

È facile verificare che le operazioni così definite soddisfano le proprietà (S_1), (S_2), (S_3), (S_4), (P_1), (P_2), (P_3), (P_4).

In particolare notiamo che l'elemento neutro è il polinomio nullo (ovvero il polinomio costantemente uguale a 0), ed il polinomio opposto di $p(x)$ per l'operazione di somma tra polinomi è il polinomio $-p(x)$ (ossia, un polinomio con tutti i coefficienti ordinatamente opposti a quelli di $p(x)$). Dunque $\mathbb{R}[x]$ è uno spazio vettoriale.

Vogliamo ora elencare alcune proprietà che sono conseguenza delle proprietà (S_1), (S_2), (S_3), (S_4), (P_1), (P_2), (P_3), (P_4) e che quindi sono vere per tutti gli spazi vettoriali.

Anche se non riportiamo la dimostrazione completa, invitiamo lo studente a verificarle nel caso in cui lo spazio vettoriale sia $\mathbb{R}^n$.

Proposizione 2.1. *Sia V uno spazio vettoriale allora*

(1) $\alpha \mathbf{0}_V = \mathbf{0}_V$.
(2) $0 \cdot \boldsymbol{v} = \mathbf{0}_V$.
(3) $(-1) \cdot \boldsymbol{v} = -\boldsymbol{v}$.
(4) *se* $\boldsymbol{v}+\boldsymbol{u} = \boldsymbol{v}+\boldsymbol{u}'$ *allora* $\boldsymbol{u} = \boldsymbol{u}'$. *[regola di cancellazione]*
(5) *se* $\alpha\boldsymbol{v} = \mathbf{0}_V$ *allora o* $\alpha = 0$ *oppure* $\boldsymbol{v} = \mathbf{0}_V$.

Dimostrazione. Convinciamoci, per esempio, della (2); per ogni $\boldsymbol{v} \in V$, si ha

$$\boldsymbol{v}+0\cdot\boldsymbol{v} = 1\boldsymbol{v}+0\cdot\boldsymbol{v} = (1+0)\cdot\boldsymbol{v} = 1\boldsymbol{v} = \boldsymbol{v}\,,$$

quindi $0\cdot\boldsymbol{v} = \mathbf{0}_V$. □

Osservazione 2.8. La proprietà (5) implica la seguente proprietà, spesso utile nella pratica: se $\boldsymbol{v} \in V$ è un vettore **non nullo** e $\alpha, \beta \in \mathbb{R}$ sono due numeri tali che

$$\alpha\boldsymbol{v} = \beta\boldsymbol{v}$$

allora necessariamente si deve avere $\alpha = \beta$.

Infatti se $\alpha\boldsymbol{v} = \beta\boldsymbol{v}$, si ha che $\alpha\boldsymbol{v}-\beta\boldsymbol{v} = \mathbf{0}_V$ e dunque $(\alpha-\beta)\boldsymbol{v} = \mathbf{0}_V$ e per la proprietà (5) deduciamo che $\alpha-\beta = 0$ ovvero $\alpha = \beta$.

3. Sottospazi vettoriali

3.1. Piani e rette in $\mathbb{E}^3_O$.

Sia r una retta **passante per l'origine** nello spazio. La retta r è l'insieme di tutti i vettori applicati in O con punto finale su r, quindi paralleli ad r; come visto nel Capitolo 1 (Proposizione 1.7), questa è indentificata con l'insieme $\mathrm{Span}(\boldsymbol{u})$, dove $\boldsymbol{u}$ è un qualsiasi vettore che ha la direzione di r.

Osserviamo che se $\boldsymbol{v}$ e $\boldsymbol{w}$ sono vettori applicati in O e paralleli a r, allora lo è anche $\boldsymbol{v}+\boldsymbol{w}$. Dunque la retta r gode della seguente proprietà:

(SS1) Comunque si fissino $\boldsymbol{v}, \boldsymbol{w} \in r$ si ha che $\boldsymbol{v}+\boldsymbol{w} \in r$.

Inoltre osserviamo che se $\boldsymbol{v}$ è un vettore applicato in O e parallelo ad r, allora lo è anche un qualsiasi suo multiplo:

(SS2) Comunque sia fissato $\boldsymbol{v} \in r$ si ha che $\lambda \boldsymbol{v} \in r$ per ogni $\lambda \in \mathbb{R}$.

Osservazione 2.9. Il fatto che la retta r passi per l'origine è un'ipotesi essenziale affinché tali proprietà siano verificate.

In effetti, consideriamo $r' \parallel r$; se r' non passa per l'origine, essa corrisponde al traslato di r (ossia di $\mathrm{Span}(\boldsymbol{u})$), mediante un vettore $\overrightarrow{OP_0}$ non parallelo ad r', ossia l'insieme dei vettori applicati in O con punto finale su r'.

In questo caso i vettori che congiungono O a r *non* sono paralleli tra loro. Inoltre, dato un vettore $\overrightarrow{OA} \in r'$ è facile verificare che il vettore $2\overrightarrow{OA}$, per esempio, non appartiene a r' (Figura 2.2).

Consideriamo ora un piano π passante per O. Anche in questo caso possiamo identificare π con l'insieme dei vettori applicati in O con punto finale su π; come visto nel Capitolo 1 (Proposizione 1.7), π si identifica con $\mathrm{Span}(\boldsymbol{u}, \boldsymbol{v})$, con $\boldsymbol{u}, \boldsymbol{v}$ linearemente indipendenti e appartenenti a π.

Utilizzando questa identificazione, i vettori in π sono esattamente tutti i vettori applicati in O paralleli a π, ossia nel piano stesso. Ora, dati due vettori $\overrightarrow{OA}, \overrightarrow{OB}$ paralleli al piano π, si ha che i punti O, A, B appartengono a π e dunque il parallelogramma $\mathcal{P}$ individuato dai vettori $\overrightarrow{OA}$, $\overrightarrow{OB}$ è contenuto in π. Segue che anche il quarto vertice Q di $\mathcal{P}$ appartiene al piano π. Ciò mostra che $\overrightarrow{OA}+\overrightarrow{OB}$ appartiene al piano. Dunque π soddisfa la proprietà (SS1).

Analogamente se $\overrightarrow{OA}$ è contenuto sul piano, allora lo saranno tutti i suoi multipli. Dunque π soddisfa anche la proprietà (SS2).

Osservazione 2.10. Ancora una volta il fatto che l'origine sia contenuta in π è un'ipotesi essenziale!

Osserviamo che un qualsiasi sottoinsieme S dello spazio, per esempio una sfera di centro O e raggio 1, può essere identificato all'insieme dei vettori applicati in O con punto finale su S. Ci si può chiedere per quali sottoinsiemi le proprietà (SS1) e (SS2) sono verificate.

Per rette e piani abbiamo già risposto: tali proprietà sono verificate se e solo se contengono l'origine.

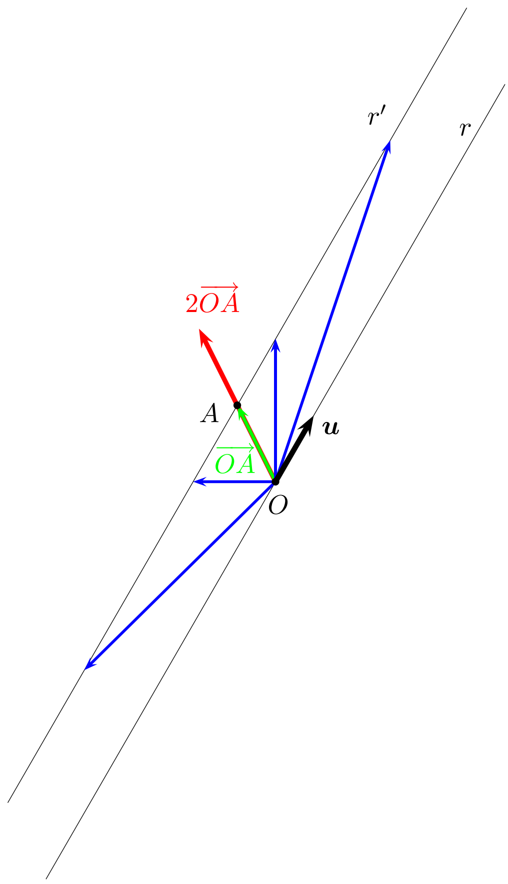

FIGURA 2.2. Se la retta r' non passa per l'origine, i vettori che individuano i suoi punti *non* sono paralleli fra loro, quindi, ad esempio, se $\overrightarrow{OA} \in r'$, si ha che $2\overrightarrow{OA} \notin r'$

Ci sono altri due casi in cui è banale verificare che tali proprietà siano soddisfatte. Il caso $S = \mathbb{E}^3_O$ (ovvero tutto lo spazio) e il caso in cui S contenga solo il vettore nullo.

In effetti più avanti dimostreremo che questi sono gli unici casi possibili:

PROPOSIZIONE 2.2. *I sottoinsiemi di $\mathbb{E}^3_O$ che soddisfano (SS1) e (SS2) sono soltanto le rette e i piani passanti per l'origine, $\mathbb{E}^3_O$ e $\left\{\overrightarrow{OO}\right\}$.*

3.2. Nozione di sottospazio.

Dalla Proposizione 2.2 segue che rette e piani sono caratterizzati dal fatto di soddisfare le proprietà (SS1) e (SS2). Osserviamo che quest'ultime sono essenzialmente proprietà di tipo algebrico, nel senso che dipendono solo dalle operazioni di somma e moltiplicazione per scalare. In particolare esse hanno senso in qualsiasi spazio vettoriale.

Ciò rende possibile allargare la nozione di retta e piano passante per l'origine ad un qualsiasi spazio vettoriale.

Più precisamente, consideriamo uno spazio vettoriale reale V.

DEFINIZIONE 2.4. Un sottoinsieme W di V è un **sottospazio vettoriale** di V se W è non vuoto e verifica (SS1) e (SS2) ovvero gode delle seguenti proprietà:

(1) per ogni $\boldsymbol{v}, \boldsymbol{w} \in W$ si ha che $\boldsymbol{v} + \boldsymbol{w} \in W$;
(2) per ogni $\boldsymbol{v} \in W$ e per ogni $\lambda \in \mathbb{R}$ si ha che $\lambda \boldsymbol{v} \in W$.

ossia, come si usa dire sinteticamente, se W è chiuso rispetto all'operazione interna di somma e rispetto all'operazione esterna di moltiplicazione per uno scalare.

OSSERVAZIONE 2.11. Risultano immediatamente le osservazioni riportate nei seguenti.

(1) Con questo linguaggio possiamo riformulare la Proposizione 2.2 in questo modo:
    I sottospazi vettoriali di $\mathbb{E}^3_O$ sono
    - Le rette per l'origine.
    - I piani per l'origine.
    - $\mathbb{E}^3_O$.
    - $\left\{\overrightarrow{OO}\right\}$
(2) Tutti i sottospazi di $\mathbb{E}^3_O$ contengono l'origine. Questo è un fatto generale: se W è un sottospazio di V allora contiene $\mathbf{0}_V$.
    Infatti fissato un qualsiasi elemento $\boldsymbol{v} \in W$ si ha che $0 \cdot \boldsymbol{v} \in W$. Del resto $0 \cdot \boldsymbol{v} = \mathbf{0}_V$ e dunque abbiamo che $\mathbf{0}_V \in W$.
(3) Ogni spazio vettoriale V contiene sempre almeno due sottospazi che sono
    - V stesso;
    -$\{\mathbf{0}_V\}$.
    Tali sottospazi sono detti **banali**.
(4) Se W è un sottospazio di V allora abbiamo che dati $\boldsymbol{v}, \boldsymbol{w} \in W$ e $\lambda \in \mathbb{R}$

$$\begin{aligned} &\boldsymbol{v} + \boldsymbol{w} \in W \\ &\lambda \boldsymbol{v} \in W \end{aligned}$$

Ciò implica che le operazioni di somma e moltiplicazione per scalare definite su V si restringono ad operazioni su W. Rispetto a tali operazioni W è esso stesso uno spazio vettoriale. Chiaramente in tal caso $\mathbf{0}_W$ coincide con $\mathbf{0}_V$.

3.3. Esempi.

(1) In $\mathbb{R}^3$ consideriamo l'insieme $W = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3 \mid x_1 = 0 \right\}$.
    Verifichiamo che W è un sottospazio.

I vettori di W sono vettori la cui prima componente è 0. Ad esempio $\begin{pmatrix}0\\1\\2\end{pmatrix}$ appartiene a W, mentre $\begin{pmatrix}1\\2\\0\end{pmatrix}$ non vi appartiene.

Per prima cosa verifico che $\mathbf{0} \in W$: questo è chiaro infatti la prima componente del vettore nullo è 0. Così facendo, non solo posso affermare che l'insieme non è vuoto, ma anche che la condizione necessaria (punto 2 dell'osservazione 2.11)

Poi devo verificare che presi due vettori $\boldsymbol{v} = \begin{pmatrix}v_1\\v_2\\v_3\end{pmatrix}$ e $\boldsymbol{w} = \begin{pmatrix}w_1\\w_2\\w_3\end{pmatrix}$ appartenenti a W, la somma $\boldsymbol{v}+\boldsymbol{w} \in W$. Ovvero devo verificare che la prima componente di $\boldsymbol{v}+\boldsymbol{w}$ sia zero.

A titolo di esempio, proviamo a fare la verifica per due specifici vettori in W

$$\boldsymbol{v} = \begin{pmatrix}0\\1\\-1\end{pmatrix},\ \boldsymbol{w} = \begin{pmatrix}0\\10\\-3\end{pmatrix}.$$

In questo caso

$$\boldsymbol{v}+\boldsymbol{w} = \begin{pmatrix}0\\11\\-4\end{pmatrix}$$

che è ancora un vettore in W.

Proviamo ora a fare la verifica per vettori generici $\boldsymbol{v}, \boldsymbol{w}$ di W.

La prima componente di $\boldsymbol{v}+\boldsymbol{w}$ è uguale a v_1+w_1. Del resto poiché per ipotesi $\boldsymbol{v} \in W$ si ha che $v_1 = 0$ e poiché $\boldsymbol{w} \in W$ abbiamo anche $w_1 = 0$. Dunque $v_1 + w_1 = 0 + 0 = 0$.

Infine devo verificare che fissato $\boldsymbol{v} = \begin{pmatrix}v_1\\v_2\\v_3\end{pmatrix} \in W$ e $\lambda \in \mathbb{R}$ si ha che $\lambda\boldsymbol{v} \in W$, ovvero la prima componente di λv deve essere uguale a 0.

Del resto si ha che $\lambda\boldsymbol{v} = \begin{pmatrix}\lambda v_1\\\lambda v_2\\\lambda v_3\end{pmatrix}$ e dunque la prima componente di $\lambda\boldsymbol{v} = \lambda v_1$. Poiché $\boldsymbol{v} \in W$ si ha che $v_1 = 0$ e dunque $\lambda v_1 = 0$.

(2) In $\mathbb{R}^3$ consideriamo il sottoinsieme $W = \left\{ \begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix} \mid x_1 = 1 \right\}$. W non è un sottospazio vettoriale. Infatti $\mathbf{0} \notin W$.

(3) In $\mathbb{R}^4$ consideriamo il sottoinsieme

$$W = \left\{ \begin{pmatrix}x_1\\x_2\\x_3\\x_4\end{pmatrix} \mid x_1 + x_2 + x_3 + x_4 = 0 \right\}.$$

Un vettore appartiene a W se la somma delle sue componenti si annulla. Per esempio $\begin{pmatrix} 2 \\ -1 \\ -1 \\ 0 \end{pmatrix}$ appartiene a W mentre $\begin{pmatrix} 2 \\ 1 \\ 1 \\ 1 \end{pmatrix}$ non appartiene a W.

Verifichiamo che W è un sottospazio vettoriale di $\mathbb{R}^4$.

$\mathbf{0} \in W$: infatti la somma delle componenti di $\mathbf{0}$ si annulla.

Dati $\boldsymbol{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{pmatrix}$ e $\boldsymbol{w} = \begin{pmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \end{pmatrix}$ appartenenti a W devo verificare che $\boldsymbol{v} + \boldsymbol{w} \in W$.

Ora $\boldsymbol{v} + \boldsymbol{w} = \begin{pmatrix} v_1 + w_1 \\ v_2 + w_2 \\ v_3 + w_3 \\ v_4 + w_4 \end{pmatrix}$ e dunque devo calcolare

$$\begin{aligned} &(v_1 + w_1) + (v_2 + w_2) + (v_3 + w_3) + (v_4 + w_4) \\ &= (v_1 + v_2 + v_3 + v_4) + (w_1 + w_2 + w_3 + w_4) \end{aligned}$$

Per ipotesi $\boldsymbol{v} \in W$ e dunque $v_1 + v_2 + v_3 + v_4 = 0$.

Analogamente poiché $\boldsymbol{w} \in W$ si ha $w_1 + w_2 + w_3 + w_4 = 0$ e dunque

$$\begin{aligned} &(v_1 + w_1) + (v_2 + w_2) + (v_3 + w_3) + (v_4 + w_4) \\ &= (v_1 + v_2 + v_3 + v_4) + (w_1 + w_2 + w_3 + w_4) = 0 \end{aligned}$$

Analogamente si verifica che $\boldsymbol{v} \in W$ allora $\lambda \boldsymbol{v} \in W$ per ogni $\lambda \in \mathbb{R}$.

(4) In $\mathbb{R}^3$ consideriamo l'insieme $W = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mid |x_1| = |x_2| \right\}$.

In questo caso un vettore appartiene a W se il valore assoluto della prima componente è uguale al valore assoluto della seconda. Vettori di W sono $\boldsymbol{v} = \begin{pmatrix} 4 \\ 4 \\ 100 \end{pmatrix}$ e $\boldsymbol{w} = \begin{pmatrix} 4 \\ -4 \\ 20 \end{pmatrix}$.

Si noti però che $\boldsymbol{v} + \boldsymbol{w} = \begin{pmatrix} 8 \\ 0 \\ 120 \end{pmatrix} \notin W$. Dunque W non è un sottospazio di $\mathbb{R}^3$.

Osservazione 2.12. Si noti che nell'ultimo esempio il vettore $\mathbf{0}$ appartiene a W, nonostante ciò W non è un sottospazio. Dunque per verificare che un dato un sottoinsieme U di uno spazio vettoriale V sia un sottospazio, **NON** è sufficiente verificare che $\mathbf{0}$ ci appartenga. Ciò significa che se $\mathbf{0}$ non appartiene ad U possiamo dedurre che U non è un sottospazio, ma se $\mathbf{0} \in U$ non possiamo dedurre niente. Più formalmente, diciamo sinteticamente che $\mathbf{0} \in W$ è *condizione necessaria* perché W sia un sottospazio vettoriale (ma non sufficiente).

OSSERVAZIONE 2.13. Un'**equazione lineare omogenea** in n incognite è un'equazione della forma

$$a_1x_1 + \ldots + a_nx_n = 0$$

dove a_i sono coefficienti numerici. Negli esempi (1), (3), (4) visti sopra W coincideva con l'insieme delle soluzioni di una particolare equazione lineare omogenea. In tutti quei casi abbiamo verificato che W è un sottospazio.

Come vedremo più avanti questo è un fatto del tutto generale. Se W è l'insieme delle soluzioni di un'equazione lineare omogenea in n incognite, allora è un sottospazio.

Osserviamo che per un'equazione lineare **non** omogenea (in cui, cioè, il termine noto è non nullo), il vettore nullo non è mai soluzione. Segue che l'insieme delle soluzioni di un'equazione lineare non omogenea non è mai un sottospazio.

3.4. Operazioni su sottospazi.

Dati due sottoinsiemi $A, B \subset E$ si possono considerare altri due sottoinsiemi $A \cup B$, $A \cap B$. Ricordiamo che $A \cup B$ è l'insieme degli elementi che appartengono ad A **o** a B, mentre $A \cap B$ è l'insieme degli elementi che appartengono ad A **e** a B (sezione 1.2 del capitolo introduttivo 0).

Sia V uno spazio vettoriale e U, W siano due sottospazi di V. Domanda: in generale $U \cap W$ è un sottospazio? $U \cup W$ è un sottospazio?

Studiamo prima l'**intersezione**.

Consideriamo il caso $\mathbb{E}^3_O$. Abbiamo visto che i sottospazi di $\mathbb{E}^3_O$ sono

- le rette per O,
- i piani per O,
- $\mathbb{E}^3_O$,
- $\left\{\overrightarrow{OO}\right\}$.

Dunque dati U e W sottospazi di $\mathbb{E}^3_O$ per capire cos'è $U \cap W$ bisogna distinguere vari casi. Consideriamone alcuni:

(1) U e W sono piani per O: allora $U \cap W$ è una retta per O (a meno che $U = W$ nel qual caso $U \cap W = U$ è un piano per O);
(2) U e W sono rette per O: allora se le due rette non coincidono $U \cap W = \left\{\overrightarrow{OO}\right\}$ se invece coincidono $U \cap W = U$ è una retta per O.
(3) U è un piano per O mentre W è una retta per O: se W è contenuta in U allora $U \cap W = W$ è una retta per O altrimenti $U \cap W = \left\{\overrightarrow{OO}\right\}$.

Come si vede in tutti questi casi $U \cap W$ è sempre un sottospazio. In effetti vale questo risultato generale:

PROPOSIZIONE 2.3. *Siano U, W sottospazi di uno spazio vettoriale V. Allora $U \cap W$ è un sottospazio.*

DIMOSTRAZIONE.

- Come sappiamo, conviene come prima cosa verificare che $\mathbf{0}_V \in U \cap W$: poiché U e W sono sottospazi allora $\mathbf{0}_V$ appartiene ad entrambi.

- Verifichiamo ora che comunque dati $\boldsymbol{v}, \boldsymbol{v}' \in U \cap W$ si ha che $\boldsymbol{v} + \boldsymbol{v}' \in U \cap W$. Riportiamo in una tabella prima tutte le informazione che abbiamo su $U, W, \boldsymbol{v}, \boldsymbol{v}'$ (le ipotesi) e poi ciò che vogliamo verificare.

Ipotesi:	1) U, W sono sottospazi, 2) $\boldsymbol{v}$ appartiene sia ad U che a W. 3) $\boldsymbol{v}'$ appartiene sia ad U che a W.
Tesi:	$\boldsymbol{v} + \boldsymbol{v}'$ appartiene sia ad U che a W.

  Ora osserviamo che poiché $\boldsymbol{v}$ e $\boldsymbol{v}'$ appartengono entrambi ad U ed U è un sottospazio, si ha che $\boldsymbol{v} + \boldsymbol{v}'$ appartiene ad U. Analogamente $\boldsymbol{v}$ e $\boldsymbol{v}'$ appartengono entrambi a W che per ipotesi è un sottospazio. Dunque $\boldsymbol{v} + \boldsymbol{v}'$ appartiene anche a W.
- Infine bisogna verificare che comunque dati $\boldsymbol{v} \in U \cap W$ e $\lambda \in \mathbb{R}$ si ha $\lambda\boldsymbol{v} \in U \cap W$. Anche in questo caso possiamo riportare una tabella come sopra:

Ipotesi:	1) U, W sono sottospazi, 2) $\boldsymbol{v}$ appartiene sia ad U che a W.
Tesi:	$\lambda\boldsymbol{v}$ appartiene sia ad U che a W.

  Poiché $\boldsymbol{v} \in U$ e U è un sottospazio, si ha che $\lambda\boldsymbol{v} \in U$. Analogamente poiché $\boldsymbol{v} \in W$ e W è un sottospazio abbiamo $\lambda\boldsymbol{v} \in W$. Dunque $\lambda\boldsymbol{v}$ appartiene contemporaneamente a U e W.

$\square$

Passiamo ora a considerare l'**unione**.

Osserviamo in questo caso che in $\mathbb{E}^3_O$ date due rette distinte U, W passanti per O l'unione $U \cup W$ non è né un piano, né una retta, né tutto né $\left\{\overrightarrow{OO}\right\}$. Dunque in generale l'unione di due sottospazi **NON** è un sottospazio!

Cerchiamo di capire più da vicino perché l'unione di due rette non è un sottospazio di $\mathbb{E}^3_O$. Notiamo che $\overrightarrow{OO} \in U \cup W$.

In effetti è anche vero che se $\boldsymbol{v} \in U \cup W$ allora $\lambda\boldsymbol{v} \in U \cup W$.

Osserviamo invece che presi due vettori non nulli $\boldsymbol{u}, \boldsymbol{w} \in U \cup W$ tali che $\boldsymbol{u} \notin W$ e $\boldsymbol{w} \notin U$, la somma $\boldsymbol{u} + \boldsymbol{w}$ non appartiene né a U né a W. Dunque $\boldsymbol{u} + \boldsymbol{w} \notin U \cup W$, il che mostra che la proprietà (SS1) **NON** è verificata (vedi Figura 2.3).

D'altra parte, abbiamo già osservato nel precedente capitolo che esiste un unico piano Z che contiene entrambe le rette U e W. Inoltre ogni vettore di Z si scompone in una parte lungo U e in una parte lungo W.

Dunque il piano Z può essere caratterizzato algebricamente come l'insieme dei vettori che si ottengono sommando un vettore di U e un vettore di W. In simboli

$$Z = \{\boldsymbol{u} + \boldsymbol{w} \mid \boldsymbol{u} \in U \text{ e } \boldsymbol{w} \in W\} .$$

In generale possiamo ripetere questa costruzione per due sottospazi di un qualsiasi spazio vettoriale.

Definizione 2.5. Siano U, W sottospazi di uno spazio vettoriale V. Allora

$$U + W = \{\boldsymbol{u} + \boldsymbol{w} \mid \boldsymbol{u} \in U \text{ e } \boldsymbol{w} \in W\} .$$

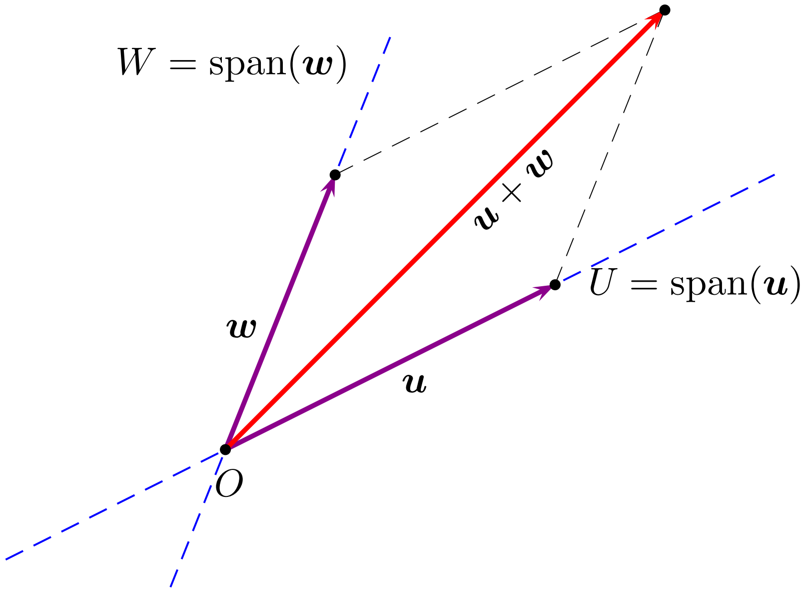

FIGURA 2.3. Presi due vettori $\boldsymbol{u}, \boldsymbol{w}$ non allineati,siano $U = \mathrm{Span}(\boldsymbol{u})$ e $W = \mathrm{Span}(\boldsymbol{w})$. Così facendo, $\boldsymbol{u}, \boldsymbol{w} \in U \cup W$, ma $\boldsymbol{u} \notin W$ e $\boldsymbol{w} \notin U$; la somma $\boldsymbol{u}+\boldsymbol{w}$ non appartiene né a W né a U.

OSSERVAZIONE 2.14. Segnaliamo un paio di precisazioni.

(1) Nel caso U, W sono due rette distinte passanti per O, allora $U+W$ ha l'interpretazione geometrica chiara di piano passante per l'origine.
(2) Attenzione a non confondere $U \cup W$ con $U+W$!

Prima di proseguire con la teoria fermiamoci un attimo a riflettere su cosa sia $U+W$. Per determinare un sottoinsieme di V bisogna determinarne gli elementi, ovvero va determinato un criterio per distinguere gli elementi che appartengono al sottoinsieme da quelli che non vi appartengono. In questo caso dati due sottospazi U e W e un vettore $\boldsymbol{v}$, se ci si chiede se $\boldsymbol{v} \in U+W$, ci si deve porre la seguente domanda:

Posso scomporre il vettore $\boldsymbol{v}$ come somma di un vettore $\boldsymbol{u} \in U$ e un vettore $\boldsymbol{w} \in W$?

Ovvero devo cercare un vettore $\boldsymbol{u} \in U$ e un vettore $\boldsymbol{w} \in W$ tali che la loro somma sia $\boldsymbol{v}$.

Se si riesce a trovare $\boldsymbol{u}$ e $\boldsymbol{w}$, allora si conclude che $\boldsymbol{v} \in U+W$, se invece ci si accorge che non è possibile trovare tali vettori allora si conclude che $\boldsymbol{v} \notin U+W$.

ESEMPIO 2.15. Consideriamo degli esempi:

(1) Siano π ed r rispettivamente un piano e una retta in $\mathbb{E}^3_O$ passanti per O e supponiamo che la retta r non sia contenuta in π.
Abbiamo visto nel capitolo precedente che tutti i vettori di $\mathbb{E}^3_O$ si scompongono come somma di un vettore in π e un vettore in r (vedi

Figura 2.4). Dunque in questo caso si ha

$$r + \pi = \mathbb{E}^3_O\,.$$

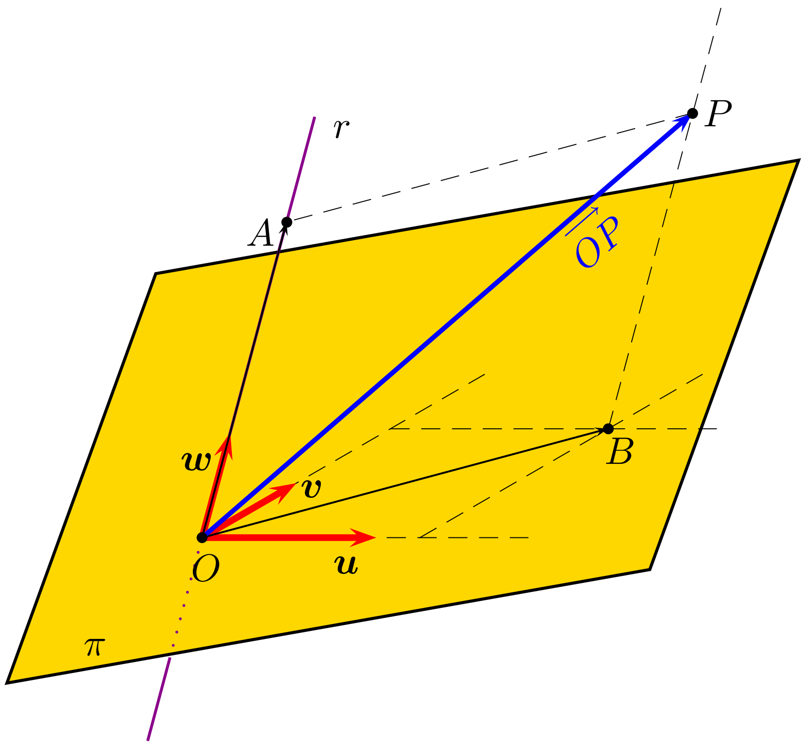

FIGURA 2.4. Ogni vettore $\overrightarrow{OP}$ di $\mathbb{E}^3_O$ può essere decomposto nella somma di un vettore lungo una retta r passante per O ed uno posto in un piano π passante per O e non contenente la retta r.

(2) In $\mathbb{R}^3$ consideriamo i sottospazi

$$U = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mid x_1 = x_2 = 0 \right\} \text{ e } W = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mid x_2 = x_3 = 0 \right\}$$

Quali tra questi vettori appartiene a $U + W$?

$$\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \quad \begin{pmatrix} 3 \\ 0 \\ -5 \end{pmatrix} \quad \begin{pmatrix} 1 \\ 1 \\ 10 \end{pmatrix}.$$

Osserviamo che i vettori di U sono del tipo $\begin{pmatrix} 0 \\ 0 \\ a \end{pmatrix}$ mentre i vettori di W sono del tipo $\begin{pmatrix} b \\ 0 \\ 0 \end{pmatrix}$.

Dunque per verificare se $\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$ appartiene a $U + W$ mi devo chiedere se trovo $a, b \in \mathbb{R}$ tali che

$$\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ a \end{pmatrix} + \begin{pmatrix} b \\ 0 \\ 0 \end{pmatrix}$$

Osserviamo che basta porre $a = -1$ e $b = 1$

Analogamente per verificare che $\begin{pmatrix} 3 \\ 0 \\ 5 \end{pmatrix}$ appartiene a $U+W$ dobbiamo trovare a, b tali che

$$\begin{pmatrix} 3 \\ 0 \\ 5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ a \end{pmatrix} + \begin{pmatrix} b \\ 0 \\ 0 \end{pmatrix}.$$

Osserviamo che basta porre $a = 5$ e $b = 3$.

Invece per verificare se $\begin{pmatrix} 1 \\ 1 \\ 10 \end{pmatrix}$ appartiene a $U+W$ dobbiamo trovare a, b tali che

$$\begin{pmatrix} 1 \\ 1 \\ 10 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ a \end{pmatrix} + \begin{pmatrix} b \\ 0 \\ 0 \end{pmatrix}.$$

Notiamo che in questo caso non è possibile trovare a e b. Ciò dipende dal fatto che se sommo vettori di U e di W ottengo comunque vettori la cui seconda componente è nulla. Viceversa dato un qualsiasi vettore la cui seconda componente è nulla riesco ad ottenerlo come somma di un vettore in U e di un vettore in W

$$\begin{pmatrix} a \\ 0 \\ b \end{pmatrix} = \begin{pmatrix} a \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ b \end{pmatrix}$$

Dunque

$$U + W = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mid x_2 = 0 \right\}$$

(3) In $\mathbb{R}^3$ consideriamo

$$U = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mid x_1 = x_2 = 0 \right\} \text{ e } W = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mid x_1 + x_2 = x_3 = 0 \right\}.$$

Proviamo a verificare quali tra questi vettori appartiene a $U + W$

$$\begin{pmatrix} 5 \\ -5 \\ 3 \end{pmatrix} \quad \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} \quad \begin{pmatrix} 7 \\ -7 \\ 2 \end{pmatrix}.$$

Un vettore di U è della forma $\begin{pmatrix} 0 \\ 0 \\ a \end{pmatrix}$ dove il numero a può essere arbitrario. D'altra parte i vettori di W sono caratterizzati dal fatto che la terza componente è nulla mentre la seconda componente è l'opposto della prima, in altre parole sono tutti vettori della forma $\begin{pmatrix} b \\ -b \\ 0 \end{pmatrix}$ dove ancora b può essere un numero qualunque.

Dunque per verificare se il vettore $\begin{pmatrix} 5 \\ -5 \\ 3 \end{pmatrix}$ appartiene ad $U + W$ ci dobbiamo chiedere se esistono due numeri a, b tali che

$$\begin{pmatrix} 5 \\ -5 \\ 3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ a \end{pmatrix} + \begin{pmatrix} b \\ -b \\ 0 \end{pmatrix}$$

Osserviamo che ponendo $a = 3$ e $b = 5$ l'uguaglianza è verificata e dunque $\begin{pmatrix} 5 \\ -5 \\ 3 \end{pmatrix} \in U + W$.

In generale lo stesso ragionamento mostra che i vettori di $U + W$ sono della forma

$$\begin{pmatrix} b \\ -b \\ a \end{pmatrix} \tag{2.7}$$

dove a e b sono numeri arbitrari.

Notiamo che i vettori della forma (2.7) hanno la proprietà che la somma delle prime due componenti si annulla (o equivalentemente che la seconda componente è l'opposto della prima). In particolare il vettore $\begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}$ non può essere scritto nella forma (2.7) e dunque non appartiene ad $U + W$.

Viceversa ogni vettore che ha soddisfa la proprietà che la somma delle prime due componenti si annulla può essere scritto nella forma

(2.7) e dunque appartiene ad $U+W$. Ad esempio il vettore $\begin{pmatrix} 7 \\ -7 \\ 2 \end{pmatrix}$ si può scrivere nella forma (2.7) ponendo $b = 7$ e $a = 2$.

In definitiva in questo caso abbiamo che

$$U + W = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} | x_1 + x_2 = 0 \right\} .$$

Vale la pena di sottolineare che il fatto che x_3 non compaia nell'equazione $x_1 + x_2 = 0$ non vuol dire che le soluzioni di tale equazioni hanno ultima componente nulla. Infatti una soluzione in $\mathbb{R}^3$ dell'equazione $x_1 + x_2 = 0$ è una terna di numeri tale che la somma delle prime due componenti è nulla. Dunque non si impone alcuna condizione su x_3 che per tale motivo è libera di assumere qualsiasi valore indipendentemente dai valori di x_1 e x_2. Sono ad esempio soluzioni i vettori:

$$\begin{pmatrix} 7 \\ -7 \\ 2 \end{pmatrix}, \quad \begin{pmatrix} 7 \\ -7 \\ 18 \end{pmatrix}, \quad \begin{pmatrix} 7 \\ -7 \\ -15 \end{pmatrix}, \ldots .$$

Dalla definizione è chiaro che $U + W$ è un ben definito sottoinsieme di V. Nel caso geometrico considerato in realtà abbiamo visto che $U + W$ è proprio un sottospazio di V. In effetti questo fatto è del tutto generale:

Proposizione 2.4. *Siano U, W sottospazi di V. Allora $U + W$ è un sottospazio di V.*

Dimostrazione. Come prima cosa mostriamo che $U+W$ è un sottospazio.

- $\mathbf{0}_V \in U + W$: per verificare che il vettore nullo appartiene a $U + W$, bisogna riuscire a scrivere il vettore nullo come somma di un elemento di $\boldsymbol{u} \in U$ e di un elemento di $\boldsymbol{w} \in W$.

  Osserviamo che $\mathbf{0}_V$ appartiene sia ad U che a W e dunque possiamo porre entrambi $\boldsymbol{u}$ e $\boldsymbol{w}$ uguali a $\mathbf{0}_V$. Chiaramente si ha

  $$\mathbf{0}_V = \mathbf{0}_V + \mathbf{0}_V$$

  e dunque $\mathbf{0}_V \in U + W$.
- Bisogna verificare che comunque presi $\boldsymbol{v}, \boldsymbol{v}' \in U+W$ si ha che $\boldsymbol{v}+\boldsymbol{v}' \in U + W$. Riassiumiamo in una tabella ipotesi e tesi:

Ipotesi:	1) U, W sono sottospazi, 2) $\boldsymbol{v} = \boldsymbol{u} + \boldsymbol{w}$ con $\boldsymbol{u} \in U$ e $\boldsymbol{w} \in W$ 3) $\boldsymbol{v}' = \boldsymbol{u}' + \boldsymbol{w}'$ con $\boldsymbol{u}' \in U$ e $\boldsymbol{w}' \in W$.
Tesi:	Esistono $\boldsymbol{u}'' \in U$ e $\boldsymbol{w}'' \in W$ tali che $\boldsymbol{v} + \boldsymbol{v}' = \boldsymbol{u}'' + \boldsymbol{w}''$.

  Osserviamo che

  $$\boldsymbol{v} + \boldsymbol{v}' = (\boldsymbol{u} + \boldsymbol{w}) + (\boldsymbol{u}' + \boldsymbol{w}') = (\boldsymbol{u} + \boldsymbol{u}') + (\boldsymbol{w} + \boldsymbol{w}')$$

  per cui se poniamo

  $$\boldsymbol{u}'' = \boldsymbol{u} + \boldsymbol{u}' \qquad\qquad \boldsymbol{w}'' = \boldsymbol{w} + \boldsymbol{w}' \in W$$

abbiamo che $\boldsymbol{u}'' \in U$ in quanto $\boldsymbol{u}, \boldsymbol{u} \in U$ e U è un sottospazio. Analogamente $\boldsymbol{w}'' \in W$ in quanto somma di $\boldsymbol{w}$ e $\boldsymbol{w}'$. Infine $\boldsymbol{v}+\boldsymbol{v}' = \boldsymbol{u}''+\boldsymbol{w}''$, e dunque siamo riusciti ad esprimere $\boldsymbol{v}+\boldsymbol{v}'$ come somma di un vettore di U e di un vettore di W.

- Bisogna verificare che se $\boldsymbol{v} \in U+W$ allora per ogni $\lambda \in \mathbb{R}$ si ha che $\lambda\boldsymbol{v} \in U+W$. Ora abbiamo

Ipotesi:	1) U, W sono sottospazi, 2) $\boldsymbol{v} = \boldsymbol{u}+\boldsymbol{w}$ con $\boldsymbol{u} \in U$ e $\boldsymbol{w} \in W$
Tesi:	Esistono $\boldsymbol{u}' \in U$ e $\boldsymbol{w}' \in W$ tali che $\lambda\boldsymbol{v} = u' + w'$.

Osserviamo che
$$\lambda\boldsymbol{v} = \lambda(\boldsymbol{u}+\boldsymbol{w}) = \lambda\boldsymbol{u} + \lambda\boldsymbol{w}$$
ed essendo U e W si ha che $\lambda\boldsymbol{u} \in U$ e $\lambda\boldsymbol{w} \in W$. Dunque posto $\boldsymbol{u}' = \lambda\boldsymbol{u}$ e $\boldsymbol{w}' = \lambda\boldsymbol{w}$ si ottiene la tesi.

□

Osservazione 2.16. Suggeriamo ora alcune osservazioni sul comportamento della somma di sottospazi:

(1) Siano U, W sottospazi di uno spazio vettoriale V. Allora U **e** W **sono entrambi contenuti in** $U+W$. Verifichiamo tale proprietà in un esempio.

Siano
$$U = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mid x_2 = x_3 = 0 \right\}$$
e
$$W = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mid 3x_1 - 7x_2 + 8x_3 = x_1 + x_2 - x_3 = 0 \right\}.$$
Verifichiamo se il vettore $\begin{pmatrix} 4 \\ 0 \\ 0 \end{pmatrix}$ appartiene a $U+W$. Come al solito mi devo chiedere se esistono vettori $\boldsymbol{u} \in U, \boldsymbol{w} \in W$ tali che
$$\begin{pmatrix} 4 \\ 0 \\ 0 \end{pmatrix} = \boldsymbol{u} + \boldsymbol{w}\,.$$
Ora osservate che $\begin{pmatrix} 4 \\ 0 \\ 0 \end{pmatrix} \in U$ e dunque se poniamo
$$\boldsymbol{u} = \begin{pmatrix} 4 \\ 0 \\ 0 \end{pmatrix} \qquad \boldsymbol{w} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$
abbiamo $\begin{pmatrix} 4 \\ 0 \\ 0 \end{pmatrix} = \boldsymbol{u}+\boldsymbol{w}$, $\boldsymbol{u} \in U$ e $\boldsymbol{w} \in W$. Dunque $\begin{pmatrix} 4 \\ 0 \\ 0 \end{pmatrix} \in U+W$.

(2) Sia U un sottospazio di uno spazio vettoriale V.
Allora si ha $\mathbf{U} + \mathbf{U} = \mathbf{U}$.

Illustriamo anche questa proprietà con un esempio.

$$U = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mid x_2 = x_3 = 0 \right\}.$$

Cerchiamo di capire chi è $U + U$. Dall'osservazione precedente $U + U$ è un sottospazio che include U.

Cerchiamo ora di capire se $U + U$ contiene vettori non appartenti ad U oppure coincide con U. Osserviamo che il vettore generico di U si scrive nella forma $\begin{pmatrix} a \\ 0 \\ 0 \end{pmatrix}$. Se sommo due vettori generici di U ottengo

$$\begin{pmatrix} a \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} b \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} a+b \\ 0 \\ 0 \end{pmatrix}.$$

Dunque se $\boldsymbol{v} \in U + U$ esso deve avere seconda e terza componente nulla. Dunque deve appartenere ad U.

In definitiva si ha $U + U = U$;

(3) Si può generalizzare il precedente risultato nel seguente modo. **se $\mathbf{U} \subset \mathbf{W}$ allora $\mathbf{U} + \mathbf{W} = \mathbf{W}$**: infatti chiaramente $W \subset U + W$. Viceversa se $\boldsymbol{v} \in U + W$, allora $\boldsymbol{v}$ si scrive come somma di un vettore $\boldsymbol{u} \in U$ e un vettore $\boldsymbol{w} \in W$. Poiché stiamo supponendo che $U \subset W$, risulta che $\boldsymbol{u} \in W$ e dunque $\boldsymbol{v} = \boldsymbol{u} + \boldsymbol{w}$ è la somma di due vettori in W e di conseguenza appartiene a W.

(4) $(U_1 + U_2) + U_3 = U_1 + (U_2 + U_3) = \{u_1 + u_2 + u_3 \mid u_i \in U_i\}$ Per questa ragione dati k sottospazi $U_1, \ldots U_k$ denotiamo con

$$\mathbf{U_1} + \mathbf{U_2} + \ldots + \mathbf{U_k} = \{\mathbf{u_1} + \mathbf{u_2} + \ldots + \mathbf{u_k} \mid \mathbf{u_i} \in \mathbf{U_i}\}$$

(5) Dati tre sottospazi U, W, Z ci si potrebbe chiedere se gli elementi di Z che appartengono a $U + W$ si possono decomporre come somma di un elemento di $U \cap Z$ e un elemento di $W \cap Z$, ovvero che valga la seguente uguaglianza $(U + W) \cap Z = (U \cap Z) + (W \cap Z)$.

Tale uguaglianza è in generale FALSA ciò vuol dire che gli elementi di Z che appartengono ad $U + W$ si decompongono come somma di un elemento $\boldsymbol{u} \in U$ e un elemento $\boldsymbol{w} \in W$, ma in generale non possiamo chiedere che $\boldsymbol{u}, \boldsymbol{w} \in Z$.

Basta considerare il seguente esempio. Siano U_1, U_2, U_3 tre rette distinte contenute in un piano W. Allora $(U_1 + U_2) = W$ e dunque $(U_1 + U_2) \cap U_3 = W \cap U_3 = U_3$. D'altra parte $U_1 \cap U_3 = \left\{\overrightarrow{OO}\right\}$ e $U_2 \cap U_3 = \left\{\overrightarrow{OO}\right\}$ e dunque $(U_1 \cap U_3) + (U_2 \cap U_3) = \left\{\overrightarrow{OO}\right\} + \left\{\overrightarrow{OO}\right\} = \left\{\overrightarrow{OO}\right\}$.

4. Il sottospazio generato da un insieme di vettori

Fin qui abbiamo introdotto il concetto di sottospazio come generalizzazione della nozione di piano e retta passante per O e abbiamo anche considerato due operazioni sui sottospazi. Nonostante fin qui abbiamo fatto alcuni esempi, ci manca ancora una teoria generale su come lavorare con spazi vettoriali astratti e su come descrivere un sottospazio.

Nel caso geometrico la descrizione algebrica di $\mathbb{E}^3_O$ è stata resa possibile dall'identificazione dei vettori applicati con triple di numeri. Tale identificazione è conseguenza dell'introduzione della nozione di base e della nozione di coordinate.

Ripercorriamo sinteticamente i principali passaggi che ci hanno permesso di associare ad un vettore applicato la terna delle coordinate.

Dato un vettore $\boldsymbol{u}_0 \in \mathbb{E}^3_O$, esso individua una retta U. Abbiamo visto che i vettori di U si ottengono tutti moltiplicando $\boldsymbol{u}_0$ per qualche scalare. Più precisamente

$$U = \{\lambda \boldsymbol{u}_0 \,|\, \lambda \in \mathbb{R}\}$$

Dati due vettori non paralleli $\boldsymbol{u}_0, \boldsymbol{v}_0$, essi individuano due rette non coincidenti U, V e abbiamo visto che il piano contenente U e V è la somma di $U + V$.

Ciò vuol dire che tutti i vettori sul piano $U + V$ possono essere scritti come la somma di un vettore di U e un vettore di V. Ovvero è possibile decomporre ogni $\boldsymbol{w} \in U + V$ in una parte $\boldsymbol{u} \in U$ e una parte $\boldsymbol{v} \in V$ in modo che $\boldsymbol{w} = \boldsymbol{u} + \boldsymbol{v}$. Del resto poiché $\boldsymbol{u} \in U$, esiste $\lambda \in \mathbb{R}$ tale che $\boldsymbol{u} = \lambda \boldsymbol{u}_0$ e analogamente esiste $\mu \in \mathbb{R}$ tale che $\boldsymbol{v} = \mu \boldsymbol{v}_0$ e dunque ogni vettore $\boldsymbol{w} \in U + V$ può essere scritto nella forma

$$\boldsymbol{w} = \lambda \boldsymbol{u}_0 + \mu \boldsymbol{v}_0\,.$$

In particolare $U + V = \{\lambda \boldsymbol{u}_0 + \mu \boldsymbol{v}_0 \,|\, \lambda, \mu \in \mathbb{R}\}$.

Ora se fissiamo un vettore $\boldsymbol{w}_0 \in \mathbb{E}^3_O$ non appartenente al piano $U + V$, esso individua una retta $W = \{\lambda \boldsymbol{w}_0 | \lambda \in \mathbb{R}\}$ che non è contenuta nel piano $U + V$. Da quanto già visto

$$(U + V) + W = \mathbb{E}^3_O$$

ovvero ogni vettore $\boldsymbol{v} \in \mathbb{E}^3_O$ può essere decomposto in una parte $\boldsymbol{u} \in U + V$ e una parte $\boldsymbol{w} \in W$ in modo che $\boldsymbol{v} = \boldsymbol{u} + \boldsymbol{z}$.

Ora il vettore $\boldsymbol{u}$ può essere scritto nella forma $\boldsymbol{u} = \lambda \boldsymbol{u}_0 + \mu \boldsymbol{v}_0$ mentre il vettore $\boldsymbol{w}$ è della forma $\boldsymbol{w} = \nu \boldsymbol{w}_0$.

In definitiva otteniamo che

$$\boldsymbol{v} = \lambda \boldsymbol{u}_0 + \mu \boldsymbol{v}_0 + \nu \boldsymbol{w}_0\,.$$

Riassumendo, per ogni vettore $\boldsymbol{v} \in \mathbb{E}^3_O$ possiamo trovare una tripla di numeri $\begin{pmatrix} \lambda \\ \mu \\ \nu \end{pmatrix}$ tali che

$$\boldsymbol{v} = \lambda \boldsymbol{u}_0 + \mu \boldsymbol{v}_0 + \nu \boldsymbol{w}_0\,.$$

I numeri λ, μ, ν sono le coordinate del vettore $\boldsymbol{v}$ rispetto alla base $\{\boldsymbol{u}_0, \boldsymbol{v}_0, \boldsymbol{w}_0\}$.

In questo modo riesco ad associare ad ogni vettore geometrico una tripla di numeri, e in tal modo costruisco un'identificazione tra lo spazio $\mathbb{E}^3_O$ e lo spazio $\mathbb{R}^3$.

Siamo ora interessati ad estendere la nozione di base e di coordinate per qualsiasi spazio vettoriale, in modo da poter avere una descrizione di qualsiasi spazio vettoriale per mezzo di n-uple di numeri.

Proviamo a riprodurre la costruzione fatta per $\mathbb{E}^3_O$ per un qualsiasi spazio vettoriale V.

Definizione 2.6. Fissato un vettore $\boldsymbol{u}$ in V possiamo definire $\mathrm{Span}(\boldsymbol{u})$ esattamente come nel caso geometrico

$$\mathrm{Span}(\boldsymbol{u}) = \{\lambda\boldsymbol{u} \mid \lambda \in \mathbb{R}\}$$

ovvero $\mathrm{Span}(\boldsymbol{u})$ è l'insieme di tutti i multipli di $\boldsymbol{u}$.

Lemma 2.5. *$\mathrm{Span}(\boldsymbol{u})$ è un sottospazio di V.*

Dimostrazione.

- Per verificare che $\mathbf{0}_V \in \mathrm{Span}(\boldsymbol{u})$, dobbiamo controllare che $\mathbf{0}_V$ si possa scrivere come multiplo di $\boldsymbol{u}$ (ovvero nella forma $\lambda\boldsymbol{u}$). Del resto ciò segue facilmente dall'identità

$$\mathbf{0}_V = 0 \cdot \boldsymbol{u}\,.$$

- Verifichiamo ora che comunque dati $\boldsymbol{v}, \boldsymbol{v}' \in \mathrm{Span}(\boldsymbol{u})$ la loro somma $\boldsymbol{v} + \boldsymbol{v}'$ appartiene ancora a $\mathrm{Span}(\boldsymbol{u})$. Riassumiamo in una tabella ipotesi e tesi.

Ipotesi:	1) $\boldsymbol{v} = \mu\boldsymbol{u}$ per un certo $\mu \in \mathbb{R}$. 2) $\boldsymbol{v}' = \mu'\boldsymbol{u}$ per un certo $\mu' \in \mathbb{R}$.
Tesi:	Esiste un numero μ'' tale che $\boldsymbol{v} + \boldsymbol{v}' = \mu''\boldsymbol{u}$.

Osserviamo che dall'ipotesi si ha

$$\boldsymbol{v} + \boldsymbol{v}' = \mu\boldsymbol{u} + \mu'\boldsymbol{u} = (\mu + \mu')\boldsymbol{u}$$

da cui si evince che anche $\boldsymbol{v} + \boldsymbol{v}'$ è multiplo di $\boldsymbol{u}$ e quindi appartiene a $\mathrm{Span}(\boldsymbol{u})$.

- Verifichiamo infine che se $\boldsymbol{v} \in \mathrm{Span}(\boldsymbol{u})$ allora $\lambda\boldsymbol{v}$ appartiene a $\mathrm{Span}(\boldsymbol{u})$.

Ipotesi:	$\boldsymbol{v} = \mu\boldsymbol{u}$ per un certo $\mu \in \mathbb{R}$.
Tesi:	Esiste un numero μ' tale che $\lambda\boldsymbol{v} = \mu'\boldsymbol{u}$.

Osserviamo che si ha $\lambda\boldsymbol{v} = \lambda(\mu\boldsymbol{u}) = (\lambda\mu)\boldsymbol{u}$. Per cui abbiamo $\lambda\boldsymbol{v} = \mu'\boldsymbol{u}$ con $\mu' = \lambda\mu$.

□

Siano ora $\boldsymbol{u}$ e $\boldsymbol{w}$ due vettori. Possiamo considerare i due spazi $U = \mathrm{Span}(\boldsymbol{u})$ e $W = \mathrm{Span}(\boldsymbol{w})$ e come prima possiamo farne la somma. Esattamente come nel caso dei vettori applicati

$$U + W = \{\lambda\boldsymbol{u} + \mu\boldsymbol{w} | \lambda, \mu \in \mathbb{R}\}$$

ovvero i vettori di $U+W$ sono quelli che si possono scrivere nella forma $\lambda\boldsymbol{u}+\mu\boldsymbol{w}$.

Ora fissati tre vettori $\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3$ possiamo considerare la somma dei tre sottospazi $U_1 = \mathrm{Span}(\boldsymbol{u}_1), U_2 = \mathrm{Span}(\boldsymbol{u}_2), U_3 = \mathrm{Span}(\boldsymbol{u}_3)$. Esso è un sottospazio e come prima si può mostrare che

$$U_1 + U_2 + U_3 = \{\lambda_1\boldsymbol{u}_1 + \lambda_2\boldsymbol{u}_2 + \lambda_3\boldsymbol{u}_3 \mid \lambda_i \in \mathbb{R}\}$$

e così via.

Tutto ciò suggerisce la seguente definizione.

Definizione 2.7. Sia V spazio vettoriale Dati k vettori $\boldsymbol{u}_1 \dots \boldsymbol{u}_k \in V$, una **combinazione lineare** di $\boldsymbol{u}_1, \dots, \boldsymbol{u}_k$ è una somma del tipo

$$\lambda_1\boldsymbol{u}_1 + \lambda_2\boldsymbol{u}_2 + \ldots + \lambda_k\boldsymbol{u}_k\,.$$

Due combinazioni lineari si dicono uguali, se i rispettivi coefficienti sono a due a due uguali.

Ad esempio fissati 4 vettori $\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3, \boldsymbol{u}_4$ in V, una possibile combinazione lineare di questi vettori è la somma

$$3\boldsymbol{u}_1 - \boldsymbol{u}_2 + 2\boldsymbol{u}_3 - 7\boldsymbol{u}_4\,.$$

In questo caso i coefficienti della combinazione sono $\lambda_1 = 3,\ \lambda_2 = -1,\ \lambda_3 = 2,\ \lambda_4 = -7$.

Anche le seguenti sono combinazioni lineari di $\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3, \boldsymbol{u}_4$

$$\begin{aligned}&\boldsymbol{u}_1 + 7\boldsymbol{u}_3 - \boldsymbol{u}_4\\&\boldsymbol{u}_1 + \boldsymbol{u}_3\\&\boldsymbol{u}_4\end{aligned}$$

infatti il fatto che non compaia il vettore $\boldsymbol{u}_2$ nella prima combinazione vuol dire che il coefficiente per cui viene moltiplicato è 0. I coefficienti di tali combinazioni lineari sono rispettivamente

$$\begin{aligned}&\lambda_1 = 1,\ \lambda_2 = 0\ \lambda_3 = 7,\ \lambda_4 = -1\\&\lambda_1 = 1,\ \lambda_2 = 0,\ \lambda_3 = 1,\ \lambda_4 = 0\\&\lambda_1 = 0,\ \lambda_2 = 0,\ \lambda_3 = 0,\ \lambda_4 = 1\,.\end{aligned}$$

Esercizio 4.1. Fissato $V = \mathbb{R}^3$ e posti

$$\boldsymbol{u}_1 = \begin{pmatrix}1\\1\\0\end{pmatrix},\ \boldsymbol{u}_2 = \begin{pmatrix}1\\-1\\0\end{pmatrix},\ \boldsymbol{u}_3 = \begin{pmatrix}0\\0\\1\end{pmatrix},\ \boldsymbol{u}_4 = \begin{pmatrix}1\\1\\1\end{pmatrix}$$

si calcoli il risultato delle combinazioni lineari considerate sopra.

Definizione 2.8. Sia V spazio vettoriale. Data una lista di vettori

$$\{\boldsymbol{u}_1, \dots, \boldsymbol{u}_k\}\,,$$

lo **span** di tale lista è l'insieme di tutti i vettori di V che si ottengono come combinazioni lineari di $\boldsymbol{u}_1, \dots, \boldsymbol{u}_k$.

In simboli

$$\mathrm{Span}(\boldsymbol{u_1}, \dots, \boldsymbol{u_k}) = \{\lambda_1\boldsymbol{u}_1 + \lambda_2\boldsymbol{u}_2 + \ldots + \lambda_k\boldsymbol{u}_k \mid \lambda_i \in \mathbb{R}\}\ .$$

Cerchiamo di capire qual è il criterio per verificare se un vettore $\boldsymbol{v}$ assegnato appartiene a $\mathrm{Span}(\boldsymbol{u}_1, \dots, \boldsymbol{u}_k)$ dove $\boldsymbol{u}_1, \dots, \boldsymbol{u}_k$ sono assegnati.

Dalla definizione risulta che il vettore $\boldsymbol{v}$ appartiene a $\mathrm{Span}(\boldsymbol{u}_1, \dots, \boldsymbol{u}_k)$ se si può scrivere come combinazione lineare di $\boldsymbol{u}_1, \dots, \boldsymbol{u}_k$ mentre non vi appartiene altrimenti.

Dunque bisogna porsi la seguente domanda:

Esistono numeri $\lambda_1, \lambda_2, \lambda_3, \dots, \lambda_k$ tali che

$$\boldsymbol{v} = \lambda_1\boldsymbol{u}_1 + \lambda_2\boldsymbol{u}_2 + \lambda_3\boldsymbol{u}_3 + \dots + \lambda_k\boldsymbol{u}_k\,?$$

Se la risposta alla domanda è positiva allora $\boldsymbol{v} \in \mathrm{Span}(\boldsymbol{u}_1 \dots, \boldsymbol{u}_k)$, se la risposta è negativa allora $\boldsymbol{v}$ non vi appartiene.

Esempio 2.17. Cerchiamo di mettere in pratica questa procedura con alcuni esempi

(1) Posto $v = \begin{pmatrix} 2 \\ 0 \\ -3 \end{pmatrix}$ proviamo a verificare se il vettore $\begin{pmatrix} 8 \\ 0 \\ -12 \end{pmatrix}$ appartiene a $\mathrm{Span}(v)$.

Ci dobbiamo chiedere se esiste un numero λ tale che

$$\begin{pmatrix} 8 \\ 0 \\ -12 \end{pmatrix} = \lambda \begin{pmatrix} 2 \\ 0 \\ -3 \end{pmatrix}$$

In tal caso la risposta è affermativa, infatti basta porre $\lambda = 4$.

(2) Verifichiamo se il vettore $\begin{pmatrix} 8 \\ 0 \\ 8 \end{pmatrix}$ appartiene a $\mathrm{Span}(v)$. In questo caso mi devo chiedere se esiste un numero λ tale che

$$\begin{pmatrix} 8 \\ 0 \\ 8 \end{pmatrix} = \lambda \begin{pmatrix} 2 \\ 0 \\ -3 \end{pmatrix} \tag{2.8}$$

Uguagliando componente per componente le componenti del membro a destra e del membro a sinistar otteniamo le equazioni (dove λ rappresenta l'incognita)

$$\begin{aligned} 8 &= 2\lambda \\ 0 &= 0 \\ 8 &= -3\lambda \end{aligned} \tag{2.9}$$

Notiamo che trovare un λ che verifica la (2.8) vuol dire trovare un λ che verifica simultaneamente tutte e tre le equazioni (2.9) (notate che la seconda $0 = 0$ non dipende da λ ed è verificata qualsiasi sia λ).

Del resto l'unico valore di λ che verifichi la prima equazione è $\lambda = 4$, mentre l'unico valore di λ che verifichi la terza equazione è $\lambda = -\frac{8}{3}$. Poiché $4 \neq -\frac{8}{3}$ segue che non ci sono valori di λ per cui tutte e tre le equazioni siano verificate e dunque non esiste λ tale che la (2.8) sia verificata.

(3) Sia $V = \mathrm{Span}\left(\begin{pmatrix}1\\1\\1\end{pmatrix}, \begin{pmatrix}2\\1\\0\end{pmatrix}\right) \subset \mathbb{R}^3$.

Verifichiamo se il vettore $\begin{pmatrix}5\\7\\4\end{pmatrix}$ appartiene a V. In questo caso dobbiamo chiederci se esistono numeri λ_1, λ_2 tali che

$$\begin{pmatrix}5\\7\\4\end{pmatrix} = \lambda_1 \begin{pmatrix}1\\1\\1\end{pmatrix} + \lambda_2 \begin{pmatrix}2\\1\\0\end{pmatrix} \tag{2.10}$$

Ora si ha che

$$\lambda_1 \begin{pmatrix}1\\1\\1\end{pmatrix} + \lambda_2 \begin{pmatrix}2\\1\\0\end{pmatrix} = \begin{pmatrix}\lambda_1 + 2\lambda_2\\ \lambda_1 + \lambda_2\\ \lambda_1\end{pmatrix}$$

e dunque affinché valga la (2.10) devono valere simultaneamente le seguenti tre equazioni

$$\begin{cases} \lambda_1 + 2\lambda_2 = 5 \\ \lambda_1 + \lambda_2 = 7 \\ \lambda_1 = 4 \end{cases} \tag{2.11}$$

Ovvero il vettore $\begin{pmatrix}5\\7\\4\end{pmatrix}$ appartiene a V se e soltanto se il sistema lineare (2.11) ammette soluzione.

Ora osserviamo che dalla ultima equazione della (2.11) si deduce che $\lambda_1 = 4$. Sostituendo $\lambda_1 = 4$ nella seconda equazione si ricava $4 + \lambda_2 = 7$ ovvero $\lambda_2 = 3$.

Dunque la coppia $\lambda_1 = 4, \lambda_2 = 3$ è l'unica coppia di numeri che rende simultaneamente vere la seconda e la terza equazione.

Poiché noi vogliamo trovare due numeri λ_1, λ_2 che verifichino simultaneamente tutte e tre le equazioni, l'unica possibilità è che la coppia $\lambda_1 = 4, \lambda_2 = 3$ verifichi anche la prima.

Del resto sostituendo nella prima equazione i numeri $4, 3$ si ottiene

$$4 + 2 \cdot 3 = 4 + 6 = 10 \neq 7$$

dunque la coppia $\lambda_1 = 4, \lambda_2 = 3$ non verifica la prima equazione, da cui si deduce che non ci sono coppie di numeri che verificano tutte e e tre le equazioni di (2.11).

Segue che il vettore $\begin{pmatrix}5\\7\\4\end{pmatrix}$ non appartiene a V.

(4) Sia V come nel precedente esempio. Dato un vettore generico $\boldsymbol{v} = \begin{pmatrix}a\\b\\c\end{pmatrix}$ cerchiamo di capire per quali scelte di a, b, c tale vettore appartiene a V.

Ripetendo il ragionamento fatto nei precedenti due esempi cerchiamo λ_1, λ_2 tali che

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \lambda_1 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + \lambda_2 \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}$$

Scrivendo componente per componente questa uguaglianza vettoriale otteniamo il sistema

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

dove le incognite sono λ_1, λ_2, mentre a, b, c sono le componenti del vettore $\boldsymbol{v}$ dato. Dall'ultima ricaviamo che $\lambda_1 = c$. Sostituendo nella seconda equazione ricaviamo come sopra $\lambda_2 = b - c$. Ovvero l'unica coppia di numeri che verifica simultaneamente la seconda e la terza equazione è $\lambda_1 = c, \lambda_2 = b - c$. Dunque il sistema ha soluzione se per tale scelta di numeri anche la prima equazione è verificata, non ha soluzione altrimenti.

Sostituendo $\lambda_1 = c, \lambda_2 = b - c$ nella prima equazione, si ottiene

$$2b - c = a$$

ovvero il vettore $\boldsymbol{v} = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$ appartiene a V se e soltanto se le sue componenti sono soluzioni di tale equazione.

Osservazione 2.18. $\mathrm{Span}(\boldsymbol{u}_1, \boldsymbol{u}_2) = \mathrm{Span}(\boldsymbol{u}_1) + \mathrm{Span}(\boldsymbol{u}_2)$.

Infatti un vettore $\boldsymbol{v}$ appartiene a $\mathrm{Span}(\boldsymbol{u}_1) + \mathrm{Span}(\boldsymbol{u}_2)$ se e soltanto se può essere scritto come somma di due addendi

$$\boldsymbol{v} = \boldsymbol{v}_1 + \boldsymbol{v}_2$$

con $\boldsymbol{v}_1 \in \mathrm{Span}(\boldsymbol{u}_1)$ e $\boldsymbol{v}_2 \in \mathrm{Span}(\boldsymbol{u}_2)$.

Del resto quest'ultima condizione precisamente significa che $\boldsymbol{v}_1$ è multiplo di $\boldsymbol{u}_1$ mentre $\boldsymbol{v}_2$ è multiplo di $\boldsymbol{u}_2$, per cui in definitiva abbiamo che $\boldsymbol{v}$ appartiene a $\mathrm{Span}(\boldsymbol{u}_1) + \mathrm{Span}(\boldsymbol{u}_2)$ se e solo se è della forma

$$\boldsymbol{v} = \lambda_1 \boldsymbol{u}_1 + \lambda_2 \boldsymbol{u}_2$$

ovvero se e solo se è combinazione lineare di $\boldsymbol{u}_1$ e $\boldsymbol{u}_2$.

Più in generale si può vedere nello stesso modo che $\mathrm{Span}(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k) = \mathrm{Span}(\boldsymbol{u}_1) + \mathrm{Span}(\boldsymbol{u}_2) + \ldots + \mathrm{Span}(\boldsymbol{u}_k)$.

Dall'osservazione ricaviamo la seguente proposizione molto importante

Proposizione 2.6. $\mathrm{Span}(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k)$ *è un sottospazio vettoriale di* V.

Definizione 2.9. $\mathrm{Span}(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k)$ è detto sottospazio generato dai vettori $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k$.

Dimostrazione. Per il Lemma 2.5 abbiamo che

$$\mathrm{Span}(\boldsymbol{u}_1), \mathrm{Span}(\boldsymbol{u}_2), \mathrm{Span}(\boldsymbol{u}_3), \ldots, \mathrm{Span}(\boldsymbol{u}_k)$$

sono tutti sottospazi di V.

Del resto abbiamo visto che

$$\mathrm{Span}(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k) = \mathrm{Span}(\boldsymbol{u}_1) + \mathrm{Span}(\boldsymbol{u}_2) + \ldots + \mathrm{Span}(\boldsymbol{u}_k)$$

e poiché la somma di sottospazi è un sottospazio, concludiamo che

$$\mathrm{Span}(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k)$$

è esso stesso un sottospazio. □

Mostriamo ora alcune proprietà dello Span che in seguito utilizzeremo spesso.

Fissiamo un sottospazio W di uno spazio vettoriale V, e consideriamo una lista di vettori $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k$ tutti contenuti in W.

Per le proprietà di sottospazio, si ricava che tutti i multipli di $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k$ sono contenuti in W. Dunque data una combinazione lineare di $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k$

$$\lambda_1 \boldsymbol{u}_1 + \ldots + \lambda_k \boldsymbol{u}_k$$

si ha che tutti gli addendi che vi compaiono sono contenuti in W. Dunque il vettore dato da tale combinazione lineare è somma di vettori in W e dunque è anch'esso un vettore di W.

In conclusione otteniamo che tutti i vettori che si ottengono come combinazione lineare di $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k$ sono tutti vettori di W. Possiamo riassumere tale proprietà nel seguente enunciato.

Lemma 2.7. *Sia W sottospazio di uno spazio vettoriale V, e siano $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k$ vettori di W. Allora* $\mathrm{Span}(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k)$ *è contenuto in W.*

Si noti che dal lemma segue che $\mathrm{Span}(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k)$ è il più piccolo sottospazio che contiene $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k$.

Dal Lemma 2.7 si possono ricavare alcune altre proprietà dello Span che utilizzeremo in seguito.

Lemma 2.8. *Sia V spazio vettoriale, $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k \in V$ e $\boldsymbol{v} \in V$ Valgono le seguenti proprietà:*

(1) *Se $\boldsymbol{v} \in \mathrm{Span}(\boldsymbol{u}_1 \ldots, \boldsymbol{u}_k)$ allora $\mathrm{Span}(\boldsymbol{v}) \subset \mathrm{Span}(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k)$.*
(2) $\mathrm{Span}(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k) \subset \mathrm{Span}(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k, \boldsymbol{v})$
(3) *Se $\boldsymbol{v} \in \mathrm{Span}(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k)$ allora*

$$\mathrm{Span}(\boldsymbol{u}_1 \ldots, \boldsymbol{u}_k) = \mathrm{Span}(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k, \boldsymbol{v}).$$

Osservazione 2.19. La proprietà (2) del lemma ha il seguente significato: se ad una lista di vettori $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k$ aggiungo un vettore $\boldsymbol{v}$, lo Span della nuova lista **contiene** lo Span della vecchia lista.

La proprietà (3) precisa che quando il vettore $\boldsymbol{v}$ è contenuto esso stesso nello Span della vecchia lista, allora lo Span della nuova lista **coincide** con lo Span della vecchia lista.

Dimostrazione. Proviamo una proprietà per volta:

(1) Poniamo $W = \mathrm{Span}(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k)$. Per ipotesi abbiamo $\boldsymbol{v} \in W$ e dunque per il Lemma 2.7, $\mathrm{Span}(\boldsymbol{v})$ è contenuto in W.

(2) Poniamo $W = \mathrm{Span}(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k, \boldsymbol{v})$. Osserviamo che ogni vettore $\boldsymbol{u}_i$ appartiene a W e dunque sempre per il Lemma 2.7 $\mathrm{Span}(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k) \subset W$.

(3) Per il precedente punto si ha che $\mathrm{Span}(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k)$ è contenuto in $\mathrm{Span}(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k, \boldsymbol{v})$. Dunque per mostare che i due insiemi coincidono bisogna verificare che tutti gli elementi di $\mathrm{Span}(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k, \boldsymbol{v})$ siano contenuti in $\mathrm{Span}(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k)$. Osserviamo che un vettore $\boldsymbol{w}$ in $\mathrm{Span}(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k, \boldsymbol{v})$ è della forma

$$\boldsymbol{w} = \lambda_1 \boldsymbol{u}_1 + \ldots + \lambda_k \boldsymbol{u}_k + \alpha \boldsymbol{v} .$$

Ma, poiché assumiamo che $\boldsymbol{v} \in \mathrm{Span}(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k)$

$$\boldsymbol{v} = \mu_1 \boldsymbol{u}_1 + \ldots + \mu_k \boldsymbol{u}_k + \mu \boldsymbol{v} ,$$

e, quindi, sostituendo e riordinando

$$\boldsymbol{w} = (\lambda_1 + \alpha\mu_1)\boldsymbol{u}_1 + \ldots + (\lambda_k + \alpha\mu_k)\boldsymbol{u}_k ,$$

dunque $\boldsymbol{w} \in \mathrm{Span}(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k)$, perché ottenuto come combinazione lineare dei vettori $\boldsymbol{u}_1 \ldots \boldsymbol{u}_k$ secondo i coefficienti $(\lambda_1 + \alpha\mu_1) \ldots (\lambda_k + \alpha\mu_k)$.

□

4.1. Generatori e spazi finitamente generati.

Torniamo ora al problema da cui si era partiti: cercare una buona generalizzazione del concetto di base e di coordinate per un qualsiasi spazio vettoriale.

Sia $\{\boldsymbol{u}_0, \boldsymbol{v}_0, \boldsymbol{w}_0\}$ una base di $\mathbb{E}^3_O$. Un punto fondamentale per poter definire le coordinate è che ogni vettore $\boldsymbol{v}$ di $\mathbb{E}^3_O$ si scrive come combinazione lineare

$$\boldsymbol{v} = \lambda \boldsymbol{u}_0 + \mu \boldsymbol{v}_0 + \nu \boldsymbol{w}_0 .$$

Ovvero con il linguaggio introdotto nella precedente sezione

$$\mathrm{Span}\,\{\boldsymbol{u}_0, \boldsymbol{v}_0, \boldsymbol{w}_0\} = \mathbb{E}^3_O .$$

Tutto ciò suggerisce di dare una definizione generale.

Definizione 2.10. Sia V uno spazio vettoriale. Un insieme di vettori $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k\}$ **genera** V se $\mathrm{Span}(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k) = V$. L'insieme $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k\}$ è detto un **insieme di generatori per** V.

Osservazione 2.20. Possiamo dire che

(1) I vettori $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k$ generano V se e solo se ogni vettore $\boldsymbol{v} \in V$ è ottenibile come combinazione lineare di $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k$, ovvero se e solo se per ogni $\boldsymbol{v} \in V$ esistono $\lambda_1, \ldots, \lambda_k$ tali che

$$\boldsymbol{v} = \lambda_1 \boldsymbol{u}_1 + \lambda_2 \boldsymbol{u}_2 + \ldots + \lambda_k \boldsymbol{u}_k .$$

(2) Per la proprietà (1) del Lemma 2.8, segue che se $L = \{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k\}$ è una lista di generatori di V, allora ogni altra lista ottenuta aggiungendo ad L altri vettori continua ad essere una lista di generatori di V.

Esempio 2.21. Vediamo alcuni casi.

(1) Una base di $\mathbb{E}^3_O$ è un insieme di generatori per $\mathbb{E}^3_O$.

(2) I vettori $\boldsymbol{e}_1 = \begin{pmatrix}1\\0\\0\end{pmatrix}, \boldsymbol{e}_2 = \begin{pmatrix}0\\1\\0\end{pmatrix}, \boldsymbol{e}_3 = \begin{pmatrix}0\\0\\1\end{pmatrix}$ formano un sistema di generatori di $\mathbb{R}^3$. Per verificare ciò ci dobbiamo chiedere se ogni vettore di $\mathbb{R}^3$ si può scrivere come combinazione lineare di $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$. Proviamo a fare la verifica su un esempio numerico. Sia $\boldsymbol{v} = \begin{pmatrix}7\\23\\-14\end{pmatrix}$. Ci poniamo il problema di scrivere $\boldsymbol{v}$ come combinazione lineare di $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$, ovvero cerchiamo tre numeri $\lambda_1, \lambda_2, \lambda_3$ tali che

$$\begin{pmatrix}7\\23\\-14\end{pmatrix} = \lambda_1 \begin{pmatrix}1\\0\\0\end{pmatrix} + \lambda_2 \begin{pmatrix}0\\1\\0\end{pmatrix} + \lambda_3 \begin{pmatrix}0\\0\\1\end{pmatrix}$$

Notiamo che l'espressione a destra è uguale al vettore $\begin{pmatrix}\lambda_1\\\lambda_2\\\lambda_3\end{pmatrix}$ e dunque scegliendo $\lambda_1 = 7$,$\lambda_2 = 23$, $\lambda_3 = -14$ abbiamo l'uguaglianza cercata.

In generale dato un vettore $\boldsymbol{v} = \begin{pmatrix}v_1\\v_2\\v_3\end{pmatrix}$, cerchiamo di scriverlo come combinazione lineare di $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$, ovvero cerchiamo $\lambda_1, \lambda_2, \lambda_3$ tali che

$$\boldsymbol{v} = \lambda_1\boldsymbol{e}_1 + \lambda_2\boldsymbol{e}_2 + \lambda_3\boldsymbol{e}_3$$

Come prima l'espressione a destra è uguale al vettore $\begin{pmatrix}\lambda_1\\\lambda_2\\\lambda_3\end{pmatrix}$, e dunque per ottenere l'uguaglianza cercata basta porre $\lambda_1 =$ alla prima componente di $\boldsymbol{v}$ (ovvero v_1), $\lambda_2 =$ seconda componente di $\boldsymbol{v}$ (ovvero v_2), $\lambda =$ terza componente di $\boldsymbol{v}$ (ovvero v_3).

(3) In un qualsiasi $\mathbb{R}^n$ i seguenti n vettori

$$\boldsymbol{e}_1 = \begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix}, \boldsymbol{e}_2 = \begin{pmatrix}0\\1\\0\\\vdots\\0\end{pmatrix}, \ldots, \boldsymbol{e}_n = \begin{pmatrix}0\\\vdots\\0\\1\end{pmatrix}$$

formano un sistema di generatori. Infatti una qualsiasi combinazione lineare di $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$

$$\lambda_1\boldsymbol{e}_1 + \ldots + \lambda_n\boldsymbol{e}_n$$

è uguale al vettore la cui prima componente è λ_1, la cui seconda componente è λ_2 e così via.

Dunque, per scrivere un dato un vettore $\boldsymbol{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$ come combinazione lineare

$$\boldsymbol{v} = \lambda_1 \boldsymbol{e}_1 + \ldots + \lambda_n \boldsymbol{e}_n$$

è sufficiente fissare $\lambda_1 = v_1$, $\lambda_2 = v_2$ e così via.

(4) Consideriamo in $\mathbb{R}^2$ i due seguenti vettori

$$\boldsymbol{u}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \boldsymbol{u}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

Verifichiamo se $\boldsymbol{u}_1$ e $\boldsymbol{u}_2$ formano un sistema di generatori di $\mathbb{R}^2$.

Dato un vettore $\boldsymbol{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$, ci poniamo il problema di scriverlo come combinazione lineare di $\boldsymbol{u}_1, \boldsymbol{u}_2$:

$$\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \lambda_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \lambda_2 \begin{pmatrix} 1 \\ -1 \end{pmatrix}. \tag{2.12}$$

Ora svolgendo l'espressione a destra troviamo il vettore

$$\begin{pmatrix} \lambda_1 + \lambda_2 \\ \lambda_1 - \lambda_2 \end{pmatrix}$$

e dunque imponendo l'uguaglianza otteniamo che i numeri λ_1, λ_2 devono soddisfare le seguenti condizioni

$$\begin{cases} \lambda_1 + \lambda_2 = v_1 \\ \lambda_1 - \lambda_2 = v_2 \end{cases}$$

Sommando le due equazioni troviamo che $2\lambda_1 = v_1 + v_2$ e dunque $\lambda_1 = \frac{v_1+v_2}{2}$.

Sottraendo la seconda equazione dalla prima troviamo $2\lambda_2 = v_1 - v_2$ e dunque $\lambda_2 = \frac{v_1-v_2}{2}$.

Dunque comunque siano fissato $\boldsymbol{v} \in \mathbb{R}^2$, riusciamo a trovare λ_1, λ_2 in modo che (2.12) sia verificata. Segue che $\boldsymbol{u}_1$, $\boldsymbol{u}_2$ formano un sistema di generatori di V.

(5) Consideriamo in $\mathbb{R}^3$ i seguenti vettori

$$\boldsymbol{u}_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \boldsymbol{u}_2 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \boldsymbol{u}_3 = \begin{pmatrix} 5 \\ 1 \\ 5 \end{pmatrix}$$

Verifichiamo se tali vettori formano un sistema di generatori di $\mathbb{R}^3$. Dato un vettore $\boldsymbol{v}$ ci poniamo il problema di scriverlo come combinazione lineare di $\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3$

$$v = \lambda_1 \boldsymbol{u}_1 + \lambda_2 \boldsymbol{u}_2 + \lambda_3 \boldsymbol{u}_3\,. \tag{2.13}$$

Svolgendo l'espressione a destra troviamo

$$\lambda_1 \boldsymbol{u}_1 + \lambda_2 \boldsymbol{u}_2 + \lambda_3 \boldsymbol{u}_3 = \begin{pmatrix} \lambda_1 + \lambda_2 + 5\lambda_3 \\ \lambda_2 + \lambda_3 \\ \lambda_1 + \lambda_2 + 5\lambda_3 \end{pmatrix}$$

E dunque imponendo l'uguaglianza (2.13) componente per componente, risulta che i numeri $\lambda_1, \lambda_2, \lambda_3$ devono soddisfare tutte e tre le seguenti condizioni

$$\begin{cases} \lambda_1 + \lambda_2 + 5\lambda_3 = v_1 \\ \lambda_2 + \lambda_3 = v_2 \\ \lambda_1 + \lambda_2 + 5\lambda_3 = v_3 \end{cases}$$

Osserviamo che se $v_1 \neq v_3$, non sarà mai possibile trovare $\lambda_1, \lambda_2, \lambda_3$ che verifichino sia la prima che la terza equazione, e dunque i vettori la cui prima e la terza componente sono diverse, non potranno essere espressi come combinazione lineare di $\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3$. Segue che $\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3$ *non* formano un insieme di generatori.

Esercizio 4.2. Si propongono i seguenti esercizi.

(1) Si discuta se $\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 5 \\ 7 \end{pmatrix}$ formano un sistema di generatori di $\mathbb{R}^2$

(2) Si discuta se $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ formano un sistema di generatori di $\mathbb{R}^3$

(3) Si discuta se $\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$ formano un sistema di generatori di $\mathbb{R}^3$

(4) Si mostri che $V = \left\{ \boldsymbol{v} \in \mathbb{R}^3 | v_1 + v_2 = 0 \right\}$ 'e un sottospazio di $\mathbb{R}^3$. Si verifichi che i vettori $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$ appartengono a V e si discuta se essi formano un sistema di generatori per V.

Definizione 2.11. Uno spazio vettoriale si dice **finitamente generato** se esiste una lista finita di vettori di V, $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n\}$ che genera V.

Tutti gli spazi $\mathbb{R}^n$ sono finitamente generati. Infatti nell'esempio 2.21(3) abbiamo costruito una lista finita di generatori per ciascun $\mathbb{R}^n$.

Ma esistono anche spazi che *non* sono finitamente generati? Vediamo esplicitamente un esempio che ci convinca della risposta affermativa.

Esempio 2.22 (di spazio NON finitamente generato). Mostriamo che lo spazio dei polinomi $\mathbb{R}[x]$ non è finitamente generato. Per mostrare ciò partiamo da un esempio:

Consideriamo la seguente lista di polinomi

$$p_1(x) = x^5 + 3,\ p_2(x) = x^7 + x^4 - 3,\ p_3(x) = x^6 - x^5 + 3x^4 - x^2\,.$$

Osserviamo che facendo combinazioni lineari dei polinomi p_1, p_2, p_3 possiamo ottenere solo polinomi di grado ≤ 7, mentre non si potranno mai ottenere polinomi di grado > 7. Segue che il polinomio $x^8 \notin \text{Span}(p_1, p_2, p_3)$ e dunque p_1, p_2, p_3 non formano un sistema di generatori.

Tale argomento si applica a qualunque lista finita di polinomi. Data una lista $p_1, \ldots, p_n$ si considera il grado massimo, d, di $p_1, \ldots, p_n$, e si osserva che le combinazioni lineari di $p_1, \ldots, p_n$ hanno tutte grado $\leq d$. Dunque il polinomio $x^{d+1} \notin \text{Span}(p_1, \ldots, p_n)$.

Segue che nessuna lista finita può essere un sistema di generatori per $\mathbb{R}[x]$.

5. Vettori linearmente indipendenti

Torniamo ora al problema da cui eravamo partiti, ovvero trovare una generalizzazione della nozione di basi e coordinate per spazi vettoriali astratti.

Nella precedente sezione abbiamo sottolineato che le basi di $\mathbb{E}^3_O$ sono sistemi di generatori. Ovvero se $\boldsymbol{u}_0, \boldsymbol{v}_0, \boldsymbol{w}_0$ è base di $\mathbb{E}^3_O$, ogni vettore $\boldsymbol{v}$ si può scrivere nella forma

$$\boldsymbol{v} = \lambda \boldsymbol{u}_0 + \mu \boldsymbol{v}_0 + \nu \boldsymbol{w}_0$$

e le coordinate del vettore $\boldsymbol{v}$ rispetto alla base $\boldsymbol{u}_0, \boldsymbol{v}_0, \boldsymbol{w}_0$ sono proprio i tre coefficienti λ, μ, ν

Mostriamo però che la proprietà di generare non è sufficiente per definire le coordinate.

Per semplicità illustreremo un esempio sul piano $\mathbb{E}^2_O$.

Dati due vettori $\boldsymbol{u}_1, \boldsymbol{u}_2 \in \mathbb{E}^2_O$ che non hanno la stessa direzione ogni vettore $\boldsymbol{v} \in \mathbb{E}^2_O$ può essere scritto come combinazione lineare di $\boldsymbol{u}_1, \boldsymbol{u}_2$. Ovvero esistono due numeri λ, μ tali che

$$\boldsymbol{v} = \lambda \boldsymbol{u}_1 + \mu \boldsymbol{u}_2 \,.$$

In questo modo abbiamo associato al vettore geometrico $\boldsymbol{v}$ la coppia di numeri (λ, μ).

In effetti tale procedura funziona purché ci sia un **unico** modo di scrivere $\boldsymbol{v}$ come combinazione lineare di $\boldsymbol{u}_1, \boldsymbol{u}_2$. Infatti se esistesse un'altra coppia di numeri (λ', μ') tali che

$$\boldsymbol{v} = \lambda' \boldsymbol{u}_1 + \mu' \boldsymbol{u}_2$$

non si saprebbe bene se le coordinate di $\boldsymbol{v}$ sono (λ, μ) oppure (λ', μ').

Per fortuna ciò non accade: infatti dovremmo avere:

$$\lambda \boldsymbol{u}_1 + \mu \boldsymbol{u}_2 = \lambda' \boldsymbol{u}_1 + \mu' \boldsymbol{u}_2$$

da cui otteremmo

$$(\lambda - \lambda')\boldsymbol{u}_1 + (\mu - \mu')\boldsymbol{u}_2 = 0$$

da cui, se $\lambda \neq \lambda'$ otterremmo

$$\boldsymbol{u}_1 = -\frac{\mu - \mu'}{\lambda - \lambda'}\boldsymbol{u}_2$$

ovvero $\boldsymbol{u}_1$ e $\boldsymbol{u}_2$ sarebbero paralleli contro la nostra ipotesi.

Se ai vettori $\boldsymbol{u}_1$, $\boldsymbol{u}_2$ aggiungo un terzo vettore $\boldsymbol{u}_3 \in \mathbb{E}^2_O$ l'insieme dei vettori $\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3$ ancora genera $\mathbb{E}^2_O$. Però dato un vettore $\boldsymbol{v}$ si può mostrare che esistono infinite combinazioni lineari tali che

$$\boldsymbol{v} = \lambda_1 \boldsymbol{u}_1 + \lambda_2 \boldsymbol{u}_2 + \lambda_3 \boldsymbol{u}_3 \tag{2.14}$$

Più precisamente si può fissare arbitrariamente il valore di λ_3 e trovare λ_1 e λ_2 che verifichino la condizione (2.14).

Se ad esempio poniamo $\lambda_3 = 100$ posso considerare il vettore $\boldsymbol{v} - 100\boldsymbol{u}_3$. Poiché $\boldsymbol{u}_1$ e $\boldsymbol{u}_2$ generano trovo dei numeri λ_1, λ_2 tali che

$$\boldsymbol{v} - 100\boldsymbol{u}_3 = \lambda_1 \boldsymbol{u}_1 + \lambda_2 \boldsymbol{u}_2$$

e dunque

$$\boldsymbol{v} = \lambda_1 \boldsymbol{u}_1 + \lambda_2 \boldsymbol{u}_2 + 100\boldsymbol{u}_3$$

Dunque anche se $\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3$ sono dei generatori, **essi non mi permettono di determinare in modo univoco delle coordinate** per $\boldsymbol{v}$. La ragione di questa indeterminatezza è che $\boldsymbol{u}_3$ è superfluo, in quanto $\boldsymbol{u}_1$ e $\boldsymbol{u}_2$ già generano tutto.

Dunque per ottenere una generalizzazione della nozione di base e coordinate abbiamo allora bisogno di un'altra proprietà che esprime la condizione che combinazioni lineari diverse producano vettori diversi

Tale proprietà va sotto il nome di lineare indipendenza:

Definizione 2.12. Sia $\{\boldsymbol{u}_1, \dots, \boldsymbol{u}_n\}$ una lista di vettori di V. I vettori $\boldsymbol{u}_1, \dots, \boldsymbol{u}_k$ si dicono **linearmente indipendenti** se sono tutti non-nulli e ciascun elemento $\boldsymbol{u}_i$ non è combinazione lineare degli altri elementi della lista.

Viceversa i vettori $\{\boldsymbol{u}_1, \dots \boldsymbol{u}_k\}$ si dicono **linearmente dipendenti** se è possibile scrivere un vettore di tale lista come combinazione lineare degli altri.

Esempio 2.23. Cerchiamo di chiarici le idee con alcuni esempi espliciti.

(1) Una base $\{\boldsymbol{u}_0, \boldsymbol{v}_0, \boldsymbol{w}_0\}$ di $\mathbb{E}^3_O$ è composta da vettori linearmente indipendenti. Infatti se ad esempio $\boldsymbol{w}_0$ fosse combinazione lineare di $\boldsymbol{u}_0$ e $\boldsymbol{v}_0$, allora necessariamente i tre vettori dovrebbero essere complanari, mentre per definizione le basi di $\mathbb{E}^3_O$ sono composte da vettori non complanari.

(2) I vettori $\boldsymbol{u}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \boldsymbol{u}_2 = \begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix}$ sono linearmente dipendenti in quanto $\boldsymbol{u}_2 = 2\boldsymbol{u}_1$.

(3) I vettori $\boldsymbol{u}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \boldsymbol{u}_2 = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}$ sono indipendenti.

Infatti $\boldsymbol{u}_2$ non è multiplo di $\boldsymbol{u}_1$ né $\boldsymbol{u}_1$ è multiplo di $\boldsymbol{u}_2$.

(4) I vettori $\boldsymbol{u}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \boldsymbol{u}_2 = \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix}, \boldsymbol{u}_3 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$ sono linearmente dipendenti, in quanto $\boldsymbol{u}_2 = 2\boldsymbol{u}_1$.

(5) I vettori $\boldsymbol{u}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \boldsymbol{u}_2 = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \boldsymbol{u}_3 = \begin{pmatrix} 2 \\ 3 \\ 2 \end{pmatrix}$ sono dipendenti in quanto $\boldsymbol{u}_3 = \boldsymbol{u}_1 + \boldsymbol{u}_2$.

(6) Verifichiamo che i vettori

$$\boldsymbol{u}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \boldsymbol{u}_2 = \begin{pmatrix} 1 \\ 1 \\ -1 \\ 0 \end{pmatrix}, \boldsymbol{u}_3 = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix}, \boldsymbol{u}_4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

sono indipendenti.

Infatti se proviamo a scrivere $\boldsymbol{u}_4$ come combinazione lineare di $\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3$

$$\boldsymbol{u}_4 = \lambda \boldsymbol{u}_1 + \mu \boldsymbol{u}_2 + \nu \boldsymbol{u}_3$$

otteniamo che λ, μ, ν sono soluzioni del sistema

$$\begin{cases} \lambda + \mu + \nu = 0 \\ \lambda + \mu - \nu = 0 \\ \lambda - \mu = 0 \\ \lambda = 1 \end{cases}$$

Osserviamo che gli unici valori di λ, μ, ν che verificano contemporaneamente le ultime tre equazioni sono $\lambda = \mu = 1, \nu = 2$. Poiché questi valori non soddisfano la prima equazione il sistema non ha soluzioni e dunque $\boldsymbol{u}_4$ non è combinazione lineare di $\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3$. Ora proviamo a scrivere $\boldsymbol{u}_3$ come combinazione lineare di $\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_4$:

$$\boldsymbol{u}_3 = \lambda \boldsymbol{u}_1 + \mu \boldsymbol{u}_2 + \nu \boldsymbol{u}_4 \,.$$

Osserviamo che necessariamente $\nu = 0$ altrimenti ricaverei

$$\boldsymbol{u}_4 = \frac{1}{\nu}(\boldsymbol{u}_3 - \lambda \boldsymbol{u}_1 - \mu \boldsymbol{u}_2) \,.$$

Dunque ponendo $\boldsymbol{u}_3 = \lambda \boldsymbol{u}_1 + \mu \boldsymbol{u}_2$, si ricava che λ e μ devono soddisfare il sistema

$$\begin{cases} \lambda + \mu = 1 \\ \lambda + \mu = -1 \\ \lambda - \mu = 0 \end{cases}$$

che ovviamente non ha soluzioni, e dunque $\boldsymbol{u}_3$ non è combinazione di $\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_4$.

Proviamo che $\boldsymbol{u}_2$ non è combinazione lineare di $\boldsymbol{u}_1, \boldsymbol{u}_3, \boldsymbol{u}_4$. Se poniamo $\boldsymbol{u}_2 = \lambda \boldsymbol{u}_1 + \mu \boldsymbol{u}_3 + \nu \boldsymbol{u}_4$, come prima si ha che $\nu = 0$ altrimenti si ricaverebbe $\boldsymbol{u}_4$ come combinazione di $\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3$. Analogamente $\mu = 0$ altrimenti potremmo ricavare $\boldsymbol{u}_3$ come combinazione di $\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_4$. Dunque l'unica possibilità è che $\boldsymbol{u}_2 = \lambda \boldsymbol{u}_1$, ma questo si esclude facilmente. Analogamente si prova che $\boldsymbol{u}_1$ non è combinazione lineare di $\boldsymbol{u}_2, \boldsymbol{u}_3, \boldsymbol{u}_4$.

Osservazione 2.24. Dall'esempio (1) risulta che due vettori sono dipendenti se sono uno multiplo dell'altro.

Da tre vettori in su, non c'è una caratterizzazione così semplice. Infatti l'esempio (4) mostra un caso in cui i tre vettori sono dipendenti ma nessuno dei tre è multiplo di un altro.

Ora gli insiemi linearmente indipendenti soddisfano la proprietà che cerchiamo. Ovvero se $\{\boldsymbol{u}_1, \dots, \boldsymbol{u}_k\}$ è un insieme di vettori linearmente indipendenti posso dedurre che combinazioni lineari diverse producono vettori diversi.

Introduciamo la seguente terminologia.

Definizione 2.13. Data una combinazione lineare

$$\lambda_1 \boldsymbol{u}_1 + \ldots + \lambda_k \boldsymbol{u}_k$$

diciamo che **vettore dei coefficienti** [1] della combinazione è la k-upla formata dai coefficienti della combinazione

$$\begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_k \end{pmatrix}.$$

Proposizione 2.9. *Sia $\{\boldsymbol{u}_1, \dots, \boldsymbol{u}_k\}$ un insieme di vettori linearmente indipendenti. Dati due vettori di coefficienti $\begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_k \end{pmatrix} \neq \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_k \end{pmatrix}$ si ha che*

$$\lambda_1 \boldsymbol{u}_1 + \ldots + \lambda_k \boldsymbol{u}_k \neq \mu_1 \boldsymbol{u}_1 + \ldots + \mu_k \boldsymbol{u}_k$$

Dimostrazione. Partiamo da un esempio: supponiamo di avere 4 vettori

$$\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3, \boldsymbol{u}_4$$

e supponiamo di avere due combinazioni lineari diverse che danno lo stesso vettore:

$$\boldsymbol{u}_1 + 3\boldsymbol{u}_2 - 7\boldsymbol{u}_3 + 4\boldsymbol{u}_4 = -10\boldsymbol{u}_1 + 2\boldsymbol{u}_2 + 3\boldsymbol{u}_3 + 6\boldsymbol{u}_4$$

Cerchiamo di verificare che in tale ipotesi $\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3, \boldsymbol{u}_4$ sono per forza dipendenti.

Raccogliendo tutti i multipli di $\boldsymbol{u}_4$ ad un membro otteniamo

$$4\boldsymbol{u}_4 - 6\boldsymbol{u}_4 = -10\boldsymbol{u}_1 + 2\boldsymbol{u}_2 + 3\boldsymbol{u}_3 - \boldsymbol{u}_1 - 3\boldsymbol{u}_2 + 7\boldsymbol{u}_3$$

ovvero riordinando i termini

$$-2\boldsymbol{u}_4 = -11\boldsymbol{u}_1 - \boldsymbol{u}_2 + 10\boldsymbol{u}_3$$

da cui deduciamo che $\boldsymbol{u}_4$ è combinazione lineare di $\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3$:

$$\boldsymbol{u}_4 = \frac{11}{2}\boldsymbol{u}_1 + \frac{1}{2}\boldsymbol{u}_2 - 5\boldsymbol{u}_3 .$$

Osserviamo che non avremmo potuto dedurre la stessa cosa se i coefficienti che moltiplicano $\boldsymbol{u}_4$ fossero stati uguali in entrambi i membri. Se ad esempio avessimo avuto

$$\boldsymbol{u}_1 + 3\boldsymbol{u}_2 - 7\boldsymbol{u}_3 + 4\boldsymbol{u}_4 = -10\boldsymbol{u}_1 + 2\boldsymbol{u}_2 + 3\boldsymbol{u}_3 + 4\boldsymbol{u}_4$$

gli addendi multipli di $\boldsymbol{u}_4$ si semplificano, lasciando un'espressione da cui non è possibile ricavare $\boldsymbol{u}_4$.

[1] ATTENZIONE: il vettore dei coefficienti **NON** appartiene a V ma ad $\mathbb{R}^k$ dove k è il numero di vettori che consideriamo.

In questo caso però avremmo potuto portare gli addendi multipli di $\boldsymbol{u}_3$ (o anche di $\boldsymbol{u}_2$ o di $\boldsymbol{u}_1$) tutti da una parte e avremmo ricavato

$$-11\boldsymbol{u}_3 = -11\boldsymbol{u}_1 - \boldsymbol{u}_2$$

e dunque si otterrebbe $\boldsymbol{u}_3$ come combinazione lineare di $\boldsymbol{u}_1$ e $\boldsymbol{u}_2$.

Consideriamo ora il caso generale.

Supponiamo che esistano due combinazioni lineari distinte di $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k$ che producono lo stesso risultato:

$$\lambda_1\boldsymbol{u}_1 + \ldots + \lambda_k\boldsymbol{u}_k = \mu_1\boldsymbol{u}_1 + \ldots + \mu_k\boldsymbol{u}_k\,. \tag{2.15}$$

Se $\lambda_k \neq \mu_k$, allora isolando a sinistra i multipli di $\boldsymbol{u}_n$ otteniamo

$$(\lambda_k - \mu_k)\boldsymbol{u}_k = \mu_1\boldsymbol{u}_1 + \ldots + \mu_{k-1}\boldsymbol{u}_{k-1} - \lambda_1\boldsymbol{u}_1 - \ldots - \lambda_{n-1}\boldsymbol{u}_{n-1}$$

e riordinando

$$(\lambda_k - \mu_k)\boldsymbol{u}_k = (\mu_1 - \lambda_1)\boldsymbol{u}_1 + \ldots + (\mu_{k-1} - \lambda_{k-1})\boldsymbol{u}_{k-1}\,.$$

Infine dividendo per $\lambda_k - \mu_k$ (lo posso fare perché sto supponendo $\lambda_k \neq \mu_k$) riesco ad esprimere $\boldsymbol{u}_k$ come combinazione lineare dei precedenti.

Nel caso in cui $\lambda_k = \mu_k$, si riesce comunque a trovare un indice i per cui si ha $\lambda_i \neq \mu_i$ (altrimenti le due combinazioni sarebbero uguali).

A questo punto osservo che posso riordinare gli addendi nella (2.15) e ottenere

$$\begin{aligned}&\left(\lambda_1\boldsymbol{u}_1 + \lambda_2\boldsymbol{u}_2 + \ldots + \lambda_{i-1}\boldsymbol{u}_{i-1} + \lambda_{i+1}\boldsymbol{u}_{i+1} + \ldots + \lambda_k\boldsymbol{u}_k\right) \ + \lambda_i\boldsymbol{u}_i = \\ &\left(\mu_1\boldsymbol{u}_1 + \mu_2\boldsymbol{u}_2 + \ldots + \mu_{i-1}\boldsymbol{u}_{i-1} + \mu_{i+1}\boldsymbol{u}_{i+1} + \ldots + \mu_k\boldsymbol{u}_k\right) + \mu_i\boldsymbol{u}_i\,.\end{aligned}$$

Isolando a sinistra gli addendi multipli di $\boldsymbol{u}_i$ abbiamo

$$\begin{aligned}&\lambda_i\boldsymbol{u}_i - \mu_i\boldsymbol{u}_i = \\ &\left(\mu_1\boldsymbol{u}_1 + \ldots + \mu_{i-1}\boldsymbol{u}_{i-1} + \mu_{i+1}\boldsymbol{u}_{i+1} + \ldots + \mu_k\boldsymbol{u}_k\right) + \\ &-\left(\lambda_1\boldsymbol{u}_1 + \ldots + \lambda_{i-1}\boldsymbol{u}_{i-1} + \lambda_{i+1}\boldsymbol{u}_{i+1} + \ldots + \lambda_k\boldsymbol{u}_k\right)\end{aligned}$$

e ripeto il ragionamento di sopra. □

OSSERVAZIONE 2.25. Notiamo che l'enunciato della proposizione non vale per liste di vettori arbitrarie. Ad esempio se considero

$$\boldsymbol{u}_1 = \begin{pmatrix}1\\0\end{pmatrix}, \boldsymbol{u}_2 = \begin{pmatrix}0\\1\end{pmatrix}, \boldsymbol{u}_3 = \begin{pmatrix}1\\1\end{pmatrix}$$

è facile constatare che

$$\boldsymbol{u}_1 + \boldsymbol{u}_2 + 2\boldsymbol{u}_3 = 2\boldsymbol{u}_1 + 2\boldsymbol{u}_2 + \boldsymbol{u}_3 = 3\boldsymbol{u}_1 + 3\boldsymbol{u}_2\,.$$

DEFINIZIONE 2.14. Diciamo che una combinazione lineare

$$\lambda_1\boldsymbol{u}_1 + \ldots + \lambda_k\boldsymbol{u}_k$$

è **banale**, se tutti i suoi coefficienti sono nulli, o, equivalentemente, se il vettore dei coefficienti è il vettore nullo di $\mathbb{R}^k$.

Chiaramente la combinazione lineare banale ha come risultato il vettore nullo $\mathbf{0}_V$.

Dalla proposizione precedente ricaviamo che le liste di vettori indipendenti $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k\}$ godono della seguente proprietà:

PROPRIETÀ 2.10. *Se $\{\boldsymbol{u}_1, \dots, \boldsymbol{u}_k\}$ è una lista di vettori **indipendenti** l'unica combinazione lineare di $\boldsymbol{u}_1, \dots, \boldsymbol{u}_k$ che produce il vettore nullo è la combinazione banale.*

Osserviamo che **solo le liste di vettori indipendenti godono di questa proprietà**. Infatti data una lista di vettori dipendenti, è sempre possibile trovare una combinazione lineare non banale che produce il vettore nullo.

A titolo esemplificativo, supponiamo di avere 4 vettori $\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3, \boldsymbol{u}_4$ dipendenti, e supponiamo ad esempio che $\boldsymbol{u}_4$ sia ottenibile come combinazione lineare di $\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3$

$$\boldsymbol{u}_4 = 3\boldsymbol{u}_1 + 2\boldsymbol{u}_2 - 7\boldsymbol{u}_3 \tag{2.16}$$

Allora portando a destra $\boldsymbol{u}_4$ otteniamo una combinazione lineare non banale il cui risultato dà il vettore nullo

$$3\boldsymbol{u}_1 + 2\boldsymbol{u}_2 - 7\boldsymbol{u}_3 - \boldsymbol{u}_4 = \mathbf{0}\,.$$

Chiaramente si può ripetere lo stesso ragionamento ogni volta che in una lista un vettore può essere scritto come combinazione lineare degli altri.

PROPOSIZIONE 2.11. *Se i vettori $\{\boldsymbol{u}_1, \dots, \boldsymbol{u}_k\}$ sono linearmente dipendenti allora esiste una combinazione lineare non banale (in cui cioè almeno uno dei coefficienti è non nullo) il cui risultato è il vettore nullo.* □

ESEMPIO 2.26. Esaminiamo alcuni esempi.

(1) Verifichiamo che i seguenti vettori di $\mathbb{R}^4$

$$\boldsymbol{e}_1 = \begin{pmatrix}1\\0\\0\\0\end{pmatrix},\ \boldsymbol{e}_2 = \begin{pmatrix}0\\1\\0\\0\end{pmatrix},\ \boldsymbol{e}_3 = \begin{pmatrix}0\\0\\1\\0\end{pmatrix},\ \boldsymbol{e}_4 = \begin{pmatrix}0\\0\\0\\1\end{pmatrix}$$

sono linearmente indipendenti. Osserviamo che il risultato di una combinazione lineare di $\boldsymbol{e}_1, \dots, \boldsymbol{e}_4$ è facilmente calcolabile:

$$\lambda\boldsymbol{e}_1 + \mu\boldsymbol{e}_2 + \nu\boldsymbol{e}_3 + \eta\boldsymbol{e}_4 = \begin{pmatrix}\lambda\\\mu\\\nu\\\eta\end{pmatrix}$$

Dunque se la combinazione lineare non è quella banale (ovvero almeno uno tra i numeri λ, μ, ν, η è diverso da 0), il risultato non è il vettore nullo.

Per la Proposizione 2.11 i vettori $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3, \boldsymbol{e}_4$ non possono essere dipendenti, per cui sono indipendenti.

(2) Si può generalizzare l'argomento dato sopra e mostrare che i vettori di $\mathbb{R}^n$

$$e_1 = \begin{pmatrix}1\\0\\0\\\vdots\\0\end{pmatrix},\ e_2 = \begin{pmatrix}0\\1\\0\\\vdots\\0\end{pmatrix},\ \dots,\ e_n = \begin{pmatrix}0\\0\\0\\\vdots\\1\end{pmatrix}$$

siano indipendenti. Infatti se è possibile calcolare facilmente una combianzione lineare di tali vettori

$$\lambda_1 \boldsymbol{e}_1 + \lambda_2 \boldsymbol{e}_2 + \ldots + \lambda_n \boldsymbol{e}_n = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}.$$

Dunque è evidente che l'unica combinazione lineare che dà come risultato il vettore nullo è la combinazione banale.

Abbiamo visto nella precedente sezione che se ad una lista di generatori aggiungiamo altri vettori, continuiamo ad avere liste di generatori. Tale proprietà chiaramente non vale per le liste di vettori indipendenti. Infatti se alla lista di vettori indipendenti $\left\{ \boldsymbol{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \boldsymbol{e}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$ aggiungo il vettore $\boldsymbol{u} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, ottengo una lista di vettori linearmente dipendenti, in quanto $\boldsymbol{u} = \boldsymbol{e}_1 + \boldsymbol{e}_2$.

In effetti per gli insiemi di vettori linearmente indipendenti vale la proprietà opposta: ovvero togliendo un po' di vettori ad una lista di vettori indipendenti, continuiamo ad avere insiemi di vettori indipendenti.

Osservazione 2.27. Sia $L = \{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k\}$ una lista di vettori linearmente indipendenti di uno spazio vettoriale V.

Se L' è un sottoinsieme e di L, allora anche L' è un insieme di vettori linearmente indipendenti

Infatti ogni combinazione lineare di vettori di L' è anche una combinazione lineare di vettori di L. Dunque se riuscissimo a scrivere un vettore di L' come combinazione lineare degli altri elementi di L' saremmo riusciti a scrivere tale vettore come combinazione lineare degli altri vettori di L contraddicendo la lineare indipendenza.

Per concludere questa sezione vogliamo descrivere un modo semplice per verificare se una lista di vettori è linearmente indipendente.

Proposizione 2.12. *Una lista di vettori $\{\boldsymbol{u}_1, \ldots \boldsymbol{u}_k\}$ in uno spazio vettoriale V è linearmente indipendente se*

- $\boldsymbol{u}_1 \neq \boldsymbol{0}$,
- $\boldsymbol{u}_2 \notin \mathrm{Span}(\boldsymbol{u}_1)$,
- $\boldsymbol{u}_3 \notin \mathrm{Span}(\boldsymbol{u}_1, \boldsymbol{u}_2)$,
- $\ldots$,
- $\boldsymbol{u}_k \notin \mathrm{Span}(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_{k-1})$.

Non daremo una dimostrazione generale di tale proposizione accontentandoci di verificarla nel caso $k = 5$. L'argomento che diamo però può essere utilizzato per liste contenenti un numero arbitrario di vettori.

Supponiamo di avere 5 vettori $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_5$ in modo che nessuno si possa scrivere come combinazione lineare dei precedenti (ovvero che valga l'ipotesi della Proposizione).

Ora per ipotesi $\boldsymbol{u}_5$ non si può scrivere come combinazione lineare di $\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3, \boldsymbol{u}_4$.

Supponiamo di poter scrivere $\boldsymbol{u}_4$ come combinazione lineare di $\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3, \boldsymbol{u}_5$

$$\boldsymbol{u}_4 = \lambda_1 \boldsymbol{u}_1 + \lambda_2 \boldsymbol{u}_2 + \lambda_3 \boldsymbol{u}_3 + \lambda_5 \boldsymbol{u}_5$$

allora ci sono due casi:

- $\lambda_5 = 0$: in questo caso risulta
$$\boldsymbol{u}_4 = \lambda_1 \boldsymbol{u}_1 + \lambda_2 \boldsymbol{u}_2 + \lambda_3 \boldsymbol{u}_3$$
il che contraddice l'assunzione che $\boldsymbol{u}_4$ non sia combinazione lineare dei precedenti.
- $\lambda_5 \neq 0$: in questo caso isolando il termine $\lambda_5 \boldsymbol{u}_5$ otteniamo
$$\lambda_5 \boldsymbol{u}_5 = \boldsymbol{u}_4 - \lambda_1 \boldsymbol{u}_1 - \lambda_2 \boldsymbol{u}_2 - \lambda_3 \boldsymbol{u}_3$$
e dividendo per λ_5 riusciremmo a scrivere $\boldsymbol{u}_5$ come combinazione lineare dei precedenti.

Poiché in entrambi otteniamo una contraddizione della nostra ipotesi, ricaviamo che $\boldsymbol{u}_4$ non è combinazione lineare di $\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3, \boldsymbol{u}_5$.

Supponiamo ora di poter scrivere $\boldsymbol{u}_3$ come combinazione lineare degli altri vettori
$$\boldsymbol{u}_3 = \lambda_1 \boldsymbol{u}_1 + \lambda_2 \boldsymbol{u}_2 + \lambda_4 \boldsymbol{u}_4 + \lambda_5 \boldsymbol{u}_5 \,.$$
Distinguiamo vari casi:

- $\lambda_4 = \lambda_5 = 0$: in tal caso avremmo che
$$\boldsymbol{u}_3 = \lambda_1 \boldsymbol{u}_1 + \lambda_2 \boldsymbol{u}_2$$
per cui $\boldsymbol{u}_3 \in \mathrm{Span}(\boldsymbol{u}_1, \boldsymbol{u}_2)$ che contraddice la nostra assunzione sulla lista di vettori.
- $\lambda_5 = 0$ ma $\lambda_4 \neq 0$: in tal caso avremmo
$$\boldsymbol{u}_3 = \lambda_1 \boldsymbol{u}_1 + \lambda_2 \boldsymbol{u}_2 + \lambda_4 \boldsymbol{u}_4$$
e dunque isolando il termine $\lambda_4 \boldsymbol{u}_4$ otteniamo
$$\lambda_4 \boldsymbol{u}_4 = \boldsymbol{u}_3 - \lambda_1 \boldsymbol{u}_1 - \lambda_2 \boldsymbol{u}_2$$
da cui dividendo per λ_4 ricaviamo che $\boldsymbol{u}_4$ è combinazione lineare dei precedenti.
- se $\lambda_5 \neq 0$ allora isolando ad un membro l'addendo $\lambda_5 \boldsymbol{u}_5$ otteniamo
$$\lambda_5 \boldsymbol{u}_5 = \boldsymbol{u}_3 - \lambda_1 \boldsymbol{u}_1 - \lambda_2 \boldsymbol{u}_2 - \lambda_4 \boldsymbol{u}_4$$
da cui ancora si ricava che $\boldsymbol{u}_5$ è combinazione lineare dei precedenti

Dunque in ciascun caso riusciamo a trovare una contraddizione rispetto alla nostra assunzione sulla lista.In definitiva non è possibile che $\boldsymbol{u}_3$ sia combinazione lineare degli altri elementi della lista.

Ragionando in maniera del tutto analoga si giunge a mostrare che né $\boldsymbol{u}_2$ né $\boldsymbol{u}_1$ sono combinazioni lineari dei precedenti.

Esercizio 5.1. Si dimostri l'enunciato della Proposizione per liste contenenti 6 vettori (ovvero $k = 6$).

6. Basi e coordinate

Abbiamo visto che una base $\{\boldsymbol{u}_0, \boldsymbol{v}_0, \boldsymbol{w}_0\}$ di $\mathbb{E}^3_O$ soddisfa le due proprietà:

- è un sistema di generatori;
- è formato da vettori indipendenti.

Vogliamo ora mostrare che tali proprietà caratterizzano le basi di $\mathbb{E}^3_O$.

PROPOSIZIONE 2.13. *Sia* $L = \{\boldsymbol{u}_1, \dots, \boldsymbol{u}_k\}$ *una lista di vettori di* $\mathbb{E}^3_O$ *tale che soddisfa le proprietà*

- *è un sistema di generatori;*
- *è formato da vettori indipendenti.*

Allora L *è costituito da* 3 *vettori non complanari (in particolare* $k = 3$*).*

DIMOSTRAZIONE. Poiché un vettore genera una retta e due vettori indipendenti generano un piano sicuramente L deve contenere almeno 3 vettori.

Poiché i vettori di L sono indipendenti allora $\boldsymbol{u}_1$ e $\boldsymbol{u}_2$ non sono paralleli, e $\boldsymbol{u}_3$ non è contenuto nel piano generato da $\boldsymbol{u}_1$ e $\boldsymbol{u}_2$.

Dunque i vettori $\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3$ non sono complanari ed in particolare

$$\mathrm{Span}(\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3) = \mathbb{E}^3_O.$$

Ciò implica che L non può contenere un altro vettore $\boldsymbol{u}_4$, perché altrimenti esso sarebbe una combinazione lineare di $\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3$ contraddicendo l'assunzione che i vettori di L siano indipendenti. □

La Proposizione 2.13 suggerisce che la definizione di base si può generalizzare ad un qualsiasi spazio vettoriale nel seguente modo.

DEFINIZIONE 2.15. Una **base** di V è un insieme $\mathcal{B} = \{\boldsymbol{u}_1, \dots \boldsymbol{u}_k\}$ di vettori di V che soddisfa contemporaneamente le due proprietà:

- $\mathcal{B}$ è un sistema di generatori di V;
- i vettori di $\mathcal{B}$ sono linearmente indipendenti.

ESEMPIO 2.28 ((Base canonica)). Abbiamo visto nelle precedenti sezioni che i vettori di $\mathbb{R}^n$

$$\boldsymbol{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ \boldsymbol{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots, \ \boldsymbol{e}_n = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

sono generatori di $\mathbb{R}^n$ linearmente indipendenti. Dunque essi formano una base di $\mathbb{R}^n$, che viene detta **base canonica** .

Ora si verifica facilmente che vale la seguente proposizione.

PROPOSIZIONE 2.14. *Sia* $\mathcal{B} = \{\boldsymbol{u}_1, \dots, \boldsymbol{u}_k\}$ *una base di* V*. Allora dato un vettore* $\boldsymbol{v} \in V$ *esiste un'unica* k*-upla di numeri* $\lambda_1, \dots, \lambda_k$ *per cui*

$$\boldsymbol{v} = \lambda_1 \boldsymbol{u}_1 + \dots + \lambda_k \boldsymbol{u}_k \,.$$

DEFINIZIONE 2.16. I numeri $\lambda_1, \dots, \lambda_k$ sono dette le **coordinate** di $\boldsymbol{v}$ rispetto alla base $\mathcal{B}$.

Il vettore delle coordinate verrà indicato nel seguente modo

$$[\boldsymbol{v}]_{\mathcal{B}} = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_k \end{pmatrix}.$$

Dimostrazione. Poiché $\mathcal{B}$ genera V, esiste una scelta di scalari $\lambda_1, \dots, \lambda_k$ tali che

$$\boldsymbol{v} = \lambda_1 \boldsymbol{u}_1 + \ldots + \lambda_k \boldsymbol{u}_k \,.$$

Poiché $\boldsymbol{u}_1, \dots, \boldsymbol{u}_k$ sono indipendenti, si può applicare la Proposizione 2.9 che assicura che qualsiasi combinazione lineare diversa da $\lambda_1 \boldsymbol{u}_1 + \ldots + \lambda_k \boldsymbol{u}_k$ produce un vettore diverso da $\boldsymbol{v}$. $\square$

Esempio 2.29. Come al solito, vediamo esplicitamente alcuni casi

(1) Calcoliamo le coordinate del vettore $\boldsymbol{v} = \begin{pmatrix} 1 \\ 5 \\ 10 \end{pmatrix}$ rispetto alla base canonica di $\mathbb{R}^3$.

Dobbiamo trovare tre numeri $\lambda_1, \lambda_2, \lambda_3$ tali che

$$\begin{pmatrix} 1 \\ 5 \\ 10 \end{pmatrix} = \lambda_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + \lambda_3 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

Osserviamo che calcolando il vettore a destra otteniamo

$$\begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix}$$

Dunque per avere l'uguaglianza che cerchiamo dobbiamo porre $\lambda_1 = 1, \lambda_2 = 5, \lambda_3 = 10$ e le coordinate del vettore $\boldsymbol{v}$ rispetto alla base canonica coincidono con le componenti del vettore.

Chiaramente la stessa cosa vale per qualsiasi vettore di $\mathbb{R}^3$ e più in generale vale il seguente risultato: **le coordinate di un vettore $\boldsymbol{v}$ di $\mathbb{R}^n$ rispetto alla base canonica coincidono con le componenti di $\boldsymbol{v}$.**

Come vedremo negli esempi successivi, se calcoliamo le coordinate di un vettore di $\mathbb{R}^n$ rispetto ad una base diversa da quella canonica non è più vero che le coordinate del vettore coincidono con le componenti del vettore.

(2) Mostriamo che i vettori $\boldsymbol{u}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \boldsymbol{u}_2 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \boldsymbol{u}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ formano una base di $\mathbb{R}^3$ e calcoliamo le coordinate del vettore $\boldsymbol{v} = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}$ rispetto a tale base.

Mostriamo che i vettori $\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3$ sono linearmente indipendenti. Chiaramente $\boldsymbol{u}_2$ non è multiplo di $\boldsymbol{u}_1$. Del resto tutte le combinazioni lineari di $\boldsymbol{u}_1$ e $\boldsymbol{u}_2$ hanno l'ultima componente nulla, e dunque $\boldsymbol{u}_3$ non può essere una combinazione lineare di tali vettori.

Mostriamo che i vettori $\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3$ generano $\mathbb{R}^3$. Dobbiam mostrare che qualsiasi vettore $\begin{pmatrix} a \\ b \\ c \end{pmatrix}$ è combinazione lineare di $\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3$. Ovvero, fissati dei valori arbitrari per a, b, c dobbiamo trovare dei valori per $\lambda_1, \lambda_2, \lambda_3$ in modo che

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \lambda_1 \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} + \lambda_3 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} .$$

Svolgendo la combinazione lineare a destra otteniamo

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} \lambda_1 + \lambda_2 \\ \lambda_1 - \lambda_2 \\ \lambda_3 \end{pmatrix}$$

dunque i numeri $\lambda_1, \lambda_2, \lambda_3$ che dobbiamo trovare devono essere soluzioni del sistema

$$\begin{cases} \lambda_1 + \lambda_2 = a \\ \lambda_1 - \lambda_2 = b \\ \lambda_3 = c \end{cases}$$

ora questo sistema ha sempre soluzione qualsiasi siano i valori a, b, c, infatti si ha che

$$\lambda_1 = \frac{a+b}{2}, \ \lambda_2 = \frac{a-b}{2}, \ \lambda_3 = c$$

dunque $\operatorname{Span}(\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3) = \mathbb{R}^3$.

Osserviamo che per calcolare le coordinate del vettore $\boldsymbol{v}$ dobbiamo determinare i numeri per cui

$$\boldsymbol{v} = \lambda_1 \boldsymbol{u}_1 + \lambda_2 \boldsymbol{u}_2 + \lambda_3 \boldsymbol{u}_3$$

e quindi dobbiamo ripetere il conto fatto per il generico vettore $\begin{pmatrix} a \\ b \\ c \end{pmatrix}$ nel caso particolare $a = 2, b = 0, c = 1$ e ricaviamo

$$\lambda_1 = \frac{2+0}{2} = 1, \lambda_2 = \frac{2-0}{2} = 1, \lambda_3 = 1$$

Esercizio 6.1. Si risolvano i seguenti esercizi proposti.

(1) Si dimostri che i vettori $\boldsymbol{u}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \boldsymbol{u}_2 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \boldsymbol{u}_3 = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$ formano una base di $\mathbb{R}^3$ e si calcolino le coordinate del vettore $\begin{pmatrix} 3 \\ 2 \\ 4 \end{pmatrix}$ rispetto a tale base.

(2) Si dimostri che i vettori $\boldsymbol{u}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, $\boldsymbol{u}_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$, $\boldsymbol{u}_3 = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$ formano una base di $\mathbb{R}^3$ e si calcolino le coordinate del vettore $\begin{pmatrix} -3 \\ 7 \\ 4 \end{pmatrix}$ rispetto a tale base.

6.1. Esistenza e costruzioni di basi.

Poniamoci la seguente domanda: **Domanda:** in uno spazio vettoriale V posso trovare sempre una base?

Chiaramente affinché si possa trovare una base è necessario che **lo spazio** V **sia finitamente generato** e $V \neq \{\mathbf{0}_V\}$.

Mostriamo ora un procedimento per costruire una base di V ogni volta che sia noto un sistema di generatori per V. In maniera un po' colloquiale possiamo affermare che tale algoritmo funziona buttando via i vettori superflui.

In effetti descriveremo un algoritmo che funziona in maniera un po' più generale. Prende come input una lista qualsiasi di vettori di V, diciamo $L = \{\boldsymbol{u}_1, \ldots \boldsymbol{u}_k\}$ e restituisce come output una sottolista L' che gode delle seguenti proprietà

(1) L' è un sistema linearmente indipendente
(2) $\operatorname{Span} L = \operatorname{Span} L'$

Algoritmo 2.15 (**di estrazione**). Partiamo illustrando il funzionamento dell'algoritmo in un esempio: Supponiamo di avere la seguente lista di vettori L di $\mathbb{R}^4$

$$\boldsymbol{u}_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \ \boldsymbol{u}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \ \boldsymbol{u}_3 = \begin{pmatrix} 2 \\ 3 \\ 2 \\ 3 \end{pmatrix}, \ \boldsymbol{u}_4 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \ \boldsymbol{u}_5 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}.$$

Come prima cosa ci chiediamo se $\boldsymbol{u}_5$ è combinazione lineare dei precedenti (ovvero se $\boldsymbol{u}_5 \in \operatorname{Span}(\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3, \boldsymbol{u}_4)$). La risposta è positiva (infatti $\boldsymbol{u}_5 = \boldsymbol{u}_1 + \boldsymbol{u}_2$) e dunque lo cancelliamo dalla lista.

Ci chiediamo ora se $\boldsymbol{u}_4 \in \operatorname{Span}(\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3)$. La risposta è negativa, per cui lo conserviamo.

Proseguiamo nello stesso modo, chiedendoci se $\boldsymbol{u}_3 \in \operatorname{Span}(\boldsymbol{u}_1, \boldsymbol{u}_2)$). La risposta è positiva, infatti $\boldsymbol{u}_3 = 2\boldsymbol{u}_1 + 3\boldsymbol{u}_2$. In questo caso cancelliamo $\boldsymbol{u}_3$ dalla lista.

Infine ci chiediamo se $\boldsymbol{u}_2 \in \operatorname{Span}(\boldsymbol{u}_1)$. La risposta è negativa e dunque lo conserviamo.

Alla fine la lista di output è

$$L' = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_4\} \ .$$

Per costruzione, nessun vettore di L' è combinazione lineare dei precedenti, per cui i vettori di L' sono linearmente indipendenti.

Avendo eliminato due vettori di L, mi devo assicurare che lo $\operatorname{Span} L'$ sia rimasto uguale a $\operatorname{Span} L$.

Poiché $\boldsymbol{u}_5 \in \operatorname{Span}(\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3, \boldsymbol{u}_4)$ allora abbiamo visto che

$$\operatorname{Span}(\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3, \boldsymbol{u}_4, \boldsymbol{u}_5) = \operatorname{Span}(\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3, \boldsymbol{u}_4)\,.$$

Del resto poiché $\boldsymbol{u}_3 \in \operatorname{Span}(\boldsymbol{u}_1, \boldsymbol{u}_2)$ abbiamo che

$$\operatorname{Span}(\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3, \boldsymbol{u}_4) = \operatorname{Span}(\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_4)\,.$$

Chiaramente si può ripetere lo stesso algoritmo per ogni lista di vettori $L = \{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k\}$.

Abbiamo un algoritmo in $(k-1)$-passi, dove a ciascun passo si controlla se un vettore della lista è combinazione lineare dei precedenti. In caso positivo lo si cancella, altrimenti lo si conserva.

L'output L' sarà formato da quei vettori di L che non sono combinazione lineare dei precedenti.

Chiaramente L' è una lista di vettori indipendenti.

Per verificare che $\operatorname{Span} L' = \operatorname{Span} L$, possiamo ragionare come sopra. Infatti osserviamo che a ciascun passo possiamo modificare la lista L, ma lo Span rimane uguale. La cosa è chiara quando non cancelliamo il vettore che stiamo considerando (in questo caso non modifichiamo la lista). Del resto, nel caso in cui cancelliamo un vettore, questo risulta essere combinazione lineare dei precedenti, e dunque eliminandolo lo Span risulta invariato.

Esercizio 6.2. Si applichi l'algoritmo di estrazione alle seguenti liste di vettori

(1) $\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}2\\0\\0\end{pmatrix}, \begin{pmatrix}1\\1\\0\end{pmatrix}, \begin{pmatrix}3\\1\\0\end{pmatrix}$

(2) $\begin{pmatrix}1\\2\\1\end{pmatrix}, \begin{pmatrix}1\\1\\1\end{pmatrix} \begin{pmatrix}1\\0\\0\end{pmatrix}$

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

Quando L è un insieme di generatori per V, lo è anche la lista L'.

Infatti si ha $\operatorname{Span} L' = \operatorname{Span} L = V$. Poiché L' è formato da vettori indipendenti, esso costituisce una base.

Con questo, abbiamo allora dimostrato il seguente risultato

Corollario 2.16. *Se V è uno spazio finitamente generato diverso da $\{\mathbf{0}_V\}$, allora ammette una base* $\square$

In effetti non è difficile mostrare che uno spazio vettoriale finitamente generato contiene infinite basi diverse.

Ad esempio, nel caso $\mathbb{E}^3_O$ abbiamo visto che le basi sono composte da tre vettori non complanari. Essendoci infiniti modi di scegliere 3 vettori non complanari, $\mathbb{E}^3_O$ contiene infinite basi.

Osservazione 2.30. Quando L non è una base di V, pssiamo considerare il sottospazio $W = \operatorname{Span} L$. Abbiamo visto che W è esso stesso uno spazio vettoriale in cui le operazioni di somma e moltiplicazione per scalare sono le restrizioni delle corrispondenti operazioni su V.

Osserviamo che L' è una lista di vettori indipendenti che genera W e dunque L' è una base di W.

Osservazione 2.31. La lista L' prodotta dall'algoritmo dipende dall'ordine in cui consideriamo i vettori.

Ad esempio se applichiamo l'algoritmo alla lista

$$\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}1\\1\end{pmatrix}$$

otteniamo la lista L' composta dai vettori

$$\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}.$$

Mentre se ordiniamo i vettori di L diversamente

$$\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}$$

otteniamo la lista L'

$$\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}.$$

Si osservi che in entrambi i casi L' è una lista di vettori indipendenti tali che $\operatorname{Span} L' = \operatorname{Span} L$ (in questo caso risultano essere basi di $\mathbb{R}^2$)

Algoritmo 2.17 (**Metodo di completamento**). Ci poniamo ora il seguente problema.

Domanda: Sia V uno spazio vettoriale finitamente generato e

$$L = \{\boldsymbol{u}_1, \dots, \boldsymbol{u}_k\}$$

un insieme di vettori linearmente indipendenti di V. È possibile aggiungere ad L un po' di vettori, in modo da ottenere una base?

Per risolvere questo problema si procede in questo modo. Si fissa arbitrariamente una base di V

$$\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}.$$

Si aggiungono i vettori $\boldsymbol{e}_i$ in coda alla lista L ottenendo una nuova lista

$$L' = \{\boldsymbol{u}_1, \dots, \boldsymbol{u}_k, \boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}.$$

Osserviamo che L' è un insieme di generatori per V in quanto $\operatorname{Span} L'$ contiene $\operatorname{Span}\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\} = V$.

Ora però la lista L' non è composta da vettori linearmente indipendenti, infatti ciascun $\boldsymbol{u}_i$ è combinazione lineare di $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$.

Dunque applichiamo l'algoritmo di estrazione è otteniamo una base L''.

Osserviamo che applicando l'algoritmo si cancellano solo i vettori di L' che sono combinazione lineare dei precedenti. Siccome $\boldsymbol{u}_1, \dots, \boldsymbol{u}_k$ sono indipendenti, nessuno di loro è combinazione dei precedenti e dunque essi sono tutti contenuti nella lista L''.

Esempio 2.32. Completiamo la lista di vettori indipendenti

$$L = \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} \right\}$$

ad una base di $\mathbb{R}^3$.

Aggiungiamo ad L la base canonica di $\mathbb{R}^3$ e otteniamo la lista

$$L' = \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}.$$

Ora applichiamo l'algoritmo di estrazione. Notiamo che

$$\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} - \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

e dunque lo eliminiamo dalla lista.

Allo stesso modo

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

e dunque va eliminato.

Invece $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \notin \mathrm{Span}(\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix})$, e dunque lo conserviamo e analogamente $\begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}$ non è multiplo di $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ e dunque viene conservato.

In definitiva la base che otteniamo è

$$L'' = \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\}$$

6.2. Teorema della base.

Abbiamo già visto che in $\mathbb{E}^3_O$ una base è formata da 3 vettori non complanari.

Dunque, anche se esistono infinite basi di $\mathbb{E}^3_O$ il numero degli elementi che costituiscono ciascuna base è sempre 3. In effetti tale numero è uguale al numero di coordinate necessarie per individuare un vettore ed è un fatto intuitivo che sono necessari 3 numeri per individuare un punto dello spazio indipendentemente dal sistema di riferimento. Per questa ragione alla fine diciamo che la dimensione di $\mathbb{E}^3_O$ è 3.

Tale osservazione sullo spazio geometrico si può generalizzare a qualsiasi spazio vettoriale. Infatti vale il seguente teorema fondamentale.

Teorema 2.18. *Sia V spazio vettoriale finitamente generato e siano $\mathcal{B}$, $\mathcal{B}'$ due basi di V. Allora il numero di elementi di $\mathcal{B}$ è uguale al numero di elementi di $\mathcal{B}'$.*

Dimostreremo il Teorema 2.18 come conseguenza di alcune proprietà delle basi che andiamo ora a verificare e riassumere nell'enunciato di due Lemmi.

LEMMA 2.19. *Aggiungendo o togliendo un vettore ad una base di V si ottiene una lista che* ***non*** *è una base.*

DIMOSTRAZIONE. Come prima cosa osserviamo che se $\mathcal{B} = \{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k\}$ è base di V, allora la lista $\mathcal{B}'$ che si ottiene cancellando il vettore $\boldsymbol{u}_k$ da $\mathcal{B}$ non può essere una lista di generatori. Infatti poiché $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k$ sono indipendenti, $\boldsymbol{u}_k$ non è combinazione lineare dei vettori di $\mathcal{B}' = \{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_{k-1}\}$. Chiaramente lo stesso discorso vale se cancelliamo da $\mathcal{B}$ un qualsiasi vettore.

D'altronde, se aggiungiamo a $\mathcal{B}$ un qualsiasi vettore $\boldsymbol{v} \in V$, otteniamo una lista $\mathcal{B}'' = \{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k, \boldsymbol{v}\}$ che non può essere un sistema linearmente indipendente. Infatti poiché $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k$ sono generatori, il vettore $\boldsymbol{v}$ è una combinazione lineare dei precedenti nella lista $\mathcal{B}''$. □

LEMMA 2.20. *Sia $\boldsymbol{v} \in V$. Sostituendo un vettore $\boldsymbol{u}_j$ della lista $\mathcal{B}$ con $\boldsymbol{v}$ si ottiene ancora una base purché la j-esima coordinata di $\boldsymbol{v}$ rispetto alla base $\mathcal{B}$ sia non nulla.*

DIMOSTRAZIONE. Vediamo ora cosa succede se provo a sostituire il vettore $\boldsymbol{u}_k$ con un altro vettore $\boldsymbol{v} \in V$, ottenendo così la nuova lista $\mathcal{B}''' = \{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_{k-1}, \boldsymbol{v}\}$.

Poiché $\mathcal{B}$ è una base, posso scrivere il vettore $\boldsymbol{v}$ come combinazione lineare di $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k$.

$$\boldsymbol{v} = \lambda_1 \boldsymbol{u}_1 + \ldots + \lambda_k \boldsymbol{u}_k$$

dove $\lambda_1, \ldots, \lambda_k$ sono le coordinate di v rispetto alla base $\mathcal{B}$.

Se $\lambda_k = 0$ risulta che $\boldsymbol{v}$ è combinazione lineare di $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_{k-1}\}$, per cui la lista $\mathcal{B}'''$ non è formata da vettori linearmente indipendenti.

Invece se $\lambda_k \neq 0$, non è possibile che $\boldsymbol{v}$ sia combinazione di $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_{k-1}$: infatti i vettori che si ottengono come combinazione lineare di $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_{k-1}$ sono caratterizzati da avere la coordinata λ_k nulla.

Dunque la lista $\mathcal{B}'''$ è composta da vettori indipendenti. D'altra parte osserviamo che $\boldsymbol{u}_k$ è combinazione lineare di $\boldsymbol{u}_1 \ldots \boldsymbol{u}_{k-1}, \boldsymbol{v}$, infatti abbiamo

$$\boldsymbol{u}_k = \frac{1}{\lambda_k}(\boldsymbol{v} - \lambda_1 \boldsymbol{u}_1 - \ldots - \lambda_{k-1} \boldsymbol{u}_{k-1})$$

e dunque il vettore $\boldsymbol{u}_k$ appartiene a $\operatorname{Span} \mathcal{B}'''$.

Presa una combinazione lineare di $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k$

$$\mu_1 \boldsymbol{u}_1 + \ldots + \mu_k \boldsymbol{u}_k\,.$$

osserviamo che tutti gli addendi che vi compaiono sono contenuti in $\operatorname{Span} \mathcal{B}'''$ e dunque lo è anche la loro somma. Poiché tutti i vettori di V si possono ottenere come combinazioni lineari di $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k$, ricaviamo che $\operatorname{Span} \mathcal{B}''' = V$.

Dunque $\mathcal{B}'''$ è anche un sistema di generatori e dunque è una base.

Chiaramente avremmo ottenuto lo stesso risultato sostituendo un qualsiasi vettore $\boldsymbol{u}_j$ di $\mathcal{B}$ con un vettore $\boldsymbol{v}$, purché la j-esima coordinata di $\boldsymbol{v}$ rispetto alla base $\mathcal{B}$ fosse non nulla □

Utilizzando le due proprietà esposte sopra nei Lemmi 2.19, 2.20 è possibile dimostrare il risultato fondamentale che ci permetterà di dimostrare il Teorema 2.18

PROPOSIZIONE 2.21. *Sia V spazio vettoriale e sia $\mathcal{B} = \{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k\}$ base di V. Una qualsiasi lista di vettori indipendenti $\mathcal{D} = \{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_k\}$ contenente lo stesso numero di elementi di $\mathcal{B}$ è una base.*

DIMOSTRAZIONE. L'idea della dimostrazione è quella di mostrare che è possibile scambiare uno alla volta gli elementi di $\mathcal{B}$ con gli elementi di $\mathcal{D}$ in modo da ottenere sempre basi. Alla fine quando avremo scambiato tutti gli elementi risulterà che $\mathcal{D}$ stessa è una base.

Per semplicità ci limiteremo a mostrare questa procedura nel caso in cui le liste $\mathcal{B}$ e $\mathcal{D}$ siano formate da 4 vettori. Il caso generico è però del tutto simile.

Sostituisco un elemento di $\mathcal{B} = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3, \boldsymbol{u}_4\}$ con $\boldsymbol{w}_1$. Per scegliere l'elemento da sostituire, guardo le coordinate di $\boldsymbol{w}_1$ rispetto a $\mathcal{B}$ e fisso l'elemento in modo che la corrispondente coordinata sia non nulla. Ottengo una nuova base $\mathcal{B}_1$ formata da $\boldsymbol{w}_1$ e tre vettori della base $\mathcal{B}$. Poiché non è importante stabilire quali sono questi tre elementi, li rinomino per semplicità $\boldsymbol{z}_1, \boldsymbol{z}_2, \boldsymbol{z}_3$. In tal modo la base $\mathcal{B}_1$ è formata da $\boldsymbol{w}_1, \boldsymbol{z}_1, \boldsymbol{z}_2, \boldsymbol{z}_3$.

A questo punto in $\mathcal{B}_1$ sostituisco uno tra $\boldsymbol{z}_1, \boldsymbol{z}_2, \boldsymbol{z}_3$ con $\boldsymbol{w}_2$. Per scegliere tale elemento scrivo $\boldsymbol{w}_2$ come combinazione lineare dei vettori $\mathcal{B}_1$, ovvero

$$\boldsymbol{w}_2 = \lambda \boldsymbol{w}_1 + \mu_1 \boldsymbol{z}_1 + \mu_2 \boldsymbol{z}_2 + \mu_3 \boldsymbol{z}_3$$

Poiché $\boldsymbol{w}_2$ non è multiplo di $\boldsymbol{w}_1$, uno tra μ_1, μ_2, μ_3 deve essere non nullo. Scelgo dunque lo $\boldsymbol{z}_i$ da sostituire in corrispondenza di un $\mu_i \neq 0$.

Facendo tale scelta, otteniamo una nuova base $\mathcal{B}_2$ formata da $\boldsymbol{w}_1, \boldsymbol{w}_2$ e da altri due vettori che rinominiamo $\boldsymbol{y}_1, \boldsymbol{y}_2$.

A questo punto sostituisco in $\mathcal{B}_2$ uno tra $\boldsymbol{y}_1, \boldsymbol{y}_2$ con $\boldsymbol{w}_3$. Per scegliere l'elemento da sostituire come al solito scrivo $\boldsymbol{w}_3$ come combinazione lineare degli elmenti di $\mathcal{B}_2$

$$\boldsymbol{w}_3 = \lambda_1 \boldsymbol{w}_1 + \lambda_2 \boldsymbol{w}_2 + \mu_1 \boldsymbol{y}_1 + \mu_2 \boldsymbol{y}_2$$

Poiché $\boldsymbol{w}_3$ non è combinazione di $\boldsymbol{w}_1$ e $\boldsymbol{w}_2$ ricavo che μ_1 e μ_2 non possono essere entrambi nulli. Dunque scelgo il $\boldsymbol{y}_i$ da sostituire in modo che il corrispondente μ_i sia non nullo.

Facendo questa sostituzione ottengo ancora una base $\mathcal{B}_3$ formata da $\boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{w}_3$ e un altro vettore che rinominiamo $\boldsymbol{v}$. Ora se scriviamo $\boldsymbol{w}_4$ come combinazione lineare degli elementi di $\mathcal{B}_4$

$$\boldsymbol{w}_4 = \lambda_1 \boldsymbol{w}_1 + \lambda_2 \boldsymbol{w}_2 + \lambda_3 \boldsymbol{w}_3 + \mu \boldsymbol{v}$$

osserviamo che $\mu \neq 0$: altrimenti $\boldsymbol{w}_4$ sarebbe combinazione lineare di $\boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{w}_3$. Segue che sostituendo $\boldsymbol{v}$ con $\boldsymbol{w}_4$ ottengo di nuovo una base.

Ho dunque dimostrato che i vettori $\{\boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{w}_3, \boldsymbol{w}_4\}$ formano una base. $\square$

Possiamo ora dimostrare facilmente il Teorema 2.18

DIMOSTRAZIONE DEL TEOREMA 2.18. Fissiamo due basi $\mathcal{B} = \{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k\}$ e $\mathcal{D} = \{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_h\}$. Supponiamo che $\mathcal{B}$ abbia meno elementi di $\mathcal{D}$ (ovvero $k \leq h$). Consideriamo la lista formata dai primi k vettori di $\mathcal{D}$

$$\mathcal{D}' = \{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_k\} .$$

Chiaramente $\mathcal{D}'$ è formato da k vettori indipendenti. Applicando la Proposizione 2.21 a $\mathcal{B}$ e $\mathcal{D}'$, deduciamo che $\mathcal{D}'$ è una base. Poiché togliendo elementi da una base non si ottiene mai una base (Lemma 2.19) concludiamo che $\mathcal{D}'$ deve contenere tutti gli elementi di $\mathcal{D}$, ovvero $\mathcal{D}' = \mathcal{D}$. Segue che $k = h$. $\square$

OSSERVAZIONE 2.33. Come conseguenza immediata del Teorema 2.18, possiamo dire che una qualsiasi base di $\mathbb{R}^n$ contiene esattamente n elementi.

In base al risultato del Teorema 2.18 resta, pertanto, giustificata la seguente

DEFINIZIONE 2.17. Sia V uno spazio vettoriale finitamente generato. Se $V \neq \{\mathbf{0}_V\}$ la **dimensione** di V –che denoteremo $\dim V$– è il numero di elementi che costituiscono una base di V.

Se $V = \{\mathbf{0}_V\}$ poniamo per convenzione $\dim V = 0$.

Dal Teorema delle basi 2.18 ricaviamo anche due implicazioni molto importanti:

PROPOSIZIONE 2.22. *Sia V spazio vettoriale con* $\dim V = n$*,* $n \geq 1$.

(1) *Una lista di vettori indipendenti contiene al massimo n elementi.*
(2) *Applicando il metodo di completamento ad una lista L di k vettori indipendenti (con $k \leq n$), si aggiungono ad L esattamente $n-k$ vettori.*
(3) *Una lista L di n vettori indipendenti è una base.*

DIMOSTRAZIONE. Osserviamo che applicando il metodo di completamento si ottiene una base $\mathcal{B}$ che contiene L. Ora $\mathcal{B}$ contiene n elementi come tutte le basi di V, e L contiene k elementi.

È chiaro che $k \leq n$ e che il numero di elementi che bisogna aggiungere ad L per ottenere una base è $n - k$.

Se $n = k$, abbiamo che $\mathcal{B}$ e L contengono lo stesso numero di elementi. Poiché L è un sottoinsieme di $\mathcal{B}$, deduciamo che $L = \mathcal{B}$ e dunque L è una base. □

Ed anche

PROPOSIZIONE 2.23. *Sia V spazio vettoriale con* $\dim V = n$*,* $n \geq 1$.

(1) *Una lista $L = \{v_1, \dots, v_k\}$ di generatori contiene almeno n elementi (ovvero $k \geq n$)*
(2) *Applicando l'algoritmo di estrazione ad una lista di k generatori (con $k > n$) si scartano esattamente $k - n$ vettori*
(3) *Una sistema di n generatori è una base.*

DIMOSTRAZIONE. Applicando l'algoritmo di estrazione si trova una sottolista $\mathcal{B}$ di L che è una base e che dunque conta n elementi. Segue che $k > n$. Inoltre il numero di elementi che si è dovuto cancellare per ottenere $\mathcal{B}$ è $n - k$ Infine se $n = k$, allora $\mathcal{B} = L$ e dunque L è una base □

Un modo per sintetizzare i contenuti delle due Proposizioni 2.22, 2.23 è il seguente.

COROLLARIO 2.24. *Sia V uno spazio vettoriale di dimensione n e sia L una lista di vettori. Se due delle seguenti proposizioni sono vere allora anche la terza è vera:*

(1) *L contiene n vettori;*
(2) *L è un insieme di generatori;*
(3) *L è un insieme linearmente indipendente.*

$\square$

ESEMPIO 2.34. Mostriamo che i vettori

$$\boldsymbol{u}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \boldsymbol{u}_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \boldsymbol{u}_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

formano una base di $\mathbb{R}^3$.

Osserviamo che poiché sono in numero giusto, sarà sufficiente mostrare o che sono indipendenti o che sono generatori di $\mathbb{R}^3$

Poiché in generale l'indipendenza è più semplice da verificare, controlliamo questa proprietà.

Chiaramente $\boldsymbol{u}_2$ non è multiplo di $\boldsymbol{u}_1$.

Vediamo se $\boldsymbol{u}_3$ è combinazione lineare di $\boldsymbol{u}_1$ e $\boldsymbol{u}_2$. Ci chiediamo se è possibile attribuire dei valori numerici a λ e μ in modo che valga la seguente uguaglianza

$$\boldsymbol{u}_3 = \lambda \boldsymbol{u}_1 + \mu \boldsymbol{u}_2$$

Ora λ e μ dovrebbero essere soluzioni del sistema

$$\begin{cases} \lambda = 1 \\ \lambda + \mu = 1 \\ \mu = 1 \end{cases}$$

Tale sistema ovviamente non ha mai soluzione e dunque i vettori $\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3$ sono una base di $\mathbb{R}^3$.

7. Dimensione e sottospazi

Abbiamo già osservato che se W è un sottospazio di uno spazio vettoriale V, allora le operazioni di somma e moltiplicazione per scalare su V inducono su W una struttura di spazio vettoriale.

In questa sezione ci porremo il problema di stabilire se e come sia possibile trovare una base di un sottospazio di V.

Osserviamo che una base di W è una lista di vettori $\{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_k\}$ di W tali che

- $\mathrm{Span}(\boldsymbol{w}_1, \ldots, \boldsymbol{w}_k) = W$.
- $\{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_k\}$ sono linearmente indipendenti.

7.1. Sottospazi di spazi finitamente generati.

Fissiamo uno spazio vettoriale finitamente generato V con $\dim V = n$. Come prima cosa ci domandiamo se i sottospazi di V sono tutti finitamente generati. In altre parole, dato un sottospazio $W \subset V$ ci chiediamo se esiste una lista di vettori $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_k$ tali che $W = \mathrm{Span}(\boldsymbol{w}_1, \ldots, \boldsymbol{w}_k)$.

PROPOSIZIONE 2.25. *Sia V uno spazio finitamente generato con* $\dim V = n$. *Allora ogni sottospazio W di V è finitamente generato. Inoltre si ha che* $\dim W \leq n$. *Se* $\dim W = n$ *allora* $W = V$.

DIMOSTRAZIONE. Fisso $\boldsymbol{w}_1 \in W$. Se $W = \operatorname{Span} \boldsymbol{w}_1$ mi fermo.

Altrimenti fisso un elemento $\boldsymbol{w}_2 \in W$ che non appartiene a $\operatorname{Span}(\boldsymbol{w}_1)$. Se $W = \operatorname{Span}(\boldsymbol{w}_1, \boldsymbol{w}_2)$ mi fermo.

Altrimenti fisso un elemento $\boldsymbol{w}_3 \in W$ che non appartiene a $\operatorname{Span}(\boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{w}_3)$ e continuo così.

Questo algoritmo si ferma solo se ad un certo punto avrò abbastanza elementi $\boldsymbol{w}_1 \dots, \boldsymbol{w}_k$ tali che $W = \operatorname{Span}(\boldsymbol{w}_1, \dots \boldsymbol{w}_k)$. In particolare se W non è finitamente generato l'algoritmo non si ferma mai.

Del resto osserviamo che dopo k passi l'algoritmo produce una lista di k vettori indipendenti (infatti per costruzione nessuno è combinazione lineare dei precedenti). Dunque per la Proposizione 2.22 mi devo fermare al massimo dopo n passi. Supponiamo di fermarci al passo m. Allora abbiamo

$$W = \operatorname{Span}(\boldsymbol{w}_1, \dots, \boldsymbol{w}_m) \tag{2.17}$$

ovvero W è generato da m vettori che sono indipendenti. Duque $\boldsymbol{w}_1, \dots, \boldsymbol{w}_m$ è una base di W e segue $\dim W = m$.

Poiché abbiamo visto che mi devo fermare al massimo dopo n passi, si ha $\dim W = m \leq n$.

Se mi fermo esattamente al passo n-esimo (ovvero $n = m$), allora $\boldsymbol{w}_1, \dots,$ $\boldsymbol{w}_m$ formano una lista di n vettori indipendenti, e sempre per la Proposizione 2.22, essi formano una base di V, per cui abbiamo anche

$$\operatorname{Span}(\boldsymbol{w}_1, \dots, \boldsymbol{w}_n) = V$$

e confrontando tale uguaglianza con (2.17) concludiamo che $W = V$. □

OSSERVAZIONE 2.35. Abbiamo ora gli strumenti per verificare che gli unici sottospazi di $\mathbb{E}^3_O$ sono

(1) $\left\{\overrightarrow{OO}\right\}$;
(2) le rette per O;
(3) i piani per O;
(4) $\mathbb{E}^3_O$.

Infatti se W è un sottospazio di $\mathbb{E}^3_O$ la dimensione di W può assumere i valori compresi tra 0 e 3.

Se W ha dimensione 0 allora coincide con il sottospazio nullo.

Se W ha dimensione 1, allora ammette una base composta da un solo vettore $\boldsymbol{u}$ e dunque $W = \operatorname{Span} \boldsymbol{u}$ è una retta per O.

Se W ha dimensione 2 allora ammette una base composta da due vettore $\boldsymbol{u}, \boldsymbol{w}$. Osserviamo che poiché sono indipendenti, $\boldsymbol{u}, \boldsymbol{w}$ non sono paralleli e dunque $W = \operatorname{Span}(\boldsymbol{u}, \boldsymbol{w})$ è un piano per O.

Infine se $\dim W = 3$ allora $W = \mathbb{E}^3_O$.

7.2. Ricerca di una base del sottospazio somma.
In questa sezione ci occuperemo del seguente problema:

Problema: Supponiamo siano dati due sottospazi vettoriali $U, W \subset V$ e che siano note una base di U e una base di W. Come calcolare una base di $U + W$?

Tale problema ha una risposta molto semplice. Infatti se $\mathcal{B} = \{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_h\}$ è base di U e $\mathcal{B}' = \{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_k\}$ è base di W allora la lista $\mathcal{B} \cup \mathcal{B}'$ ottenuta mettendo insieme $\mathcal{B}$ e $\mathcal{B}'$ è **un sistema di generatori** di $U + W$. E dunque applicando l'algoritmo di estrazione a $\mathcal{B} \cup \mathcal{B}'$ si ottiene una base di $U + W$.

Per verificare che $\mathcal{B} \cup \mathcal{B}'$ è un insieme di generatori per $U + W$ si osservi che ogni vettore $\boldsymbol{v} \in U + W$ può essere decomposto in

$$\boldsymbol{v} = \boldsymbol{u} + \boldsymbol{w}$$

con $\boldsymbol{u} \in U$ e $\boldsymbol{w} \in W$. Del resto poiché $\mathcal{B}$ è base di U abbiamo che il vettore $\boldsymbol{u}$ è combinazione lineare di $\boldsymbol{u}_1 \ldots \boldsymbol{u}_h$

$$\boldsymbol{u} = \lambda_1 \boldsymbol{u}_1 + \ldots + \lambda_h \boldsymbol{u}_h$$

e analogamente avremo

$$\boldsymbol{w} = \mu_1 \boldsymbol{w}_1 + \ldots + \mu_k \boldsymbol{w}_k$$

Mettendo insieme queste uguaglianze otteniamo

$$\boldsymbol{v} = \lambda_1 \boldsymbol{u}_1 + \ldots + \lambda_h \boldsymbol{u}_h + \mu_1 \boldsymbol{w}_1 + \ldots + \mu_k \boldsymbol{w}_k$$

ovvero $\boldsymbol{v}$ è combinazione lineare dei vettori $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_h, \boldsymbol{w}_1, \ldots, \boldsymbol{w}_k$.

Esempio 2.36. 2.43 Consideriamo il caso $U = \operatorname{Span} \mathcal{B}$ e $W = \operatorname{Span} \mathcal{B}'$ dove

$$\mathcal{B} = \left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \right\}, \quad \mathcal{B}' = \left\{ \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ -1 \\ 0 \end{pmatrix} \right\}$$

In questo caso $\mathcal{B} \cup \mathcal{B}'$ è la lista

$$\left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ -1 \\ 0 \end{pmatrix} \right\}.$$

A questo punto applicando l'algoritmo di estrazione si verifica che il secondo vettore non è multiplo del primo né il terzo è combinazione lineare dei primi due. Invece il quarto si può scrivere come

$$\begin{pmatrix} 2 \\ 0 \\ -1 \\ 0 \end{pmatrix} = 2 \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix}.$$

e dunque lo scartiamo. In definitiva una base di $U + W$ è data dai vettori

$$\left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix} \right\}$$

OSSERVAZIONE 2.37. L'esempio mostra che la lista $\mathcal{B} \cup \mathcal{B}'$ non è in generale formata da vettori indipendenti, dunque **in generale non è una base di** $U + W$.

Per ottenere una base bisogna applicare a tale lista l'algoritmo di estrazione.

7.3. La formula di Grassmann.

In questa sezione dimostreremo un'importante relazione numerica — detta formula di Grassmann — che lega la $\dim U$ e la $\dim W$ con la $\dim(U + W)$ e la $\dim(U \cap W)$.

PROPOSIZIONE 2.26. *Siano U e W due sottospazi vettoriali di uno spazio vettoriale V finitamente generato. Allora si ha*

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

DIMOSTRAZIONE. Come prima cosa fissiamo una base di $U \cap W$

$$\mathcal{B}_{U \cap W} = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$$

in tal modo si ha che

$$\dim(U \cap W) = k \,. \tag{2.18}$$

Ora poiché $U \cap W$ è un sottospazio di U, possiamo applicare l'algoritmo di completamento a $\mathcal{B}_{U \cap W}$ e aggiungendo un certo numero di vettori $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_p$ di U otteniamo una base di U della forma

$$\mathcal{B}_U = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k, \boldsymbol{u}_1, \ldots, \boldsymbol{u}_p\}$$

Osserviamo che tale lista contiene $k + p$ vettori (k della base di $U \cap W$ più p aggiunti dall'algoritmo di completamento). Dunque si ha

$$\dim U = k + p \,. \tag{2.19}$$

Poiché $U \cap W$ è anche sottospazio di W possiamo analogamente aggiungere a $\mathcal{B}_{U \cap W}$ un certo numero di vettori $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_q$ di W e ottenere una base di W della forma

$$\mathcal{B}_W = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k, \boldsymbol{w}_1, \ldots, \boldsymbol{w}_q\} \ .$$

Questa lista contiene $k + q$ elementi per cui si ha

$$\dim W = n + q \tag{2.20}$$

Da quanto visto nella sezione precedente, mettendo insieme $\mathcal{B}_U$ e una base di $\mathcal{B}_W$ si ottiene un insieme di generatori di $U + W$. Ovvero la lista

$$\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k,\ \boldsymbol{u}_1, \ldots, \boldsymbol{u}_p,\ \boldsymbol{v}_1, \ldots, \boldsymbol{v}_k,\ \boldsymbol{w}_1, \ldots, \boldsymbol{w}_q\}$$

è un insieme di generatori di $U + W$. Notiamo che in tale lista gli elementi $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k$ appaiono due volte, per cui eliminandoli la seconda volta che appaiono otteniamo una lista

$$L = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k,\ \boldsymbol{u}_1, \ldots, \boldsymbol{u}_p,\ \boldsymbol{w}_1, \ldots, \boldsymbol{w}_q\}$$

che è ancora una lista di generatori di $U + W$.

Affermiamo ora che i vettori in L sono linearmente indipendenti. Prima di dimostrare tale affermazione, vediamo come si può concludere la dimostarzione. Poiché i vettori di L sono indipendenti e generano $U+W$, essi costituiscono una base di $U+W$. Dunque la dimensione di $U+W$ è pari al numero di elementi che costituiscono L Tale numero è uguale a $k+p+q$ e dunque si ha

$$\dim(U+W) = k+p+q\,. \tag{2.21}$$

Ora dall'equazioni (2.19) e (2.20) si ha

$$\dim U + \dim W = k+p+k+q = 2k+p+q$$

Del resto per le equazioni (2.18) e (2.21) otteniamo che

$$\dim(U\cap W) + \dim(U+W) = k+k+p+q = 2k+p+q\,.$$

Per cui confornatando queste uguaglianze deduciamo la formula

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

Per concludere la dimostrazione è allora sufficiente verificare che i vettori della lista L sono indipendenti. Verificheremo che se abbiamo una combinazione lineare nulla

$$\lambda_1\boldsymbol{v}_1 + \ldots + \lambda_k\boldsymbol{v}_k + \mu_1\boldsymbol{u}_1 + \ldots + \mu_p\boldsymbol{u}_p + \nu_1\boldsymbol{w}_1 + \ldots + \nu_q\boldsymbol{w}_q = \mathbf{0}_V \tag{2.22}$$

allora necessariamente tutti i coefficienti sono uguali a 0.

Come prima cosa mostreremo che sono nulli i coefficienti $\mu_1,\ldots,\mu_p$. Consideriamo i due vettori

$$\mathbf{x} = \lambda_1\boldsymbol{v}_1 + \ldots \lambda_k\boldsymbol{v}_k + \mu_1\boldsymbol{u}_1 + \ldots \mu_p\boldsymbol{u}_p \qquad \mathbf{y} = \nu_1\boldsymbol{w}_1 + \ldots \nu_q\boldsymbol{w}_q \tag{2.23}$$

Osserviamo che $\mathbf{x}\in U$ perchè è combinazione lineare di elementi di $\mathcal{B}_U$, mentre $\mathbf{y}\in W$ in quanto combinazione lineare di elementi di $\mathcal{B}_W$. Del resto per ipotesi si ha $\mathbf{x}+\mathbf{y}=\mathbf{0}_V$, per cui si ha

$$\mathbf{x} = -\mathbf{y}$$

Dunque deduciamo che $\mathbf{x}$ appartiene anche a W e dunque $\mathbf{x}\in U\cap W$. Questo vuol dire che $\mathbf{x}$ si scrive come combinazione lineare dei soli $\boldsymbol{v}_1,\ldots,\boldsymbol{v}_k$

$$\mathbf{x} = \alpha_1\boldsymbol{v}_1 + \ldots + \boldsymbol{v}_k\,.$$

Del resto essendo $\mathcal{B}_U$ una base di U, se confrontiamo la decomposizione di $\mathbf{x}$ in (2.23) con questa decomposizione deduciamo che i coefficienti $\mu_1,\ldots,\mu_p$ sono tutti nulli (e che i coefficienti λ_i sono uguali ai corrispondenti coefficienti α_i).

Ora, se i coefficienti $\mu_1,\ldots,\mu_p$ sono tutti nulli, la combinazione lineare (2.22) diventa

$$\lambda_1\boldsymbol{v}_1 + \ldots + \lambda_k\boldsymbol{v}_k + \nu_1\boldsymbol{w}_1 + \ldots + \nu_q\boldsymbol{w}_q = \mathbf{0}_V$$

I vettori che appaiono in tale combinazione lineare sono tutti vettori di $\mathcal{B}_W$, e dunque per ipotesi sono indipendenti. Si conclude che anche i coefficienti $\lambda_1,\ldots\lambda_k,\nu_1,\ldots,\nu_q$ sono tutti nulli. □

Osservazione 2.38. Definiamo la dimensione del sottospazio nullo uguale a 0. Con tale scelta la formula di Grassmann rimane valida anche quando $U\cap W$ risulta essere il sottospazio nullo.

Verifichiamo questa formula in alcuni casi particolari:

- Se $U=W$ allora abbiamo visto che $U+W=U+U=U$ mentre $U\cap W = U\cap U = U$, per cui la formula è banalmente vera.
- Se $U\subset W$, allora abbiamo visto che $U+W=W$ mentre $U\cap W=U$. Ancora la formula risulta banalmente verificata.

- Se $U \cap W = \{0_V\}$. In questo caso non ci sono vettori non nulli nell'intersezione. Per tale motivo la lista $\mathcal{B} \cup \mathcal{B}'$ risulta composta da vettori linearmente indipendenti (altrimenti il procedimento esposto nella sezione precedente produrrebbe un vettore non nullo nell'intersezione). Dunque in questo acso abbiamo $\dim(U + W) = \dim U + \dim W$.

La formula di Grassmann ha alcune conseguenze molto semplici:

(1) In $\mathbb{E}^3_O$ due piani passanti per l'origine (e non coincidenti) si intersecano lungo la retta passante per 0. Infatti siano U e W tali piani. Allora $\dim U = \dim W = 2$. Dunque per la formula di Grassmann

$$4 = \dim(U + W) + \dim(U \cap W)$$

Osserviamo che $U + W$ è un sottospazio di $\mathbb{E}^3_O$ e dunque $\dim(U + W) \leq 3$. Dunque $\dim(U \cap W) \geq 1$. Dunque o $\dim U \cap W = 1$ oppure $\dim(U \cap W) = 2$. In quest'ultimo caso $U \cap W$ sarebbe un sottospazio di U di dimensione 2. Poiché U ha dimensione 2 segue che $U \cap W = U$, ovvero $U \subset W$, ma U e W hanno la stessa dimensione, per cui concluderei che $U = W$. Dunque $\dim U \cap W = 1$.

(2) Se U, W sono sottospazi di $\mathbb{R}^{10}$ di dimensioni 7 e 6 allora $\dim(U \cap W) \geq 3$: osserviamo in questo caso che $\dim(U + W) \leq 10$ e dunque dalla formula di Grassmann otteniamo

$$13 - \dim(U \cap W) \leq 10$$

da cui $\dim(U \cap W) \geq 3$.

(3) Analogamente al precedente caso se U, W sono sottospazi di $\mathbb{R}^7$ e U ha dimensione 4 e W ha dimensione 5 allora $\dim(U \cap W) \geq 2$

(4) In generale abbiamo che se $\dim(U) + \dim(W) \geq \dim(V)$ allora $\dim(U \cap W) \geq \dim(U) + \dim(W) - \dim V$. In particolare se $U \cap W = \{0\}$ allora $\dim U + \dim W \leq \dim V$.

8. Somma diretta di sottospazi

DEFINIZIONE 2.18. Siano U, W sottospazi di V. Diremo che i due sottospazi sono in **somma diretta** se $U \cap W = \{0\}$. In questo caso il sottospazio $U + W$ verrà denotato $U \oplus W$.

Dunque con il simbolo $U \oplus W$ denoteremo $U + W$ **solo** nel caso in cui $U \cap W = \{0\}$.

ESEMPIO 2.39. Due rette non coincidenti in $\mathbb{E}^3_O$ sono sempre in somma diretta.

Osserviamo invce che due piani non sono mai in somma diretta.

Infatti che se $\dim U + \dim W > \dim V$ non è possibile che U e W siano in somma diretta.

Osservazione 2.40. Se U e W sono in somma diretta, dalla formula di Grassmann si ricava facilmente che

$$\dim(U \oplus W) = \dim U + \dim W\,.$$

In particolare date una base $\mathcal{B}$ di U e una base $\mathcal{B}'$ di W, la lista di vettori $\mathcal{B} \cup \mathcal{B}'$ è una base di $U \oplus W$ (infatti è un insieme di generatori che ha il numero giusto di elementi).

Osservazione 2.41. Supponiamo ora di avere due sottospazi U, W per cui vale la relazione $\dim(U+W) = \dim U + \dim W$. Allora sempre dalla formula di Grassmann deduciamo che $\dim U \cap W = 0$. Dunque $U \cap W$ è necessariamente il sottospazio nullo, per cui U e W sono in somma diretta.

In altre parole i sottospazi in somma diretta sono caratterizzati dalla proprietà che la dimensione di $U+W$ è uguale alla somma delle dimensioni. Ovvero che l'unione di una base di U e di una base di W è una base di $U+W$.

Le due osservazioni precedenti suggeriscono che la nozione di somma diretta si possa estendere al caso di molti sottospazi.

Definizione 2.19. Un insieme di sottospazi $U_1 \ldots, U_k$ si dice in somma diretta se comunque sia fissata una base $\mathcal{B}_1$ di U_1, una base $\mathcal{B}_2$ di U_2 etc. l'unione di tali basi $\mathcal{B} = \mathcal{B}_1 \cup \ldots \cup \mathcal{B}_2$ è una base per $U_1 + U_2 + \ldots + U_k$. In tal caso la somma di questi sottospazi verrà denotata con

$$U_1 \oplus U_2 \oplus \ldots \oplus U_k\,.$$

Osservazione 2.42. Se $U_1 \ldots, U_k$ sono in somma diretta allora

$$\dim(U_1 \oplus U_2 \oplus \ldots \oplus U_k) = \dim U_1 + \dim U_2 + \ldots + \dim U_k\,.$$

ATTENZIONE:
per verificare che $U_1, \ldots, U_k$ siano in somma diretta **NON** è sufficiente mostrare che siano a due a due in somma diretta. Ad esempio se

$$U_1 = \operatorname{Span} \boldsymbol{u}_1, U_2 = \operatorname{Span} \boldsymbol{u}_2, U_3 = \operatorname{Span} \boldsymbol{u}_3$$

sono rette su un piano allora chiaramente si ha che a due a due si intersecano nell'origine (e dunque U_1 e U_2 sono in somma diretta, così come lo sono i due sottospazi U_1 e U_3 e i due sottospazi U_2 e U_3)

Del resto $U_1 + U_2 + U_3$ ìl piano che contiene le tre rette ma non è vero che $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3\}$ è una base del piano.

La caratterizzazione più semplice di spazi in somma diretta è la seguente.

Proposizione 2.27. *$U_1, \ldots, U_k$ sono in somma diretta se*

- $U_2 \cap U_1 = \{\mathbf{0}_V\}$;
- $U_3 \cap (U_1 \oplus U_2) = \{\mathbf{0}_V\}$;
- $U_4 \cap (U_1 \oplus U_2 \oplus U_3) = \{\mathbf{0}_V\}$
- *e così via.*

Definizione 2.20. Sia U sottospazio di V. Un sottospazio W si dice **complementare** di U in V se

$$V = U \oplus W$$

Se U è un piano passante per l'origine in $\mathbb{E}^3_O$, un complementare di U è una qualsiasi retta non contenuta in U passante per l'origine.

Analogamente se U è una retta passante per l'origine, un qualsiasi piano che non la contiene è un suo complementare.

In particolare si ha che in generale **un sottospazio ammette infiniti complementari**.

Se V ha dimensione n e U ha dimensione k, un complementare di U ha dimensione $n-k$. Infatti abbiamo visto che per spazi in somma diretta vale la proprietà che il sottospazio somma ha dimensione uguale alla somma dei sottospazi per cui si ha

$$\dim V = \dim(U \oplus W) = \dim U + \dim W = k + \dim W$$

Viceversa se U e W sono in somma diretta e $\dim W = n-k$, allora risulta che $\dim(U \oplus W) = k + (n-k) = n$ e dunque $U \oplus W = V$.

Possiamo riassumere questa osservazione con il seguente enunciato.

PROPOSIZIONE 2.28. *Sia V spazio vettoriale di dimensione n e U sottospazio di V di dimensione k. Un sottospazio W di V è un complementare di U se e solo se $U \cap W = \{0\}$ e $\dim W = n-k$*

Problema: dato un sottospazio U di W, come faccio a trovarne un complementare?

Fisso una base $\mathcal{B} = \{\boldsymbol{u}_1, \dots, \boldsymbol{u}_k\}$ di W. Con il metodo di completamento ottengo una base $\mathcal{B}' = \{\boldsymbol{u}_1, \dots, \boldsymbol{u}_k, \boldsymbol{w}_1, \dots, \boldsymbol{w}_{n-k}\}$ di V che contiene $\mathcal{B}$. Considero la lista dei vettori che sono stati aggiunti a $\mathcal{B}$

$$\mathcal{B}'' = \{\boldsymbol{w}_1, \dots, \boldsymbol{w}_{n-k}\}$$

Pongo $W = \operatorname{Span} \mathcal{B}''$.

Si osservi che $U + W = V$ infatti la lista $\mathcal{B} = \mathcal{B}'' = \mathcal{B}'$ è un sistema di generatori per V. Inoltre osserviamo che $\dim U = k$, $\dim W = n-k$ e $\dim(U+W) = n$ per cui U e W sono in somma diretta.

APPENDICE

9. Ricerca di una base del sottospazio intersezione

Nella sezione 7.2 abbiamo mostrato come trovare una base della somma, ci poniamo il problema di determinare una base dell'interesezione. Come vedremo la soluzione per questo problema è lievemente più sofisticata.

Consideriamo l'esempio 2.43. In quell'esempio avevamo $U = \operatorname{Span}\mathcal{B}$ e $W = \operatorname{Span}\mathcal{B}'$ dove

$$\mathcal{B} = \left\{ \begin{pmatrix}1\\1\\0\\1\end{pmatrix}, \begin{pmatrix}1\\1\\1\\1\end{pmatrix} \right\}, \quad \mathcal{B}' = \left\{ \begin{pmatrix}1\\0\\-1\\0\end{pmatrix}, \begin{pmatrix}2\\0\\-1\\0\end{pmatrix} \right\}$$

Applicando l'algoritmo di estrazione a $\mathcal{B}\cup\mathcal{B}'$ abbiamo scartato il vettore $\begin{pmatrix}2\\0\\-1\\0\end{pmatrix}$ in quanto questo vettore si può scrivere come combinazione lineare dei precedenti

$$\begin{pmatrix}2\\0\\-1\\0\end{pmatrix} = 2\begin{pmatrix}1\\0\\-1\\0\end{pmatrix} + \begin{pmatrix}1\\1\\1\\1\end{pmatrix} - \begin{pmatrix}1\\1\\0\\1\end{pmatrix}.$$

Portando a sinistra il vettore di $\mathcal{B}'$ ricaviamo che

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

Tali uguaglianze mostrano che il vettore $\begin{pmatrix}0\\0\\1\\0\end{pmatrix}$ può essere scritto sia come combinazione lineare dei vettori della lista $\mathcal{B}$ sia come combinazione lineare dei vettori di $\mathcal{B}'$.

Dunque tale vettore appartiene sia ad U che a W, ovvero abbiamo determinato un vettore che appartiene all'intersezione $U \cap W$.

Ciò succede in generale. Sia $\mathcal{B} = \{\boldsymbol{u}_1, \dots, \boldsymbol{u}_h\}$ base di un sottospazio U e $\mathcal{B}' = \{\boldsymbol{w}_1, \dots, \boldsymbol{w}_k\}$ base di W. Applicando l'algoritmo di estrazione a $\mathcal{B} \cup \mathcal{B}'$, ogni volta che viene scartato un vettore, se ne produce contestualmente **un altro** che appartiene a $U \cap W$.

Proviamo ad analizzare un altro esempio

Esempio 2.43. Supponiamo che U, W siano sottospazi di $\mathbb{R}^4$ e che una base di U sia

$$\mathcal{B} = \left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \\ -1 \end{pmatrix} \right\}$$

mentre una base di W sia data dai vettori

$$\mathcal{B}' = \left\{ \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \\ 2 \\ 1 \end{pmatrix} \right\}.$$

Applicando l'algoritmo di estrazione a $\mathcal{B} \cup \mathcal{B}'$ scartiamo il vettore $\begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}$ in quanto

$$\begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 1 \\ -1 \\ 0 \\ -1 \end{pmatrix} \tag{2.24}$$

e il vettore $\begin{pmatrix} 2 \\ 2 \\ 2 \\ 1 \end{pmatrix}$ in quanto

$$\begin{pmatrix} 2 \\ 2 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \frac{1}{2} \begin{pmatrix} 1 \\ -1 \\ 0 \\ -1 \end{pmatrix} \tag{2.25}$$

Ora se nelle uguaglianze (2.24) e (2.25) raggruppiamo nel primo membro tutti gli addendi corrispondenti a elementi di $\mathcal{B}'$ e nel secondo membro tutti gli addendi corrispondenti a elementi di $\mathcal{B}$ otteniamo le uguaglianze

$$\begin{aligned} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix} - \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} &= \tfrac{1}{2} \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix} + \tfrac{1}{2} \begin{pmatrix} 1 \\ -1 \\ 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \\ \begin{pmatrix} 2 \\ 2 \\ 2 \\ 1 \end{pmatrix} - \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} &= \tfrac{1}{2} \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \tfrac{1}{2} \begin{pmatrix} 1 \\ -1 \\ 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 1 \\ 1 \end{pmatrix} \end{aligned}$$

Come abbiamo scritto i vettori $\begin{pmatrix}1\\0\\0\\0\end{pmatrix}$ e $\begin{pmatrix}2\\1\\1\\1\end{pmatrix}$ sia come combinazione lineare di elementi di $\mathcal{B}$ sia come combinazione lineare di elementi di $\mathcal{B}'$. Dunque essi appartengono a $U \cap W$.

Osservazione 2.44. Vogliamo insistere sul fatto che i vettori che abbiamo ottenuto in $U \cap W$ **NON** sono i vettori della lista $\mathcal{B} \cup \mathcal{B}'$ che abbiamo scartato.

Consideriamo ora il caso generale in cui

$$\mathcal{B} = \{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_h\}$$

e

$$\mathcal{B}' = \{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_k\}.$$

Applico l'algoritmo di estrazione a $\mathcal{B} \cup \mathcal{B}' = \{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_h, \boldsymbol{w}_1, \ldots, \boldsymbol{w}_k\}$.

Osserviamo che nessuno dei vettori $\boldsymbol{u}_i$ ̀combinazione lineare dei precedenti per cui nessuno di tali vettori verrà scartato nel processo algoritmico.

Supponiamo che uno dei $\boldsymbol{w}_j$ venga scartato. Allora lui sarà combinazione lineare dei precedenti $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_h, \boldsymbol{w}_1, \ldots, \boldsymbol{w}_{j-1}$, ovvero potrò scrivere $\boldsymbol{w}_j$ nel seguente modo

$$\boldsymbol{w}_j = \lambda_1 \boldsymbol{u}_1 + \ldots + \lambda_h \boldsymbol{u}_h + \mu_1 \boldsymbol{w}_1 + \ldots + \mu_{j-1} \boldsymbol{w}_{j-1}.$$

Riordinando i termini di questa uguaglianza in modo che tutti i vettori della lista $\mathcal{B}$ si trovino a destra a e tutti i vettori della lista $\mathcal{B}'$ si trovino a sinistra otteniamo

$$\boldsymbol{w}_j - \mu_1 \boldsymbol{w}_1 - \ldots - \mu_{j-1} \boldsymbol{w}_{j-1} = \lambda_1 \boldsymbol{u}_1 + \ldots + \lambda_h \boldsymbol{u}_h$$

I due membri di questa uguaglianza sono rispettivamente una combinazione lineare dei vettori di $\mathcal{B}'$ e una combinazione lineare dei vettori di $\mathcal{B}$ che di conseguenza producono lo stesso vettore $\boldsymbol{v}$. Tale vettore è dunque rappresentabile sia come combinazione lineare di $\mathcal{B}$ che come combinazione lineare di $\mathcal{B}'$ e dunque appartiene a $U \cap W$

Ora applicando questo processo ogni volta che scarto un vettore nella lista $\mathcal{B} \cup \mathcal{B}'$ ottengo una certa lista di vettori di $U \cap W$.

$$\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m\}$$

(dove ancora una volta rimarchiamo che in generale i vettori $\boldsymbol{v}_j$ non compaiono nella lista $\mathcal{B} \cup \mathcal{B}'$).

Vale la seguente proprietà che non dimostreremo

Proposizione 2.29. *La lista* $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m\}$ *è una base di* $U \cap W$.

CAPITOLO 3

Matrici

1. Lo spazio delle matrici $k \times n$

Definizione 3.1. Una **matrice** A è una tabella rettangolare di numeri, detti **entrate** (o **elementi**, o anche, talora, coefficienti) della matrice, diciamo che A ha **ordine** $k \times n$ se ha k righe e n colonne.
Una matrice si dice **quadrata** se il numero di righe è uguale al numero di colonne.
Una matrice $1 \times n$ (con una sola riga) è detta **vettore riga**, una matrice $k \times 1$ (con una sola colonna) è detta **vettore colonna**.

Esempio 3.1.
Ecco un esempio di una matrice 3×4:

$$A = \begin{pmatrix} 1 & 2 & -1 & 24 \\ 7 & 18 & 15 & -19 \\ -1 & 0 & -1 & 0 \end{pmatrix}.$$

Questa è una matrice quadrata 3×3:

$$B = \begin{pmatrix} 1 & -1 & 32 \\ 72 & -10 & 77 \\ 9 & 1 & 10 \end{pmatrix}.$$

Questa è una matrice 1×4, quindi un vettore riga:

$$\begin{pmatrix} 1 & 2 & -1 & 17 \end{pmatrix},$$

mentre questa è una matrice 5×1, quindi un vettore colonna: $\begin{pmatrix} 1 \\ -1 \\ 10 \\ 57 \\ 11 \end{pmatrix}$.

Si noti che due matrici hanno lo stesso ordine se hanno sia lo stesso numero di righe che lo stesso numero di colonne. Ad esempio le matrici seguenti hanno ordine diverso:

$$\begin{pmatrix} 1 & -1 & 32 & 7 & 4 \\ 72 & -10 & 77 & 15 & 17 \\ 9 & 1 & 10 & 16 & 11 \end{pmatrix} \qquad \begin{pmatrix} 1 & 10 & 3 \\ 7 & -13 & 7 \\ 9 & 1 & 10 \\ -9 & -1 & -7 \\ 3 & -3 & 15 \end{pmatrix}.$$

Si osservi che il loro ordine è rispettivamente 3×5 e 5×3.

Osservazione 3.2. Una matrice 1×1 è formata da un solo numero (una riga e una colonna), identificheremo la matrice col numero stesso.
I vettori colonna $k \times 1$ possono essere identificati con i vettori di $\mathbb{R}^k$, visti nel capitolo 2.

Le matrici sono normalmente indicate con lettere maiuscole. Osserviamo che gli elementi di $\mathbb{R}^k$ possono essere denotati sia in grassetto (come normalmente denotiamo i vettori) che con una lettera maiuscola (pensandoli come matrici colonna). Nel seguito utilizzeremo entrambe le notazioni in maniera del tutto equivalente.

Per i vettori di $\mathbb{R}^n$, c'è una chiara numerazione delle entrate. Per cui, ad esempio, la terza entrata del vettore

$$\boldsymbol{w} = \begin{pmatrix} 1 \\ 7 \\ -5 \\ 3 \\ 10 \\ 15 \end{pmatrix}$$

è il numero $w_3 = -5$.

Per individuare un'entrata di una matrice c'è bisogno di una doppia numerazione: bisogna infatti specificare a quale riga e a quale colonna appartiene. Ad esempio se consideriamo la matrice

$$A = \begin{pmatrix} 1 & -1 & 32 & 7 & 4 \\ 72 & -10 & 77 & 15 & 17 \\ 9 & 1 & 10 & 16 & 11 \end{pmatrix}$$

l'entrata in posizione $(2,3)$ viene individuata considerando il numero sulla seconda riga della terza colonna. In questo caso scriviamo

$$a_{2\,3} = 77$$

In generale:

- a_{ij} è l'entrata della matrice A che si trova sulla riga i e colonna j.

Le posizioni delle entrate di una matrice sono dunque individuate da due **indici**. Il primo viene detto **indice di riga** e indica la riga in cui l'entrata è posta. Il secondo numero è detto **indice di colonna** e indica la colonna in cui l'entrata è posta. Osserviamo che se la matrice ha ordine $k \times n$, l'indice di riga è un numero intero compreso tra 1 e k, mentre l'indice di colonna è un numero intero compreso tra 1 e n. In altre parole le entrate della matrice sono tutte della forma a_{ij} dove $i = 1, \ldots, k$ e $j = 1, \ldots, n$.
Dunque la matrice generica $k \times n$ può essere scritta nel seguente modo

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & a_{k3} & \dots & a_{kn} \end{pmatrix},$$

oppure $A = (a_{ij}), \quad i = 1, \ldots k \; j = 1, \ldots, n.$

1.1. Righe e colonne di una matrice.

NOTAZIONE 3.2.

- Data una matrice A di ordine $k \times n$ denotiamo con $A^1, A^2, \ldots, A^n$ le sue colonne e scriviamo
$$A = (A^1|A^2|\ldots|A^n)\,.$$
- Data una matrice A di ordine $k \times n$ denotiamo con $A_1, \ldots, A_k$ le sue righe e scriviamo:
$$A = \begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_k \end{pmatrix}\,.$$

Ad esempio se consideriamo la matrice 3×4
$$A = \begin{pmatrix} 8 & 7 & 5 & 18 \\ 11 & -3 & -65 & 36 \\ 3 & 8 & 9 & 51 \end{pmatrix}$$
avremo
$$A^1 = \begin{pmatrix} 8 \\ 11 \\ 3 \end{pmatrix}\,,\ A^2 = \begin{pmatrix} 7 \\ -3 \\ 8 \end{pmatrix}\,,\ A^3 = \begin{pmatrix} 5 \\ -65 \\ 9 \end{pmatrix}\,,\ A^4 = \begin{pmatrix} 18 \\ 36 \\ 51 \end{pmatrix}\,.$$
Osserviamo che le colonne di una matrice A di ordine $k \times n$ sono vettori colonna $k \times 1$, contenenti esattamente k numeri (uno per ciascuna riga), dunque sono vettori di $\mathbb{R}^k$.
La componente i del vettore A^1 è l'entrata sulla riga i della prima colonna di A e dunque $(A^1)_i = a_{i1}$. In generale si verifica che la componente i-esima del vettore A^j è
$$(A^j)_i = a_{ij}\,.$$
Inoltre, se consideriamo l'esempio precedente abbiamo
$$A_1 = (8, 7, 5, 18)\,,\ A_2 = (11, -3, -65, 36),\ A_3 = (3, 8, 9, 51)\,.$$
Osserviamo che le righe di A sono vettori riga $1 \times n$, contenenti ciascuno n entrate (una per ogni colonna).
La componente j-esima della prima riga è l'entrata sulla colonna j riga 1. Dunque si ha $(A_1)_j = a_{1j}$. Analogamente la componente j-esima della riga A_i è
$$(A_i)_j = a_{ij}\,.$$

1.2. Struttura di spazio vettoriale sullo spazio delle matrici $k \times n$.

Definizione 3.3. Dati una matrice A di tipo $k \times n$ ed un numero reale λ, si definisce la matrice λA come la matrice ottenuta moltiplicando tutte le colonne di A per il numero λ.
Se $A = (A^1|A^2|\cdots|A^n)$, allora $\lambda A = (\lambda A^1|\lambda A^2|\cdots|\lambda A^n)$.

Osserviamo che:
(1) λA è una matrice di ordine $k \times n$, come la matrice A;
(2) λA si ottiene moltiplicando tutte le entrate di A per il numero λ, cioè l'entrata di posto (i,j) della matrice λA è λa_{ij}.

Ad esempio abbiamo:
$$3\begin{pmatrix} 1 & -1 & 32 & 7 & 4 \\ 72 & -10 & 77 & 15 & 17 \\ 9 & 1 & 10 & 16 & 11 \end{pmatrix} = \begin{pmatrix} 3 & -3 & 96 & 21 & 12 \\ 216 & -30 & 231 & 45 & 51 \\ 27 & 3 & 30 & 48 & 33 \end{pmatrix}$$

Definizione 3.4. Date due matrici A, B **dello stesso ordine**, $k \times n$, la matrice somma $A+B$ è definita come la matrice ottenuta sommando le colonne di A con le rispettive colonne di B.
Se $A = (A^1|A^2|\ldots|A^n)$ e $B = (B^1|B^2|\ldots|B^n)$, allora
$$A + B = (A^1 + B^1|A^2 + B^2|\ldots|A^n + B^n).$$

Osserviamo che:
(1) $A + B$ è una matrice di ordine $k \times n$, come le matrici A e B;
(2) $A + B$ si ottiene sommando tutte le entrate di A con le entrate di B posto per posto, cioè l'entrata di posto (i,j) della matrice $A + B$ è $a_{ij} + b_{ij}$.

Ad esempio abbiamo:
$$\begin{pmatrix} 1 & -1 & 30 \\ 10 & 15 & -3 \\ 11 & 7 & -2 \end{pmatrix} + \begin{pmatrix} 7 & 1 & 3 \\ 5 & 9 & 4 \\ -3 & 6 & 8 \end{pmatrix} = \begin{pmatrix} 8 & 0 & 33 \\ 15 & 24 & 1 \\ 8 & 13 & 6 \end{pmatrix}.$$

Nota bene: Due matrici di ordini diversi **NON** possono essere sommate.

Definizione 3.5. Indichiamo con
$$\mathcal{M}_{\mathbb{R}}(k,n) = \{\text{matrici di ordine } k \times n \text{ ed entrate reali}\}$$
l'insieme di tutte le matrici di ordine $k \times n$ ad elementi in $\mathbb{R}$.

Nel caso di matrici quadrate semplificheremo la notazione indicando con $\mathcal{M}_{\mathbb{R}}(n)$ l'insieme delle matrici $n \times n$.

Nell'insieme $\mathcal{M}_{\mathbb{R}}(k, n)$ abbiamo definito due operazioni:
(1) la somma di matrici,
(2) la moltiplicazione di una matrice per un numero reale.

È semplice verificare che tali operazioni verificano tutte le proprietà (2.3) ... (2.3) e (2.3) ... (2.3) e enunciate nella Definizione 2.3 del Capitolo 2.

In particolare l'elemento neutro della somma –detta **matrice nulla** e denotata con $O_{k\times n}$– è la matrice $k \times n$ le cui entrate sono tutte nulle.

Proposizione 3.1. *$\mathcal{M}_{\mathbb{R}}(k, n)$ è uno spazio vettoriale su $\mathbb{R}$.*

Lasciamo la verifica di questa proposizione allo studente interessato.

Ci preoccupiamo ora di studiare la dimensione dello spazio $\mathcal{M}_{\mathbb{R}}(k, n)$. Consideriamo le matrici di $\mathcal{M}_{\mathbb{R}}(k, n)$ le cui entrate sono tutte nulle tranne una che è uguale a 1. In particolare, fissata una posizione (i, j), denotiamo con $E_{i\,j}$ la matrice composta da tutti 0 tranne 1 nella posizione (i, j).

Ad esempio nel caso $k = 3$, $n = 2$ abbiamo

$$
\begin{array}{cc}
E_{1\,1} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} & E_{1\,2} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \\
E_{2\,1} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \end{pmatrix} & E_{2\,2} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \\
E_{3\,1} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \end{pmatrix} & E_{3\,2} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix}
\end{array}
$$

Consideriamo ora una combinazione lineare di tutte le matrici $E_{i\,j}$. Ad esempio ne caso $k = 2, n = 3$ consideriamo la combinazione lineare

$$
\begin{aligned}
10 \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} + 17 \begin{pmatrix} 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} + 2 \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \end{pmatrix} - 8 \begin{pmatrix} 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \\
+ 7 \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \end{pmatrix} + 15 \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 10 & 17 \\ 2 & -8 \\ 7 & 15 \end{pmatrix}.
\end{aligned}
$$

Osserviamo che i risultato di questa combinazione è la matrice le cui entrate sono proprio i coefficienti della combinazione lineare. Più precisamente l'entrata in posizione (i, j) del risultato è uguale al coefficiente che moltiplica la matrice $E_{i\,j}$.

Questo è un fatto generale. Ovvero ogni combinazione lineare delle matrici $E_{i\,j}$ di $\mathcal{M}_{\mathbb{R}}(k, n)$

$$\lambda_{1\,1}E_{1\,1} + \lambda_{2\,1}E_{2\,1} + \ldots + \lambda_{12}E_{1\,2} + \ldots + \lambda_{2\,n}E_{2\,n} + \ldots + \lambda_{k\,1}E_{n\,1} + \ldots \lambda_{k\,n}E_{k\,n}$$

dà come risultato la matrice A la cui entrata $a_{i\,j}$ coincide con il coefficiente $\lambda_{i,j}$.

Ciò mostra che l'unica combinazione lineare delle matrici $E_{i\,j}$ che produce la matrice nulla è la combinazione banale (ovvero corrisponde alla scelta di tutti i coefficienti della combinazione nulli). Segue che le matrici $E_{i\,j}$ sono indipendenti.

Inoltre è possibile esprimere una qualsiasi matrice A di ordine $k \times n$ come combinazione delle matrici E_{ij} fissando i coefficienti λ_{ij} della combinazione uguale all'entrata a_{ij} della matrice. Possiamo quindi concludere con la seguente proprietà:

Proposizione 3.2. *Le matrici E_{ij} formano una base dello spazio $M_{\mathbb{R}}(k,n)$.*
In particolare $\mathcal{M}_{\mathbb{R}}(k,n)$ è uno spazio vettoriale finitamente generato di dimensione kn.

2. Moltiplicazione tra matrici

2.1. Moltiplicazione matrice-vettore.

Prima di definire la moltiplicazione tra matrici definiamo il prodotto tra matrice e vettore colonna.

Definizione 3.6. Siano $A \in \mathcal{M}_{\mathbb{R}}(k,n)$ e $X \in \mathbb{R}^n$.

Il prodotto AX è per definizione la combinazione lineare delle colonne di A i cui coefficienti sono le entrate di X. In simboli:

$$AX = (A^1|A^2|\ldots|A^n)\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = x_1A^1 + x_2A^2 + \ldots + x_nA^n.$$

Osserviamo che:

(1) Il prodotto tra una matrice A e un vettore colonna X è definito solo quando **il numero di colonne di A è uguale al numero di entrate di X**.
(2) Nel caso in cui si possa fare il prodotto tra la matrice A e il vettore X, **il risultato AX è un vettore colonna in $\mathbb{R}^k$ dove k è il numero di righe di A**.

Esempio 3.3.

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

In questo caso abbiamo

$$AX = 2\begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} + 1\begin{pmatrix} -1 \\ -5 \\ 1 \end{pmatrix} + 0\begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} - 1\begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} -2 \\ -3 \\ -2 \end{pmatrix}$$

Elenchiamo le principali proprietà del prodotto matrice per vettore.

PROPRIETÀ 3.3.

(1) **Proprietà distributiva rispetto alla somma di matrici**. Per ogni coppia di matrici A, B di ordine $k \times n$ e per ogni $X \in \mathbb{R}^n$ vale:
$$(A + B)X = AX + BX.$$

Infatti da una parte abbiamo che
$$(A + B)X = x_1(A + B)^1 + x_2(A + B)^2 + \ldots + x_n(A + B)^n .$$
Poiché le colonne di $A+B$ sono la somma delle corrispondenti colonne di A e di B possiamo concludere
$$(A + B)X = x_1(A^1 + B^1) + x_2(A^2 + B^2) + \ldots + x_n(A^n + B^n) . \tag{3.1}$$
Del resto si ha che
$$\begin{aligned} AX + BX = &(x_1A^1 + x_2A^2 + \ldots + x_nA^n) \\ &+ (x_1B^1 + x_2B^2 + \ldots + x_nB^n) \end{aligned}$$
e riordinando i termini otteniamo
$$AX + BX = x_1(A^1 + B^1) + x_2(A^2 + B^2) + \ldots + x_n(A^n + B^n) . \tag{3.2}$$
Confrontando il secondo membro della (3.1) con il secondo membro della (3.2) si ottiene l'uguaglianza cercata.

(2) **Proprietà distributiva rispetto alla somma di vettori**. Per ogni matrice A di ordine $k \times n$ e per ogni $X, Y \in \mathbb{R}^n$ vale:
$$A(X + Y) = AX + AY.$$
Infatti da una parte abbiamo
$$A(X + Y) = (x_1 + y_1)A^1 + (x_2 + y_2)A^2 + \ldots + (x_n + y_n)A^n . \tag{3.3}$$
Dall'altra si ha
$$\begin{aligned} AX + AY = &(x_1A^1 + x_2A^2 \ldots + x_nA^n) \\ &+ (y_1A^1 + y_2A^2 + \ldots + y_nA^n) \end{aligned} \tag{3.4}$$
e confrontando la (3.3) con la (3.4) si ha l'uguaglianza cercata.

(3) **Omogeneità**. Per ogni matrice A di ordine $k \times n$, per ogni $X \in \mathbb{R}^n$ e per ogni λ numero reale si ha:
$$A(\lambda X) = (\lambda A)X = \lambda(AX).$$
Infatti osserviamo che il primo membro è uguale a
$$(\lambda x_1)A^1 + (\lambda x_2)A^2 + \ldots + (\lambda x_n)A^n$$
il secondo membro è uguale a
$$x_1\lambda A^1 + x_2\lambda A^2 + \ldots + x_n\lambda A^n$$

mentre infine il terzo membro è uguale a

$$\lambda(x_1 A^1 + x_2 A^2 + \ldots + x_n A^n)\,.$$

Confrontando queste espressioni l'uguaglianza cercata segue facilmente.

(4) Per ogni matrice A di ordine $k \times n$, si ha:

$$A\mathbf{0}_n = \mathbf{0}_k.$$

Infatti $A\mathbf{0}_n$ è per definizione la combinazione banale delle colonne di A.

(5) Per ogni $X \in \mathbb{R}^n$, si ha:

$$O_{k\times n}X = \mathbf{0}_k.$$

Infatti le colonne di $O_{k\times n}$ coincidono tutte con il vettore nullo di $\mathbb{R}^k$.

(6) Sia $\boldsymbol{e}_j$ il vettore j-esimo della base canonica di $\mathbb{R}^n$. Per ogni matrice A di ordine $k \times n$ si ha:

$$A\boldsymbol{e}_j = A^j\,.$$

Osserviamo che il prodotto di A con $\boldsymbol{e}_1$ è dato da:

$$(A^1|A^2|\ldots|A^n)\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix} = 1A^1 + 0A^2 + \ldots + 0A^n = A^1$$

ovvero la prima colonna di A.

(7) Per ogni vettore riga A di ordine $1 \times n$ e per ogni $X \in \mathbb{R}^n$, il prodotto AX è un numero reale.

Esempio 3.4. $A = (1, -1, 0, 3)$ e $X = \begin{pmatrix}2\\1\\0\\-1\end{pmatrix}$. In questo caso

$$AX = 2 \cdot 1 + 1 \cdot -1 + 0 \cdot 0 + (-1)3 = -2\,.$$

Osservazione 3.5. Ora cerchiamo di capire nel caso generale come calcolare la prima componente del prodotto AX. Osserviamo che si ha

$$AX = x_1A^1 + \ldots + x_nA^n = x_1\begin{pmatrix}a_{1\,1}\\ *\\ \vdots\\ *\end{pmatrix} + x_2\begin{pmatrix}a_{1\,2}\\ *\\ \vdots\\ *\end{pmatrix} + x_3\begin{pmatrix}a_{1\,3}\\ *\\ \vdots\\ *\end{pmatrix} + \ldots + x_n\begin{pmatrix}a_{1\,n}\\ *\\ \vdots\\ *\end{pmatrix}$$

per cui la prima entrata del vettore AX si ottiene sommando i prodotti delle entrate della prima riga di A con le corrispondenti entrate del vettore X. Ovvero

$$(AX)_1 = A_1X\,.$$

Analogamente la seconda entrata di AX si ottiene moltiplicando la seconda riga di A per X e così via. Sinteticamente abbiamo

$$AX = \begin{pmatrix} A_1X \\ A_2X \\ \vdots \\ A_kX \end{pmatrix}.$$

2.2. Prodotto tra matrici.

Definizione 3.7. Siano A una matrice di ordine $k \times n$ e B una matrice di ordine $n \times h$. Definiamo il prodotto

$$AB = (AB^1|AB^2|\ldots|AB^h),$$

dove B^1,...,B^h sono le colonne di B.

Nota bene: Ogni colonna B^i della matrice B può essere considerata un vettore di $\mathbb{R}^n$, e dunque posso moltiplicare tale colonna per la matrice A. Il prodotto AB^i è quindi un vettore di $\mathbb{R}^k$, quindi la matrice AB è una matrice di ordine $k \times h$.

Osservazione 3.6. Date due matrici A, B posso fare il prodotto AB se e solo se **il numero di colonne di A è uguale al numero di righe di B**. Inoltre il numero di righe di AB è uguale al numero di righe di A mentre il numero di colonne di AB è uguale al numero di colonne di B.

Ovvero a livello di ordini si ha

$$\begin{array}{cccc} A & B & = & AB \\ k \times n & n \times h & \to & k \times h \end{array}.$$

Si noti che se A è un vettore riga allora anche AB è un vettore riga con un numero di elementi pari al numero di colonne di B.

Invece se B è un vettore colonna allora anche AB è un vettore colonna con un numero di elementi pari al numero di righe di A.

Esempio 3.7. Consideriamo le due matrici

$$A = \begin{pmatrix} 1 & -2 \\ 7 & 3 \\ 3 & -2 \end{pmatrix} \qquad B = \begin{pmatrix} 1 & 3 & 4 & -1 \\ 1 & -1 & 2 & -2 \end{pmatrix}$$

In questo caso il numero di colonne di A è uguale al numero di righe di B e dunque si può considerare il prodotto AB. Inoltre l'ordine di AB è uguale a 3×4.

Calcolando otteniamo

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

$$\Bigg(\;1\begin{pmatrix}1\\7\\3\end{pmatrix}+1\begin{pmatrix}-2\\3\\-2\end{pmatrix}\;\Bigg|\;3\begin{pmatrix}1\\7\\3\end{pmatrix}-1\begin{pmatrix}-2\\3\\-2\end{pmatrix}\;\Bigg|$$

$$\Bigg|\;4\begin{pmatrix}1\\7\\3\end{pmatrix}+2\begin{pmatrix}-2\\3\\-2\end{pmatrix}\;\Bigg|\;-1\begin{pmatrix}1\\7\\3\end{pmatrix}-2\begin{pmatrix}-2\\3\\-2\end{pmatrix}\;\Bigg)=$$

$$\begin{pmatrix}-1&5&0&3\\10&18&34&-13\\1&11&8&1\end{pmatrix}.$$

Osservazione 3.8.

(1) L'elemento (i,j) della matrice AB e' l'i-esimo elemento della colonna j. La colonna j è il risultato del prodotto AB^j. Da quanto visto l'i-esimo elemento di tale prodotto si ottiene moltiplicando la riga i-esima di A per il vettore B^i. Ovvero abbiamo

$$(AB)_{ij} = A_iB^j \tag{3.5}$$

Per questo motivo la moltiplicazione di matrici è a volte detta **moltiplicazione riga per colonna**.

(2) La prima riga della matrice AB si ottiene moltiplicando la prima riga di A per la matrice B ovvero si ha

$$(AB)_1 = A_1B\,.$$

Infatti per la (3.5) abbiamo

$$(AB)_1 = (A_1B^1|A_1B^2|\dots|A_1B^h)$$

del resto si ha

$$A_1B = (A_1B^1|A_1B^2|\dots|A_1B^h)\,.$$

Analogamente si può mostrare che la seconda riga di AB si ottiene moltiplicando la seconda riga di A per la matrice B. In definitiva si ha:

$$AB=\begin{pmatrix}A_1B\\A_2B\\\vdots\\A_kB\end{pmatrix}.$$

2.3. Proprietà del prodotto tra matrici.

Enunciamo le principali proprietà del prodotto tra matrici:

PROPRIETÀ 3.4.

(1) **Proprietà distributiva rispetto al primo fattore**. Per ogni coppia di matrici A, B di ordine $k \times n$ e per ogni matrice C di odine $n \times h$ si ha:
$$(A + B)C = AC + BC.$$
Osserviamo che la prima colonna di $(A+B)C$ è uguale a $(A+B)C^1$. Per la proprietà distributiva del prodotto matrice per vettore rispetto alla somma di matrici abbiamo che
$$((A + B)C)^1 = AC^1 + BC^1$$
il che mostra che la prima colonna di $(A + B)C$ è uguale alla prima colonna di $AC + BC$. Ragionando in questo modo si mostra che le colonne delle matrici $(A+B)C$ e $AC+BC$ sono a due a due uguali e dunque le due matrici coincidono.

(2) **Proprietà distributiva rispetto al secondo fattore** . Per ogni matrice A di ordine $k \times n$ e per ogni coppia di matrici B e C di ordine $n \times h$, si ha:
$$A(B + C) = AB + AC.$$
Confrontiamo la prima colonna di $A(B+C)$ con la prima colonna di $AB + AC$. Risulta che esse sono uguali rispettivamente a $A(B^1 + C^1)$ e $AB^1 + AC^1$. Per la proprietà distributiva del prodotto matrice per vettore rispetto alla somma i vettori si conclude che le prime due colonne sono uguali. Ragionando in questo modo si mostra che le colonne di $A(B + C)$ sono uguali alle colonne di $AB + AC$ e dunque le due matrici sono uguali.

(3) **Omogeneità del prodotto** . Per ogni matrice A di ordine $k \times n$, per ogni matrice B di ordine $n \times h$ e per ogni numero reale λ, si ha:
$$(\lambda A)B = A(\lambda B) = \lambda(AB).$$
Infatti confrontando le colonne di queste matrici, esse risultano uguali per l'omogeneità del prodotto matrice per vettore.

(4) **Proprietà associativa del prodotto.** Se A, B e C sono matrici tali che si possa fare il prodotto $(AB)C$, allora è possibile anche fare il prodotto $A(BC)$ e inoltre si ha
$$A(BC) = (AB)C.$$
Mostriamo che se è possibile fare il prodotto $(AB)C$ allora posso anche fare il prodotto $A(BC)$ e gli ordini dei risultati sono uguali. Osserviamo che se A è di ordine $k \times n$ B sarà del ordine $n \times h$ e il

prodotto AB è di ordine $k \times h$. Siccome posso produrre AB con C allora C è di ordine $h \times p$ e il risultato $(AB)C$ è di ordine $k \times p$. Dunque in definitiva abbiamo

$$\begin{array}{cccccc} (& A & B &) & C & \\ & k \times n & n \times h & & h \times p & \to k \times p \end{array}$$

Osserviamo dunque che possiamo fare il prodotto (BC) che sarà di ordine $n \times p$ e in definitiva è possibile fare il prodotto $A(BC)$ e il risultato è sempre di ordine $k \times p$.

$$\begin{array}{cccccc} A & (& B & C &) & \\ k \times n & & n \times h & h \times p & & \to k \times p. \end{array}$$

Verifica della proprietà associativa.

- Prima verificheremo l'uguaglianza nel caso particolare che C sia un vettore della base canonica di $\mathbb{R}^h$.
- In un secondo momento la verificheremo nel caso in cui C sia un qualsiasi vettore colonna di $\mathbb{R}^h$ (ovvero $p = 1$).
- Infine lo verificheremo nel caso generale.

Caso 1: Assumiamo che C sia un vettore $\boldsymbol{e}_i$ della base canonica di $\mathbb{R}^h$. Allora si ha $(AB)\boldsymbol{e}_i = (AB)^i = AB^i$. Del resto $A(B\boldsymbol{e}_i) = AB^i$ e dunque ricaviamo

$$(AB)\boldsymbol{e}_i = A(B\boldsymbol{e}_i)\,. \tag{3.6}$$

Caso 2: Assumiamo che C sia un vettore colonna X di $\mathbb{R}^h$. Considerando la decomposizione di X

$$X = x_1\boldsymbol{e}_1 + \ldots + x_n\boldsymbol{e}_n$$

si ottiene

$$\begin{aligned} (AB)X &= (AB)(x_1\boldsymbol{e}_1 + \ldots + x_n\boldsymbol{e}_n) \\ &= (AB)(x_1\boldsymbol{e}_1) + (AB)(x_2\boldsymbol{e}_2) + \ldots + (AB)(x_n\boldsymbol{e}_n) \end{aligned}$$

dove l'ultima uguaglianza si ottiene applicando la proprietà distributiva. Per l'omogeneità del prodotto abbiamo ancora che

$$(AB)X = x_1(AB)\boldsymbol{e}_1 + x_2(AB)\boldsymbol{e}_2 + \ldots + x_n(AB)\boldsymbol{e}_n\,.$$

Analogamente si calcola che

$$A(BX) = x_1A(B\boldsymbol{e}_1) + x_2A(B\boldsymbol{e}_2) + \ldots + x_nA(B\boldsymbol{e}_n)$$

e per la (3.6) si conclude che

$$(AB)X = A(BX)\,. \tag{3.7}$$

Caso 3: Supponiamo ora che C sia una qualsiasi matrice $h \times p$. Allora la prima colonna della matrice $(AB)C$ è il vettore $(AB)C^1$. Del resto la prima colonna di $A(BC)$ è uguale a $A(BC)^1 = A(BC^1)$. Per la (3.7), la prima colonna di $(AB)C$ e la prima colonna di $A(BC)$ coincidono. Ragionando in maniera analoga si mostra che tutte le colonne di $(AB)C$ coincidono con le corrispondenti colonne di $A(BC)$ e dunque le due matrici sono uguali.

Esiste anche un modo molto compatto per verificare la proprietà associativa: occorre scrivere il prodotto righe per colonne fra matrici usando una notazione sintetica di sommatoria, si veda l'appendice 8 alla fine del capitolo.

NOTA BENE: Il prodotto tra matrici **NON** è commutativo. Infatti:

- Anzitutto, in generale se è possibile fare AB non è detto che si possa fare BA. Ad esempio se A è 2×3 e B è 3×4 è possibile fare AB ma non è possibile fare BA.
- Anche quando è possibile fare AB e BA non è detto che queste due matrici abbiano lo stesso ordine. Ad esempio se A è 2×3 e B è 3×2 si ha che AB è 2×2 mentre BA è 3×3.
- Anche quando sia possibile fare i prodotti AB e BA e i risultati abbiano lo stesso ordine non è detto che $AB = BA$. Ad esempio abbiamo
$$\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 2 \end{pmatrix},$$
mentre si ha
$$\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix}.$$

2.4. Matrice identità.

Definizione 3.8. La **matrice identità** $n \times n$ è la matrice quadrata $n \times n$ le cui colonne sono i vettori della base canonica di $\mathbb{R}^n$:
$$I_n = (\boldsymbol{e}_1|\boldsymbol{e}_2|\dots|\boldsymbol{e}_n) = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & & & \vdots \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & & 1 & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix}$$

Tale matrice ha alcune proprietà notevoli che vogliamo mettere in evidenza:

(1) Per ogni $X \in \mathbb{R}^n$ si ha:
$$I_n X = X\,.$$
Infatti abbiamo
$$I_n X = (\boldsymbol{e}_1|\boldsymbol{e}_2|\dots|\boldsymbol{e}_n)X = x_1\boldsymbol{e}_1 + x_2\boldsymbol{e}_2 + \dots + x_n\boldsymbol{e}_n = X$$

(2) Per ogni matrice A di ordine $n \times h$ si ha:
$$I_n A = A\,.$$
Infatti $I_n A = (I_n A^1|I_n A^2|\dots|I_n A^h)$ e per il punto precedente le colonne di $I_n A$ sono uguali ad $A^1, A^2, \dots, A^h$. Per cui in definitiva $I_n A = A$.

(3) Per ogni matrice A di ordine $k \times n$ si ha:
$$AI_n = A\,.$$
Infatti $AI_n = A(\boldsymbol{e}_1|\boldsymbol{e}_2|\dots|\boldsymbol{e}_n) = (A\boldsymbol{e}_1|A\boldsymbol{e}_2|\dots|A\boldsymbol{e}_n)$ e poiché $A\boldsymbol{e}_1 = A^1$, $A\boldsymbol{e}_2 = A^2$ e così via si conclude che $AI_n = A$.

3. Il prodotto tra matrici quadrate e l'invertibilità

Osserviamo che se A e B sono matrici quadrate $n \times n$ allora si può fare il prodotto tra A e B e il risultato AB è ancora una matrice $n \times n$. In altre parole il prodotto tra matrici è un'*operazione interna* di $\mathcal{M}_{\mathbb{R}}(n)$.

Vogliamo studiarne le principali proprietà confrontandole con le proprietà del prodotto tra numeri. Abbiamo già osservato che questa operazione è **associativa** e **distributiva** rispetto alla somma, mentre non è commutativa. Inoltre osserviamo che la matrice identità I_n è **l'elemento neutro**, infatti dalla precedente sezione risulta che per ogni matrice $A \in \mathcal{M}_{\mathbb{R}}(n)$:

$$I_n A = A I_n = A.$$

Questo ci consente di affermare la seguente

PROPOSIZIONE 3.5. *L'insieme $\mathcal{M}_{\mathbb{R}}(n)$ delle matrici quadrate reali di ordine n con le operazioni interne di addizione e moltiplicazione è un anello unitario (non commutativo).*

Scopriremo in seguito che **NON** è integro

L'inverso di un numero a è un numero $\frac{1}{a}$ caratterizzato dalla proprietà che $a \cdot \frac{1}{a} = 1$. Possiamo allora definire l'inverso di una matrice in modo del tutto simile.

DEFINIZIONE 3.9. Sia $A \in \mathcal{M}_{\mathbb{R}}(n)$, un'**inversa** per A è una matrice $B \in \mathcal{M}_{\mathbb{R}}(n)$, tale che

$$AB = BA = I_n \,. \tag{3.8}$$

OSSERVAZIONE 3.9.

(1) Una matrice A non può ammettere due inverse distinte. Infatti supponiamo che B e C siano inverse di A, e calcoliamo in due modi diversi il prodotto BAC: da una parte $BAC = (BA)C = I_nC = C$ dall'altra $BAC = B(AC) = BI_n = B$ e dunque dovremmo avere $C = B$.
(2) Se una matrice A ammette un'inversa, questa inversa è unica e verrà denotata con A^{-1}.

Ci poniamo ora il seguente problema: quali matrici ammettono un'inversa? Come si calcola l'inversa?

DEFINIZIONE 3.10. Una matrice $A \in \mathcal{M}_{\mathbb{R}}(n)$ che ammette un'inversa si dice **invertibile**.

Come vedremo non tutte le matrici sono invertibili. Per mostrarlo verificheremo prima di tutto che le matrici invertibili soddisfano alcune interessanti proprietà.

Data una matrice $A \in \mathcal{M}_{\mathbb{R}}(n)$, possiamo considerare l'equazione vettoriale

$$AX = \mathbf{0}_n \,. \tag{3.9}$$

Le soluzioni di tale equazione sono i vettori di $\mathbb{R}^n$ che moltiplicati per la matrice A danno come risultato il vettore nullo. Chiaramente il vettore nullo è sempre

una soluzione dell'equazione (3.9), ma in generale non è l'unica soluzione. Ad esempio se
$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$
si verifica facilmente, per esempio, che anche $X = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ è soluzione (in realtà, lo sono tutti i vettori della forma $X = \begin{pmatrix} a \\ -a \end{pmatrix}$, con un qualunque numero a).

LEMMA 3.6. *Sia $A \in \mathcal{M}_{\mathbb{R}}(n)$ una matrice invertibile. Allora l'equazione vettoriale*
$$AX = \mathbf{0}_n$$
ammette come ***unica*** *soluzione $X = \mathbf{0}_n$.*

DIMOSTRAZIONE. Sia X una soluzione dell'equazione e calcoliamo $A^{-1}AX$ in due modi distinti. Da una parte osserviamo che
$$A^{-1}AX = A^{-1}(AX) = A^{-1}\mathbf{0}_n = \mathbf{0}_n.$$
Dall'altra si ha
$$A^{-1}AX = (A^{-1}A)X = I_nX = X.$$
Da cui otteniamo $X = \mathbf{0}_n$. □

Osserviamo subito che non può esistere l'inversa della matrice $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$. Infatti abbiamo visto che l'equazione $AX = \mathbf{0}_2$ ha soluzioni diverse dal vettore nullo.

Il fatto che la matrice A nell'esempio precedente non sia invertibile dipende dal fatto che le colonne di A sono dipendenti e dunque esistono combinazioni lineari non banali che danno come risultato il vettore nullo.

Infatti vale la seguente semplice osservazione

LEMMA 3.7. *Sia $A \in \mathcal{M}_{\mathbb{R}}(n)$. Se $X = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$ è soluzione dell'equazione $AX = \mathbf{0}_n$ allora la combinazione lineare delle colonne di A con coefficienti le componenti di X ha come i risultato il vettore nullo:*
$$\alpha_1 A^1 + \alpha_2 A^2 + \ldots + \alpha_n A^n = \mathbf{0}_n\,.$$

Viceversa data una combinazione lineare di $A^1, \ldots, A^n$ che produce il vettore nullo
$$\lambda_1 A^1 + \lambda_2 A^2 + \ldots + \lambda_n A^n = \mathbf{0}_n$$
allora il vettore dei coefficienti di tale combinazione è una soluzione dell'equazione $AX = \mathbf{0}_n$.

DIMOSTRAZIONE. La dimostrazione segue direttamente dell'osservazione che per definizione si ha
$$AX = x_1 A^1 + \ldots + x_n A^n\,.$$
□

Osservazione 3.10. Le colonne di una matrice sono linearmente indipendenti se e soltanto se l'unica soluzione dell'equazione $AX = \mathbf{0}_n$ è $X = \mathbf{0}_n$. Infatti abbiamo visto che la proprietà che caratterizza le liste di vettori indipendenti è che l'unica combinazione lineare che produce il vettore nullo è la combinazione banale (con tutti i coefficienti nulli).

Proposizione 3.8. *Sia $A \in \mathcal{M}_{\mathbb{R}}(n)$ una matrice invertibile, allora le colonne di A formano una base di $\mathbb{R}^n$.*

Dimostrazione. Per il Lemma 3.6 l'equazione $AX = \mathbf{0}_n$ ha come unica soluzione il vettore nullo $\mathbf{0}_n$. Per l'Osservazione 3.10 le colonne di A sono linearmente indipendenti. Dunque poiché sono esattamente n, esse formano una base di $\mathbb{R}^n$. $\square$

La Proposizione 3.8 ci dà una condizione necessaria affinché una matrice sia invertibile. Mostriamo ora che la condizione è anche sufficiente e che quindi l'insieme delle matrici invertibili coincide con l'insieme della matrici le cui colonne formano una base di $\mathbb{R}^n$.

Proposizione 3.9. *Sia $A \in \mathcal{M}_{\mathbb{R}}(n)$. Se le colonne di A formano una base $\mathcal{B}_A$ di $\mathbb{R}^n$ allora A è invertibile e le colonne di A^{-1} sono le coordinate degli elementi della base canonica rispetto alla base $\mathcal{B}_A$.*

Dimostrazione. Consideriamo la matrice quadrata B di ordine n la cui prima colonna è il vettore delle coordinate di $\boldsymbol{e}_1$ rispetto a $\mathcal{B}_A$, la cui seconda colonna è il vettore delle coordinate di $\boldsymbol{e}_2$ rispetto a $\mathcal{B}_A$ e così via:

$$B = ([\boldsymbol{e}_1]_{\mathcal{B}_A} \,|\, [\boldsymbol{e}_2]_{\mathcal{B}_A} \,|, \dots \,|\, [\boldsymbol{e}_n]_{\mathcal{B}_A}).$$

Allora si ha

$$AB = (A[\boldsymbol{e}_1]_{\mathcal{B}_A} \,|\, A[\boldsymbol{e}_2]_{\mathcal{B}_A} \,|\, \dots \,|\, A[\boldsymbol{e}_n]_{\mathcal{B}_A}).$$

Ora se le cordinate di $\boldsymbol{e}_1$ rispetto $\mathcal{B}_A$ sono $\lambda_1, \dots, \lambda_n$, si ha

$$\boldsymbol{e}_1 = \lambda_1 A^1 + \lambda_2 A^2 + \dots + \lambda_n A^n\,.$$

Osserviamo del resto che

$$A[\boldsymbol{e}_1]_{\mathcal{B}_A} = A \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} = \lambda_1 A^1 + \lambda_1 A^2 + \dots + \lambda_n A^n,$$

per cui deduciamo che $A[\boldsymbol{e}_1]_{\mathcal{B}_A} = \boldsymbol{e}_1$. Allo stesso modo si deduce che $A[\boldsymbol{e}_i]_{\mathcal{B}_A} = \boldsymbol{e}_i$, per $i = 2, \dots, n$. In particolare abbiamo che

$$AB = (\boldsymbol{e}_1|\boldsymbol{e}_2|\dots|\boldsymbol{e}_n) = I_n\,.$$

Ora per mostrare che B è l'inversa di A dobbiamo ancora verificare che $BA = I_n$.

Mostriamo come prima cosa che le colonne di B sono indipendenti. Per l'Osservazione 3.10 è sufficiente verificare che l'unica soluzione dell'equazione $BX = \mathbf{0}_n$ è $X = \mathbf{0}_n$. In effetti se $X = \boldsymbol{u}$ è una soluzione di tale equazione allora abbiamo da una parte che $AB\boldsymbol{u} = A(B\boldsymbol{u}) = A\mathbf{0}_n = \mathbf{0}_n$ e dall'altra $AB\boldsymbol{u} = (AB)\boldsymbol{u} = I_n\boldsymbol{u} = \boldsymbol{u}$ da cui segue che $\boldsymbol{u}$ è necessariamente $\mathbf{0}_n$.

In particolare, le colonne della matrice B formano una base di $\mathbb{R}^n$ e ripetendo lo stesso ragionamento fatto per A, esiste una matrice C tale che $BC = I_n$. Ora sappiamo che $AB = BC = I_n$. Calcoliamo ABC in due modi diversi. Da una parte si ha $ABC = (AB)C = I_nC = C$. Dall'altra abbiamo invece $ABC = A(BC) = AI_n = A$. Da ciò si deduce $A = C$ per cui $BA = BC = I_n$. □

Esempio 3.11. Esaminiamo qualche esempio.

(1) Consideriamo la matrice 2×2

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

Osserviamo che le sue colonne

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

sono indipendenti e dunque sono una base di $\mathbb{R}^2$. La matrice A è dunque invertibile. Proviamo a calcolarne l'inversa: dobbiamo trovare le coordinate degli elementi della base canonica di $\mathbb{R}^2$ rispetto alla base $\mathcal{B}_A = \{A^1, A^2\}$.

Calcoliamo prima di tutto le coordinate $[\boldsymbol{e}_1]_{\mathcal{B}_A}$. Dobbiamo trovare λ, μ tali che

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \lambda \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \mu \begin{pmatrix} 1 \\ -1 \end{pmatrix} .$$

Uguagliando cmponente per componente questa identità vettoriale otteniamo il sistema

$$\begin{cases} \lambda + \mu = 1 \\ \lambda - \mu = 0 \end{cases}$$

la cui unica soluzione è $\lambda = \frac{1}{2}$ e $\mu = \frac{1}{2}$. Ovvero $[\boldsymbol{e}_1]_{\mathcal{B}_A} = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}$.

Analogamente per calcolare le coordinate del vettore $\boldsymbol{e}_2$ rispetto alla base $\mathcal{B}_A$ bisogna trovare λ, μ tali che

$$\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \lambda \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \mu \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

e calcolando come prima si ottiene $\lambda = \frac{1}{2}$ e $\mu = -\frac{1}{2}$, per cui $[\boldsymbol{e}_2]_{\mathcal{B}_A} = \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \end{pmatrix}$.

Dunque abbiamo

$$A^{-1} = ([\boldsymbol{e}_1]_{\mathcal{B}_A} | [\boldsymbol{e}_2]_{\mathcal{B}_A}) = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix} .$$

(2) Consideriamo la matrice 3×3

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

Verifichiamo se le colonne di A sono indipendenti. Chiaramente A^2 non è multiplo di A^1. Se proviamo a scrivere A^3 come combinazione

lineare di A^1 e A^2 si ha
$$\begin{pmatrix}1\\1\\1\end{pmatrix} = \lambda \begin{pmatrix}1\\1\\0\end{pmatrix} + \mu \begin{pmatrix}0\\1\\1\end{pmatrix}.$$
Per cui λ e μ dovrebbero essere soluzioni del sistema
$$\begin{cases} \lambda = 1 \\ \lambda + \mu = 1 \\ \mu = 1 \end{cases}$$
Chiaramente questo sistema non ha soluzioni per cui A^1, A^2, A^3 sono indipendenti. Segue che la matrice A è invertibile.

Calcoliamo l'inversa.

Come prima dobbiamo calcolare le coordinate dei vettori $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$ della base canonica di $\mathbb{R}^3$ rispetto alla base $\mathcal{B}_A$.

Ponendo
$$\begin{pmatrix}1\\0\\0\end{pmatrix} = \lambda \begin{pmatrix}1\\1\\0\end{pmatrix} + \mu \begin{pmatrix}0\\1\\1\end{pmatrix} + \nu \begin{pmatrix}1\\1\\1\end{pmatrix}$$
otteniamo che $\lambda = 0, \mu = -1, \nu = 1$ ovvero $[\boldsymbol{e}_1]_{\mathcal{B}_A} = \begin{pmatrix}0\\-1\\1\end{pmatrix}$.

Ponendo
$$\begin{pmatrix}0\\1\\0\end{pmatrix} = \lambda \begin{pmatrix}1\\1\\0\end{pmatrix} + \mu \begin{pmatrix}0\\1\\1\end{pmatrix} + \nu \begin{pmatrix}1\\1\\1\end{pmatrix}$$
otteniamo $\lambda = 1, \mu = 1, \nu = -1$ ovvero $[\boldsymbol{e}_2]_{\mathcal{B}_A} = \begin{pmatrix}1\\1\\-1\end{pmatrix}$.

Infine ponendo
$$\begin{pmatrix}0\\0\\1\end{pmatrix} = \lambda \begin{pmatrix}1\\1\\0\end{pmatrix} + \mu \begin{pmatrix}0\\1\\1\end{pmatrix} + \nu \begin{pmatrix}1\\1\\1\end{pmatrix}$$
si ricava $\lambda = -1, \mu = 0, \nu = 1$ ovvero $[\boldsymbol{e}_3]_{\mathcal{B}_A} = \begin{pmatrix}-1\\0\\1\end{pmatrix}$.

In definitiva si ha
$$A^{-1} = ([\boldsymbol{e}_1]_{\mathcal{B}_A} | [\boldsymbol{e}_2]_{\mathcal{B}_A} | [\boldsymbol{e}_3]_{\mathcal{B}_A}) = \begin{pmatrix} 0 & 1 & -1 \\ -1 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix}.$$

Possiamo riassumere tutta la questione sull'invertibilità qui discussa nel seguente modo.

TEOREMA 3.10. *Sia $A \in \mathcal{M}_{\mathbb{R}}(n)$, le seguenti proprietà sono equivalenti:*

(1) *A è invertibile;*

(2) *le colonne di A formano una base di $\mathbb{R}^n$;*
(3) *l'equazione vettoriale $AX = \mathbf{0}_n$ ha come unica soluzione $X = \mathbf{0}_n$.*

OSSERVAZIONE 3.12. Consideriamo un'equazione vettoriale del tipo

$$AX = \mathbf{B}, \tag{3.10}$$

dove $A \in \mathcal{M}_{\mathbb{R}}(n)$, è una matrice di ordine $n \times n$ e $\mathbf{B}$ è un vettore colonna di $\mathbb{R}^n$.

Vogliamo mostrare che se A è invertibile allora l'equazione ammette come unica soluzione $\boldsymbol{u} = A^{-1}\mathbf{B}$.

Prima di tutto mostriamo che $\boldsymbol{u}$ è soluzione. Infatti abbiamo

$$A\boldsymbol{u} = A(A^{-1}\mathbf{B}) = (AA^{-1})\mathbf{B} = I_n\mathbf{B} = \mathbf{B}\,.$$

Del resto se $\boldsymbol{w}$ è un'altra soluzione dell'equazione (3.10) allora avremmo

$$A(\boldsymbol{u} - \boldsymbol{w}) = A\boldsymbol{u} - A\boldsymbol{w} = \mathbf{B} - \mathbf{B} = \mathbf{0}_n$$

ovvero $\boldsymbol{u} - \boldsymbol{w}$ sarebbe soluzione dell'equazione $AX = \mathbf{0}_n$. Poiché abbiamo visto che questa equazione ha solo soluzione banale nel caso in cui A sia invertibile ne deduciamo che $\boldsymbol{u} - \boldsymbol{w} = \mathbf{0}_n$, il che contraddice l'assunzione che $\boldsymbol{w}$ sia diversa da $\boldsymbol{u}$.

Riassumendo: nel caso in cui la matrice A sia invertibile, la soluzione dell'equazione (3.10) si ottiene moltiplicando il termine noto $\mathbf{B}$ per l'inversa della matrice A.

Enunciamo alcune delle proprietà di cui godono le matrici invertibili.

(1) Se A è invertibile allora anche la matrice inversa A^{-1} è invertibile e si ha $(A^{-1})^{-1} = A$.

(2) Se A è invertibile e $\lambda \neq 0$ allora λA è invertibile e si ha $(\lambda A)^{-1} = \frac{1}{\lambda}A^{-1}$.

(3) Se A e B sono matrici invertibili di ordine $n \times n$ allora lo è anche il prodotto AB e si ha $(AB)^{-1} = B^{-1}A^{-1}$.
Infatti se poniamo $C = B^{-1}A^{-1}$, allora
$$(AB)C = (AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AI_nA^{-1} = AA^{-1} = I_n$$
e analogamente $C(AB) = B^{-1}A^{-1}AB = I_n$.

(4) Se A è invertibile e $AB = O_{n\times n}$ allora $B = O_{n\times n}$.
Infatti calcolando $A^{-1}AB$ si ottiene da una parte $(A^{-1}A)B = I_nB = B$ e dall'altra $A^{-1}(AB) = A^{-1}O_{n\times n} = O_{n\times n}$, da cui deduciamo che $B = O_{n\times n}$.

OSSERVAZIONE 3.13. L'ultima proprietà non vale per le matrici non invertibili. Infatti abbiamo visto che se A non è invertibile allora esiste un vettore non nullo X tale che $AX = \mathbf{0}_n$. Se consideriamo la matrice $n \times n$ le cui colonne sono tutte uguali al vettore X

$$B = (X|X|X|\dots|X)$$

si verifica facilmente che $AB = O_{n\times n}$.

Esempio 3.14. Si considerino le matrici

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

È facile verificare che $AB = O_{2\times 2}$.

Osservazione 3.15. Riassumendo, possiamo affermare, alla luce di quanto detto, che l'anello $(\mathcal{M}_{\mathbb{R}}(n), +, *)$ non è integro (Definizione 0.19).

Esercizio 3.1. Sia

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

Lo studente determini una matrice 2×2 B tale che $AB = O_{2\times 2}$ (con $B \neq B = O_{2\times 2}$).

Osservazione 3.16. Se A e B sono matrici $n \times n$ invertibili non è detto che lo sia la loro somma. Ad esempio se $B = -A$ si ha che $A + B = O_{n\times n}$.

Osservazione 3.17. Se il prodotto AB di due matrici $n \times n$ è invertibile allora lo sono anche A e B. Infatti sia $\boldsymbol{u}$ una soluzione dell'equazione $BX = \mathbf{0}_n$. Allora abbiamo

$$(AB)\boldsymbol{u} = A(B\boldsymbol{u}) = A\mathbf{0}_n = \mathbf{0}_n .$$

Dunque abbiamo che $\boldsymbol{u}$ è soluzione dell'equazione $(AB)X = \mathbf{0}_n$. Poiché la matrice AB è invertibile l'unica soluzione di tale equazione è $\mathbf{0}_n$ e dunque necessariamente si deve avere $\boldsymbol{u} = \mathbf{0}_n$. In particolare abbiamo verificato che l'unica soluzione dell'equazione $BX = \mathbf{0}_n$ è $\mathbf{0}_n$. Ciò implica che la matrice B è invertibile.

Per verificare che A è invertibile osserviamo che possiamo scrivere

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

e dunque A è il prodotto di due matrici invertibili.

In base alle osservazioni precedenti è possibile mostrare la seguente proposizione; la verifica viene lasciata allo studente per esercizio.

Proposizione 3.11. *L'insieme*

$$\mathcal{GL}(n, \mathbb{R}) = \{M \in \mathcal{M}_{\mathbb{R}}(n) \mid M \ \textit{è invertibile}\} , \tag{3.11}$$

delle matrici quadrate reali di ordine $n\times n$ invertibili*, è un gruppo (non commutativo) rispetto all'operazione di composizione interna data dal prodotto righe per colonne fra matrici. Il gruppo viene denominato* gruppo lineare *delle matrici di ordine n.*

Dimostrazione. Le proprietà di gruppo sono immediatamente verificate; l'elemento neutro del gruppo è la matrice identità, e, per ogni matrice M l'elemento opposto (o inverso) è proprio la matrice inversa M^{-1}, che appartiene sempre all'insieme $\mathcal{GL}(n, \mathbb{R})$ (proprietà (1) della matrice inversa). $\square$

4. Cambiamenti di base

4.1. Calcolo delle coordinate di un vettore di $\mathbb{R}^n$ rispetto ad una base qualsiasi.

Sia $\mathcal{B} = \{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n\}$ una base di $\mathbb{R}^n$. Dato un vettore $\boldsymbol{v} \in \mathbb{R}^n$, le sue coordinate rispetto alla base $\mathcal{B}$ sono gli scalari $\lambda_1, \lambda_2, \ldots, \lambda_n$ tali che

$$\boldsymbol{v} = \lambda_1 \boldsymbol{u}_1 + \lambda_2 \boldsymbol{u}_2 + \ldots + \lambda_n \boldsymbol{u}_n \,. \tag{3.12}$$

Se consideriamo la matrice A le cui colonne sono $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n$:

$$A = (\boldsymbol{u}_1 | \boldsymbol{u}_2 | \ldots | \boldsymbol{u}_n)\,,$$

si osserva che il secondo membro della (3.12) si ottiene moltiplicando la matrice A per il vettore delle coordinate di $\boldsymbol{v}$ (ovvero il vettore ottenuto mettendo per colonna i numeri λ_i).

Segue che il vettore delle coordinate di $\boldsymbol{v}$ rispetto alla base $\mathcal{B}$ (ovvero il vettore $[\boldsymbol{v}]_{\mathcal{B}} = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}$) è soluzione dell'equazione

$$AX = \boldsymbol{v}\,.$$

Del resto, poiché la matrice A è invertibile (le sue colonne sono linearmente indipendenti; vd Teorema 3.10) la soluzione di questa equazione può essere calcolata moltiplicando il vettore $\boldsymbol{v}$ per la matrice A^{-1}.

In definitiva deduciamo che per calcolare le coordinate di $\boldsymbol{v}$ rispetto alla base $\mathcal{B}$ è sufficiente moltiplicare $\boldsymbol{v}$ per la matrice A^{-1}.

Definizione 3.11. La matrice A^{-1} è detta **matrice di cambiamento di base** e viene denotata con $M_{\mathcal{B}}$.

Tale matrice ha la proprietà che moltiplicata per un vettore $\boldsymbol{v}$ restituisce le coordinate del vettore rispetto alla base $\mathcal{B}$

$$[\boldsymbol{v}]_{\mathcal{B}} = M_{\mathcal{B}} \boldsymbol{v}\,.$$

Esempio 3.18. Consideriamo la base di $\mathbb{R}^3$

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

In questo caso la matrice A è uguale a

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

L'inversa di A è stata calcolata nell'Esempio 3.11, per cui si ha

$$M_{\mathcal{B}} = A^{-1} = \begin{pmatrix} 0 & 1 & -1 \\ -1 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix}.$$

Proviamo a calcolare le coordinate del vettore $\boldsymbol{v} = \begin{pmatrix} 1 \\ 4 \\ 2 \end{pmatrix}$ rispetto alla base $\mathcal{B}$. Si ha

$$[\boldsymbol{v}]_{\mathcal{B}} = M_{\mathcal{B}}\boldsymbol{v} = \begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix} .$$

In effetti si verifica facilmente che

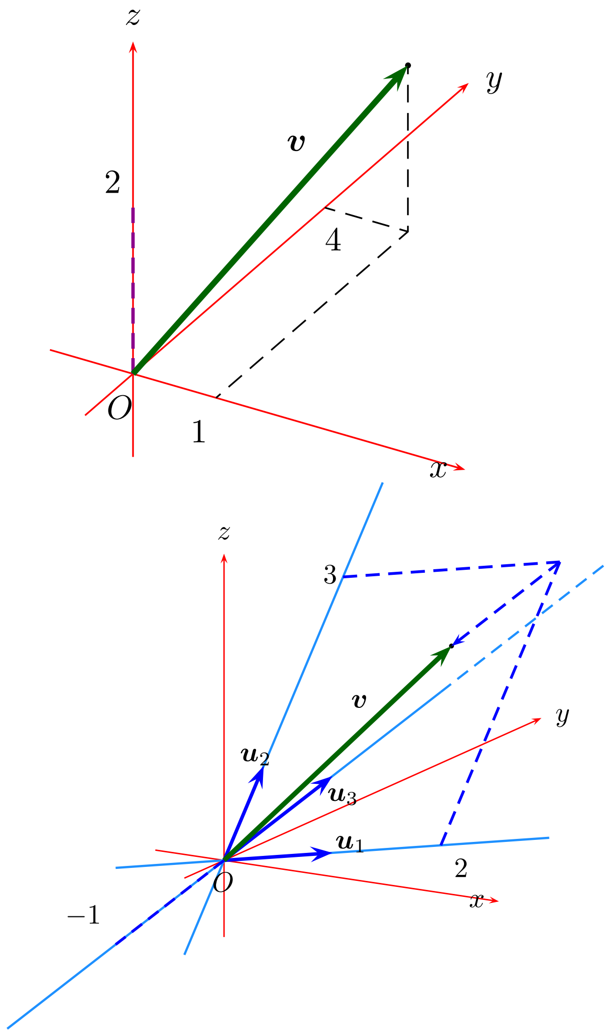

Figura 3.1. Il vettore $\boldsymbol{v}$ nella base standard ha le coordinate $(1, 4, 2)$, e nella nuova base $\mathcal{B} = \{(1, 1, 0), (0, 1, 1), (1, 1, 1)\}$ ha coordinate $(2, 3, -1)$.

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

4.2. Cambiamento di base nel caso generico.

Sia V uno spazio vettoriale di dimesione n. Consideriamo due basi $\mathcal{B} = \{\boldsymbol{u}_1, \dots, \boldsymbol{u}_n\}$ e $\mathcal{D} = \{\boldsymbol{w}_1, \dots, \boldsymbol{w}_n\}$ in V.

Dato un vettore $\boldsymbol{v} \in V$ è possibile calcolare le sue coordinate sia rispetto alla base $\mathcal{B}$ – ovvero $[\boldsymbol{v}]_{\mathcal{B}}$ – che rispetto alla base $\mathcal{D}$ – ovvero $[\boldsymbol{v}]_{\mathcal{D}}$.

In questa sezione ci porremo il problema di stabilire un'esplicita relazione tra questi due set di coordinate.

Più precisamente, supponiamo di conoscere le coordinate dei vettori della base $\mathcal{D}$ rispetto alla base $\mathcal{B}$, ovvero che siano noti i vettori di $\mathbb{R}^n$

$$[\boldsymbol{w}_1]_{\mathcal{B}}, [\boldsymbol{w}_2]_{\mathcal{B}}, \dots, [\boldsymbol{w}_n]_{\mathcal{B}}$$

e consideriamo le seguenti domande:

- Supposto siano note le coordinate di $\boldsymbol{v}$ rispetto a $\mathcal{D}$ come calcolare le coordinate dello stesso vettore rispetto a $\mathcal{B}$?
- Viceversa, supposto che siano note le coordinate di $\boldsymbol{v}$ rispetto a $\mathcal{B}$ come calcolare le coordinate rispetto a $\mathcal{D}$?

La risposta alla prima domanda è abbastanza semplice. Infatti le coordinate di un vettore $\boldsymbol{v} \in V$ sono gli scalari $\mu_1, \mu_2, \dots, \mu_n$ tali che

$$\boldsymbol{v} = \mu_1 \boldsymbol{w}_1 + \mu_2 \boldsymbol{w}_2 + \ldots + \mu_n \boldsymbol{w}_n \,.$$

Esprimendo questa uguaglianza nelle coordinate date dalla base $\mathcal{B}$ si ottiene

$$[\boldsymbol{v}]_{\mathcal{B}} = \mu_1 [\boldsymbol{w}_1]_{\mathcal{B}} + \mu_2 [\boldsymbol{w}_2]_{\mathcal{B}} + \ldots + \mu_n [\boldsymbol{w}_n]_{\mathcal{B}} \tag{3.13}$$

Ora se consideriamo

$$A = ([\boldsymbol{w}_1]_{\mathcal{B}} | [\boldsymbol{w}_2]_{\mathcal{B}} | \ldots | [\boldsymbol{w}_n]_{\mathcal{B}}) \,.$$

Il secondo membro della (3.13) si ottiene moltiplicando la matrice A per il vettore $\begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_n \end{pmatrix}$, che è il vettore delle coordinate di $\boldsymbol{v}$ rispetto alla base $\mathcal{D}$.

Ovvero otteniamo la relazione

$$[\boldsymbol{v}]_{\mathcal{B}} = A[\boldsymbol{v}]_{\mathcal{D}} \tag{3.14}$$

che esprime la semplice proprietà che per calcolare le coordinate di un vettore rispetto alla base $\mathcal{B}$ è sufficiente moltiplicare per la matrice A le coordinate del vettore rispetto alla base $\mathcal{D}$.

Per rispondere alla seconda domanda verifichiamo preliminarmente che le colonne della matrice A sono indipendenti (ovvero che la matrice A è invertibile). Infatti se prendiamo una combinazione lineare delle colonne di A che dà come risultato il vettore nullo

$$\lambda_1 [\boldsymbol{w}_1]_{\mathcal{B}} + \lambda_2 [\boldsymbol{w}_2]_{\mathcal{B}} + \ldots + \lambda_n [\boldsymbol{w}_n]_{\mathcal{B}} = \mathbf{0}_n \,,$$

la corrispondente combinazione lineare dei vettori $\boldsymbol{w}_i$ dà come risultato il vettore nullo

$$\lambda_1 \boldsymbol{w}_1 + \lambda_2 \boldsymbol{w}_2 + \ldots + \lambda_n \boldsymbol{w}_n = \mathbf{0}_V$$

e poiché i $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n$ sono indipendenti, i coefficienti $\lambda_1, \lambda_2, \ldots, \lambda_n$ sono tutti nulli. In particolare l'unica combinazione dei vettori numerici $[\boldsymbol{w}_i]_{\mathcal{B}}$ che dà il vettore nullo è la combinazione banale. Ciò mostra che la matrice A è invertibile.

Dunque dalla (3.14) ricaviamo che

$$[\boldsymbol{v}]_{\mathcal{D}} = A^{-1}[\boldsymbol{v}]_{\mathcal{B}}$$

che esprime la proprietà che per calcolare le coordinate di un vettore $\boldsymbol{v}$ rispetto alla base $\mathcal{D}$ è sufficiente moltiplicare per l'inversa della matrice A le coordinate del vettore rispetto alla base $\mathcal{B}$

Definizione 3.12.
- La matrice A è detta **matrice di cambiamento di base da $\mathcal{D}$ a $\mathcal{B}$** e viene denotata con $A = M_{\mathcal{B},\mathcal{D}}$.
- La matrice A^{-1} è invece detta **matrice di cambiamento di base da $\mathcal{B}$ a $\mathcal{D}$** e viene denotata con $A^{-1} = M_{\mathcal{D},\mathcal{B}}$.

La matrice $M_{\mathcal{B},\mathcal{D}}$ ha la proprietà di trasformare le coordinate rispetto alla base $\mathcal{D}$ nelle coordinate rispetto alla base $\mathcal{B}$, nel senso che vale la seguente identità:

$$[\boldsymbol{v}]_{\mathcal{B}} = M_{\mathcal{B},\mathcal{D}}[\boldsymbol{v}]_{\mathcal{D}} \tag{3.15}$$

La matrice $M_{\mathcal{D},\mathcal{B}}$ e ha la proprietà di trasformare le coordinate rispetto alla base $\mathcal{B}$ nelle coordinate rispetto alla base $\mathcal{D}$, nel senso che vale la seguente identità:

$$[\boldsymbol{v}]_{\mathcal{D}} = M_{\mathcal{D},\mathcal{B}}[\boldsymbol{v}]_{\mathcal{B}}\,. \tag{3.16}$$

Esempio 3.19. (1) Si consideri nel piano $\mathbb{E}^2_O$ due basi $\mathcal{B} = \{\hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}\}$ e $\mathcal{D} = \{\boldsymbol{u}_0, \boldsymbol{w}_0\}$ dove $\boldsymbol{u}_0 = \hat{\boldsymbol{\imath}} + \hat{\boldsymbol{\jmath}}$ e $\boldsymbol{w}_0 = \hat{\boldsymbol{\imath}} - \hat{\boldsymbol{\jmath}}$. Dato un vettore $\boldsymbol{v} = \overline{OP}$ indichiamo con x, y le coordinate rispetto a $\mathcal{B}$ e con x', y' le coordinate rispetto a $\mathcal{D}$.

Osserviamo che si ha

$$\boldsymbol{v} = x'\boldsymbol{u}_0 + y'\boldsymbol{w}_0$$

e passando nelle coordinate di $\mathcal{B}$ tale relazione diventa

$$\begin{pmatrix} x \\ y \end{pmatrix} = x' \begin{pmatrix} 1 \\ 1 \end{pmatrix} + y' \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

ovvero

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix} \tag{3.17}$$

In particolare la matrice di cambiamento di base dalla base $\mathcal{D}$ alla base $\mathcal{B}$ è semplicemente la matrice

$$M_{\mathcal{B},\mathcal{D}} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

(si noti che le colonne di tale matrice sono proprio i vettori delle coordinate di $\boldsymbol{u}_0$, $\boldsymbol{w}_0$ rispetto alla base $\mathcal{B}$).

Utilizzando la (3.17) è possibile esprimere le coordinate (x, y) in termini delle coordinate x', y' infatti si ha

$$\begin{cases} x = x' + y' \\ y = x' - y' \end{cases} .$$

Per esprimere le coordinate x', y' in termini delle coordinate x, y dalla (3.17) si ricava che

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \left(\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\right)^{-1} \begin{pmatrix} x \\ y \end{pmatrix} \tag{3.18}$$

Dunque ritroviamo che la matrice di cambiamento di base da $\mathcal{B}$ a $\mathcal{D}$ è l'inversa di $M_{\mathcal{B},\mathcal{D}}$. Tale inversa è stata calcolata nell'Esempio 3.11 per cui risulta

$$M_{\mathcal{D},\mathcal{B}} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix}$$

e dalla 3.18 ricaviamo

$$\begin{cases} x' = \frac{x+y}{2} \\ y' = \frac{x-y}{2} \end{cases} \tag{3.19}$$

che esprime le coordinate x', y' in termini delle coordinate x, y.

(2) Siano $\mathcal{B}$ e $\mathcal{D}$ come nell'esempio precedente. Si consideri la retta r nel piano che ha equazione parametrica (nelle coordinate date dalla base $\mathcal{B}$)

$$\begin{cases} x = 2t + 3 \\ y = t - 1 \end{cases} .$$

Calcoliamo l'equazione parametrica di r nelle coordinate x', y'. Osserviamo che l'equazione data esprime le coordinate x, y dei punti sulla retta in termini di un parametro t. Ora per la (3.18) il vettore $\overline{OP}$ che ha coordinate $x = 2t + 3$ e $y = t - 1$ (rispetto a $\mathcal{B}$) ha coordinate

$$x' = \frac{x+y}{2} = \frac{3}{2}t + 1 \qquad y' = \frac{x-y}{2} = \frac{1}{2}t + 1$$

(rispetto a $\mathcal{D}$). Osserviamo che in tal modo abbiamo espresso le coordinate x', y' dei punti sulla retta r in termini del parametro t e dunque abbiamo trovato l'equazione parametrica cercata.

Finora abbiamo supposto di lavorare **conoscendo le coordinate dei vettori della base $\mathcal{D}$ rispetto alla base $\mathcal{B}$**.

In tali ipotesi risulta piuttosto semplice calcolare la matrice di cambiamento di base da $\mathcal{D}$ a $\mathcal{B}$ (ovvero la matrice $M_{\mathcal{B},\mathcal{D}}$). Infatti le colonne di tale matrice sono semplicemente i vettori delle coordinate dei vettori di $\mathcal{D}$ rispetto a $\mathcal{B}$.

La matrice di cambiamento di base da $\mathcal{B}$ a $\mathcal{D}$ è invece più difficile da calcolare in quanto bisogna trovare l'inversa della matrice $M_{\mathcal{B},\mathcal{D}}$ (e il calcolo dell'inversa di una matrice è un'operazione dispendiosa da un punto di vista computazionale).

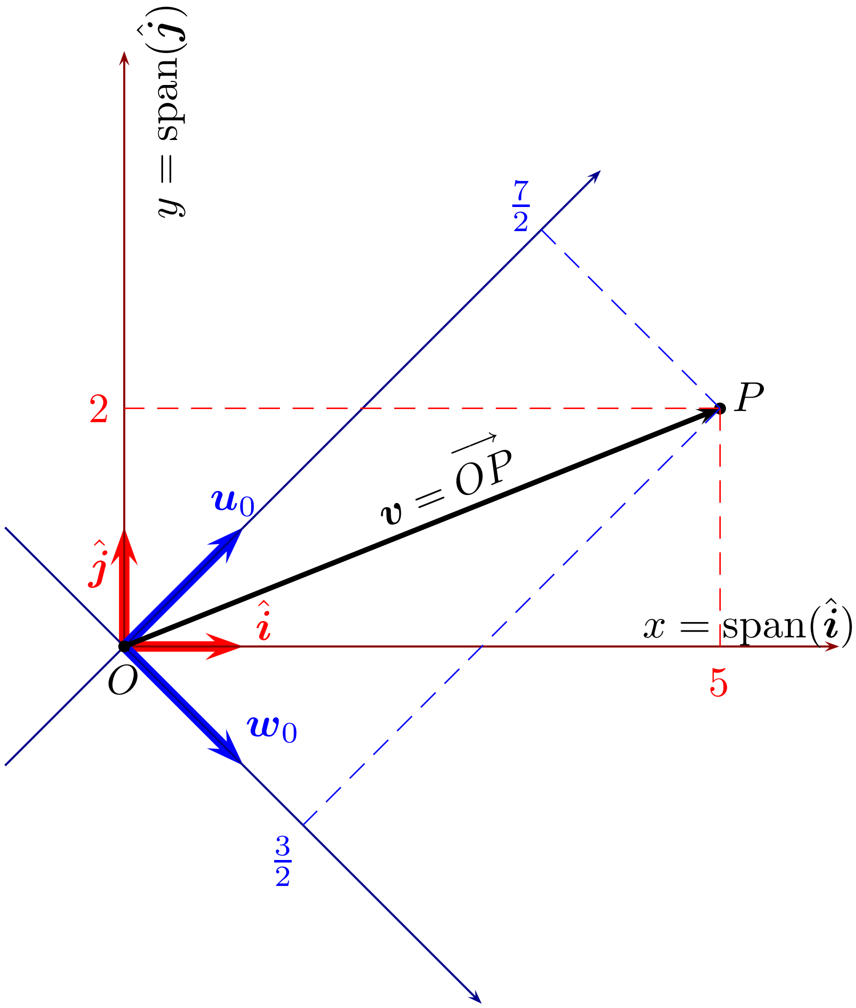

Figura 3.2. Le diverse coordinate di un medesimo vettore nel piano $\mathbb{E}^2_O$ si ottengono proiettando versi assi il vettore. Per esempio, si consideri il vettore $\boldsymbol{v} = \overrightarrow{OP}$ corrispondente al punto di coordinate $(5, 2)$, coincidenti con le sue componenti sulla base canonica $\mathcal{B} = \{\hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}\}$. È facile convincersi che ha le coordinate $(7/2, 3/2)$ nella base $\mathcal{D} = \{\boldsymbol{u}_0, \boldsymbol{w}_0\}$, con $\boldsymbol{u}_0 = \hat{\boldsymbol{\imath}} + \hat{\boldsymbol{\jmath}}$ e $\boldsymbol{w}_0 = \hat{\boldsymbol{\imath}} - \hat{\boldsymbol{\jmath}}$.

Chiaramente se lavorassimo nelle ipotesi opposte (ovvero conoscendo le coordinate dei vettori della base $\mathcal{B}$ rispetto alla base $\mathcal{D}$) avremmo esattamente la situazione opposta:

- la matrice di cambiamento di base da $\mathcal{B}$ a $\mathcal{D}$ si ottiene semplicemente mettendo per colonna le coordinate dei vettori $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n$ rispetto alla base $\mathcal{D}$
- la matrice di cambiamento di base da $\mathcal{D}$ a $\mathcal{B}$ si ottiene calcolando l'inversa della precedente matrice.

5. L'operazione di trasposizione

DEFINIZIONE 3.13. Sia A una matrice di ordine $k \times n$, definiamo la matrice **trasposta** di A che indicheremo con A^{T}:
- se A è un vettore riga, A^{T} è semplicemente il vettore A messo per colonna;
- se A è un vettore colonna, A^{T} è il vettore A messo per riga;
- in generale la matrice A^{T} è la matrice le cui colonne sono le righe di A trasposte:

$$A^{\mathrm{T}} = ((A_1)^{\mathrm{T}}|(A_2)^{\mathrm{T}}|\dots|(A_k)^{\mathrm{T}}).$$

ESEMPIO 3.20.

$$(3, 4, 7)^{\mathrm{T}} = \begin{pmatrix} 3 \\ 4 \\ 7 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 7 & 3 \\ 4 & 8 & 12 \end{pmatrix}^{\mathrm{T}} = \begin{pmatrix} 1 & 4 \\ 7 & 8 \\ 3 & 12 \end{pmatrix}$$

Discutiamo alcune semplici proprietà della trasposizione.

(1) Se A è una matrice $k \times n$ allora A^{T} è una matrice $n \times k$. Ovvero il numero di colonne di A^{T} è uguale al numero di righe di A e viceversa il numero di righe di A^{T} è ugual al numero di colonne di A.

(2) Per definizione le colonne di A^{T} sono le righe di A messe per colonna ovvero

$$(A^{\mathrm{T}})^i = (A_i)^{\mathrm{T}} \qquad \text{per } i = 1 \dots k\,.$$

Viceversa le righe di A^{T} sono le colonne di A messe per riga, ovvero

$$(A^{\mathrm{T}})_i = (A^i)^{\mathrm{T}} \qquad \text{per } i = 1 \dots n\,.$$

(3) Nella posizione $(1, 2)$ della matrice A^{T} c'è la prima componente della seconda colonna. Poiché la seconda colonna della matrice A^{T} è uguale alla seconda riga della matrice A si ha che l'elemento $(1, 2)$ della matrice A^{T} è uguale alla prima componente della seconda riga della matrice A, ovvero all'elemento $a_{2\,1}$. In simboli abbiamo $(A^{\mathrm{T}})_{1\,2} = a_{2\,1}$. In maniera analoga si verifica che **vale la formula generale**

$$(A^{\mathrm{T}})_{ij} = a_{ji}\,.$$

(4) Trasponendo due volte una matrice si ritrova la matrice stessa. Ovvero

$$(A^{\mathrm{T}})^{\mathrm{T}} = A\,.$$

Infatti la prima riga della matrice $(A^{\mathrm{T}})^{\mathrm{T}}$ è uguale alla prima colonna della matrice A^{T}. Del resto la prima colonna della matrice A^{T} è uguale alla prima riga della matrice A, per cui si ha che la prima riga della matrice $(A^{\mathrm{T}})^{\mathrm{T}}$ è uguale alla prima riga della matrice A.

Analogamente si verifica che tutte le righe di $(A^{\mathrm{T}})^{\mathrm{T}}$ sono uguali alle corrispondenti righe di A.

(5) La trasposta della matrice λA è uguale alla matrice che si ottiene moltiplicando per λ la trasposta della matrice A. In simboli
$$(\lambda A)^{\mathrm{T}} = \lambda A^{\mathrm{T}}\,.$$
(6) La trasposta di una somma di due matrici $A+B$ è uguale alla somma delle trasposte di A e B. In simboli
$$(A+B)^{\mathrm{T}} = A^{\mathrm{T}} + B^{\mathrm{T}}\,.$$
(7) Se A è una matrice $k\times n$ e B è una matrice $n\times h$ allora vale la seguente formula
$$(AB)^{\mathrm{T}} = B^{\mathrm{T}}A^{\mathrm{T}}\,.$$
Prima di tutto verifichiamo che valga la formula a livello di ordini. La matrice AB ha ordine $k\times h$ per cui la matrice $(AB)^{\mathrm{T}}$ ha ordine $h\times k$.

Del resto le matrici B^{T} e A^{T} hanno ordini rispettivamente $h\times n$ e $n\times k$, per cui il loro prodotto $B^{\mathrm{T}}A^{\mathrm{T}}$ ha ordine $h\times k$.

Verifichiamo ora la formula. Se A è un vettore riga e B è un vettore colonna (ovvero $k=h=1$) si ha che AB e $B^{\mathrm{T}}A^{\mathrm{T}}$ coincidono con il numero ottenuto sommando il prodotto delle componenti di A per le rispettive componenti di B. In questo caso inoltre $(AB)^{\mathrm{T}} = AB$ in quanto il trasposto di un numero è il numero stesso. Dunque la formula è vera nel caso di prodotto di una riga per una colonna.

Consideriamo ora il caso generico. Calcoliamo l'elemento $(2,3)$ della matrice $(AB)^{\mathrm{T}}$. Si ha che
$$((AB)^{\mathrm{T}})_{2\,3} = (AB)_{3\,2} = A_3B^2 \tag{3.20}$$
dove nell'ultimo passaggio abbiamo utilizzato la regola della moltiplicazione riga per colonna.

Calcoliamo ora l'elemento $(2,3)$ della matrice $B^{\mathrm{T}}A^{\mathrm{T}}$. Tale elemento è uguale al prodotto della seconda riga della matrice B^{T} per la terza colonna della matrice A^{T} ovvero
$$(B^{\mathrm{T}}A^{\mathrm{T}})_{2\,3} = (B^{\mathrm{T}})_2(A^{\mathrm{T}})^3\,.$$
Del resto la seconda riga della matrice B^{T} è uguale alla seconda colonna della matrice B (messa per riga) e analogamente la terza colonna della matrice A^{T} è uguale alla terza riga della matrice A (messa per colonna). Dunque ricaviamo
$$(B^{\mathrm{T}}A^{\mathrm{T}})_{2\,3} = (B^2)^{\mathrm{T}}(A_3)^{\mathrm{T}} = A_3B^2 \tag{3.21}$$
dove l'ultimo passaggio segue dal fatto che abbiamo verificato la formula nel caso del prodotto di una riga per una colonna. Confrontando la (3.20) con la (3.21) si ottiene che l'entrata (2,3) della matrice $(AB)^{\mathrm{T}}$ coincide con le corrispondenti entrate della matrice $B^{\mathrm{T}}A^{\mathrm{T}}$.

In maniera simile si verifica che tutte le altre entrate coincidono e dunque le due matrici sono uguali.

(8) La trasposta della matrice identità è ancora la matrice identità:
$$(I_n)^{\mathrm{T}} = I_n\,.$$

(9) Se A è quadrata ed invertibile allora anche A^{T} è invertibile e l'inversa di A^{T} è la traposta di A^{-1}. In simboli:

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

Infatti posto $B = A^{-1}$ abbiamo

$$B^{\mathrm{T}}A^{\mathrm{T}} = (AB)^{\mathrm{T}} = (AA^{-1})^{\mathrm{T}} = (I_n)^{\mathrm{T}} = I_n$$

e analogamente si verifica che $A^{\mathrm{T}}B^{\mathrm{T}} = I_n$.

6. Il determinante

Definizione 3.14. Ad ogni matrice **quadrata** A è possibile associare uno scalare detto il **determinante** della matrice, che indichiamo con $\det A$ o con $|A|$.

• Se A è una matrice 1×1, $A = (a)$, poniamo $\det A = a$.

• Se A è una matrice 2×2:

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

il determinante viene calcolato nel seguente modo. Si considera la prima colonna di A. Si moltiplica il primo coefficiente di tale colonna (ovvero a) per il numero ottenuto cancellando la riga e la colonna che contengono a (ovvero d). Analogamente si moltiplica il secondo coefficiente (ovvero c) per il numero ottenuto cancellando la riga e la colonna che contengono c (ovvero b) Il determinante di A è semplicemente la differenza dei numeri così trovati. Ovvero

$$\det A = ad - cb.$$

• Se A è una matrice 3×3:

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

anche in questo caso consideriamo la prima colonna della matrice A. In corrispondenza del numero a_{11} denotiamo con $A_{[1,1]}$ la matrice 2×2 ottenuta cancellando la riga e la colonna che contengono a_{11}. Chiaramente si ha

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \qquad \to \qquad A_{[1,1]} = \begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix}$$

Analogamente in corrispondenza del numero a_{21} consideriamo la matrice $A_{[2,1]}$ ottenuta cancellando la riga e la colonna che contengono a_{21}. In questo caso si ha

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \qquad \to \qquad A_{[2,1]} = \begin{pmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{pmatrix}.$$

Infine consideriamo la matrice $A_{[3,1]}$ ottenuta cancellando la riga e la colonna che contengono l'elemento a_{31} ovvero

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \qquad \to \qquad A_{[31]} = \begin{pmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{pmatrix} .$$

Abbiamo visto come calcolare il determinante delle matrici 2×2 e dunque in particolare possiamo calcolare il determinante delle matrici $A_{[1,1]}$, $A_{[2,1]}$ e $A_{[3,1]}$. Poniamo allora

$$\det A = a_{11} \det A_{[1,1]} - a_{21} \det A_{[2,1]} + a_{31} \det A_{[3,1]} .$$

- Se A è una matrice $n \times n$, consideriamo la prima colonna di A

$$A = \begin{pmatrix} a_{11} & * & * & \dots & * \\ a_{21} & * & * & \dots & * \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & * & * & \dots & * \end{pmatrix} ,$$

Conviene definire in modo generale la notazione introdotta prima.

Notazione 3.15. Denotiamo con $A_{[i,j]}$ la matrice $(n-1) \times (n-1)$ ottenuta cancellando da A la riga e la colonna che contengono a_{ij}.

Allora poniamo:

$$\det A = a_{11} \det A_{[1,1]} - a_{21} \det A_{[2,1]} + a_{31} \det A_{[3,1]} + \ldots + (-1)^{n+1} a_{n1} \det A_{[n,1]} .$$

Tale formula è detta **sviluppo del determinante lungo la prima colonna**, e può essere anche scritta nel seguente modo:

$$\det A = \sum_{i=1}^{n} (-1)^{i+1} a_{i1} \det A_{[i,1]} .$$

Osservazione 3.21. Osserviamo che:

- questa formula permette di scrivere il determinante di A come una somma a **segni alterni** di n addendi. L'i-esimo addendo è:

$$(-1)^{i+1} a_{i1} \det A_{[i,1]} ,$$

  dove

$$(-1)^{i+1} = \begin{cases} +1 & \text{se } i+1 \text{ è pari} \\ -1 & \text{se } i+1 \text{ è dispari} \end{cases} ;$$

- questa formula permette di ridurre il calcolo del determinante di A al calcolo di determinanti delle matrici $A_{[i,1]}$ che hanno ordine $(n-1) \times (n-1)$. A sua volta lo sviluppo dei determinanti delle matrici $A_{[i,1]}$ permette di ridurre il problema al calcolo di determinanti di matrici di ordine ancora più basso. È possibile riapplicare lo sviluppo fino a che non ci si riduce al calcolo di determinanti di matrici 1×1 per le quali il determinante si calcola banalmente.

Esempio 3.22. Consideriamo il calcolo dei seguenti determinanti.

(1)
$$\det\begin{pmatrix}1 & 2\\ 3 & 4\end{pmatrix} = 1\cdot 4 - 3\cdot 2 = -4-6=-2\,.$$

(2) Calcoliamo il determinante della matrice $A=\begin{pmatrix}1 & 0 & 1\\ 1 & 1 & 1\\ 0 & 1 & 1\end{pmatrix}$

In questo caso si ha
$$\det A = 1\cdot\begin{vmatrix}1 & 1\\ 1 & 1\end{vmatrix} - 1\cdot\begin{vmatrix}0 & 1\\ 1 & 1\end{vmatrix} + 0\cdot\begin{vmatrix}0 & 1\\ 1 & 1\end{vmatrix}$$

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

da cui segue che
$$\det A = 1\cdot 0 - 1\cdot(-1) + 0\cdot(-1) = 1$$

(3) Calcoliamo il determinante della matrice $A=\begin{pmatrix}4 & 1 & 1 & 1\\ 2 & 1 & 1 & 0\\ 5 & 3 & 0 & 0\\ 1 & 0 & 0 & 0\end{pmatrix}$.

Sviluppando lungo la prima colonna risulta
$$\det A = 4\det A_{[1,1]} - 2\det A_{[2,1]} + 3\det A_{[3,1]} - \det A_{[4,1]}\,.$$

Dunque dobbiamo calcolare il determinante delle matrici $A_{[i,1]}$.

Abbiamo
$$\det A_{[1,1]} = \begin{vmatrix}1 & 1 & 0\\ 3 & 0 & 0\\ 0 & 0 & 0\end{vmatrix} = 1\cdot\begin{vmatrix}0 & 0\\ 0 & 0\end{vmatrix} - 3\cdot\begin{vmatrix}1 & 0\\ 0 & 0\end{vmatrix} = 0\,,$$

$$\det A_{[2,1]} = \begin{vmatrix}1 & 1 & 1\\ 3 & 0 & 0\\ 0 & 0 & 0\end{vmatrix} = 1\begin{vmatrix}0 & 0\\ 0 & 0\end{vmatrix} - 3\cdot\begin{vmatrix}1 & 1\\ 0 & 0\end{vmatrix} = 0\,,$$

$$\det A_{[3,1]} = \begin{vmatrix}1 & 1 & 1\\ 1 & 1 & 0\\ 0 & 0 & 0\end{vmatrix} = 1\cdot\begin{vmatrix}1 & 0\\ 0 & 0\end{vmatrix} - 1\cdot\begin{vmatrix}1 & 1\\ 0 & 0\end{vmatrix} = 0\,,$$

$$\det A_{[4,1]} = \begin{vmatrix}1 & 1 & 1\\ 1 & 1 & 0\\ 3 & 0 & 0\end{vmatrix} = 1\cdot\begin{vmatrix}1 & 0\\ 0 & 0\end{vmatrix} - 1\cdot\begin{vmatrix}1 & 1\\ 0 & 0\end{vmatrix} + 3\begin{vmatrix}1 & 1\\ 1 & 0\end{vmatrix} = -3$$

e dunque
$$\det A = -(-3) = 3\,.$$

Ci si potrebbe domandare cosa si ottiene se invece di considerare lo sviluppo lungo la prima colonna si considera lo sviluppo lungo ad esempio la seconda colonna o la terza riga. Più precisamente, sia $A_{[i,j]}$ la matrice ottenuta cancellando la riga e la colonna che contengono l'entrata a_{ij} (ovvero la riga i e la colonna j).

Allora, se consideriamo la seconda colonna di A

$$A = \begin{pmatrix} * & a_{12} & * & \dots & * \\ & a_{22} & * & \dots & * \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ * & a_{1n} & * & \dots & * \end{pmatrix}$$

svolgendo il calcolo lungo tale colonna, seguendo il procedimento illustrato prima, otteniamo:

$$-a_{12} \det A_{[1,2]} + a_{22} \det A_{[2,2]} - a_{32} \det A_{[3,2]} + \ldots + (-1)^{n+2} a_{n2} \det A_{[n,2]},$$

analogamente, se consideriamo la terza riga

$$A = \begin{pmatrix} * & * & \dots & * \\ * & * & \dots & * \\ a_{31} & a_{32} & \dots & a_{3n} \\ * & * & \dots & * \\ \vdots & \vdots & \ddots & \vdots \end{pmatrix}$$

otteniamo:

$$a_{3\,1} \det A_{[3,1]} - a_{32} \det A_{[3,2]} + a_{33} \det A_{[3,3]} + \ldots + (-1)^{n+3} a_{3n} \det A_{[3,n]}\,.$$

Il seguente teorema ci dice che sviluppando lungo una qualsiasi colonna o riga di A si ottiene $\det A$.

Teorema 3.12 (di Laplace). *Sia A una matrice $n \times n$.*

- *Fissata la colonna j vale che*

$$\det A = (-1)^{j+1} a_{1j} \det A_{[1,j]} + (-1)^{j+2} a_{2j} \det A_{[2,j]} + \ldots + (-1)^{j+n} a_{n,j} \det A_{[n,j]}\,.$$

- *Fissata la riga i vale che*

$$\det A = (-1)^{i+1} a_{i1} \det A_{[i,1]} + (-1)^{i+2} a_{i2} \det A_{[i,2]} + \ldots + (-1)^{i+n} a_{in} \det A_{[i,n]}\,.$$

In forma compatta, usando il simbolo di sommatoria:

$$\det A = \sum_{i=1}^{n} (-1)^{i+j}\, a_{ij} \det A_{[i,j]} = \sum_{j=1}^{n} (-1)^{i+j}\, a_{ij} \det A_{[i,j]}. \tag{3.22}$$

Omettiamo la dimostrazione del teorema, un po' tecnica; in Appendice (Sezione 9) enunceremo e dimostreremo che il calcolo del determinante può essere effettuato sviluppando lungo la prima riga; la dimostrazione del Teorema di Laplace 3.12 può essere condotta con tecniche analoghe.
Osserviamo che:
calcolando il determinante di A, sviluppando lungo una qualsiasi riga o una qualsiasi colonna di A, è importante stabilire il segno dell' addendo $a_{ij} \det A_{[i,j]}$.
Diciamo che:

- a_{ij} occupa **un posto pari** se $i+j$ è pari,
- a_{ij} occupa **un posto dispari** se $i+j$ è dispari.

Il segno del termine $a_{ij} \det A_{[i,j]}$ è quindi $+$ se a_{ij} occupa **un posto pari**, $-$ se a_{ij} occupa **un posto dispari**.
Un modo semplice per ricordarsi il segno, è costruirsi una tabella di segni alterni come la seguente (costruita nel caso 6×6):

$$\begin{pmatrix} + & - & + & - & + & - \\ - & + & - & + & - & + \\ + & - & + & - & + & - \\ - & + & - & + & - & + \\ + & - & + & - & + & - \\ - & + & - & + & - & + \end{pmatrix},$$

il segno cercato è esattamente quello che appare nella posizione (i,j) di tale tabella. Ad esempio il segno che moltiplica il termine $a_{3,2}|A_{[3,2]}|$ è $-$.

Vediamo ora alcuni esempi.

Esempio 3.23.

(1) Proviamo a ricalcolare il determinante della matrice $A = \begin{pmatrix} 4 & 1 & 1 & 1 \\ 2 & 1 & 1 & 0 \\ 5 & 3 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$.
considerata nell'Esempio 3.22 sviluppando lungo la quarta colonna.

Otteniamo

$$\det A = -(1) \cdot \det A_{[1,4]} + 0 \cdot \det A_{[2,4]} - 0 \cdot A_{[3,4]} + 0 \cdot A_{[4,4]} = -\det A_{[1,4]}$$

Ora abbiamo

$$A_{[1,4]} = \begin{pmatrix} 2 & 1 & 1 \\ 5 & 3 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

e dunque sviluppando il determinante di $A_{[1,4]}$ lungo la terza colonna otteniamo

$$\det A_{[1,4]} = 1 \cdot \det \begin{vmatrix} 5 & 3 \\ 1 & 0 \end{vmatrix} = -3\,.$$

Da cui si conclude che $\det A = 3$.

Osserviamo che in questo caso lo sviluppo lungo la quarta colonna è risultato molto più semplice rispetto allo sviluppo lungo la prima colonna. Ciò è essenzialmente dovuto alla presenza degli zeri lungo la quarta colonna che ci hanno permesso di calcolare unicamente un determinante 3×3.

(2) Calcoliamo il determinante della matrice $A = \begin{pmatrix} 1 & 1 & 3 & 4 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 4 & 5 \\ 7 & 3 & 13 & 6 \end{pmatrix}$.

Osserviamo che in questo caso conviene sviluppare lungo la seconda riga. Otteniamo:

$$\det A = -0 \cdot \det A_{[2,1]} + 0 \cdot \det A_{[2,2]} - \det A_{[2,3]} + 0 \cdot A_{[2,4]} = -\det A_{[2,3]}.$$

Ora si ha

$$\det A_{[2,3]} = \begin{vmatrix} 1 & 1 & 4 \\ 0 & 1 & 5 \\ 7 & 3 & 6 \end{vmatrix} .$$

Sviluppando lungo la seconda riga si ottiene

$$\det A_{[2,3]} = \begin{vmatrix} 1 & 4 \\ 7 & 6 \end{vmatrix} - 5 \begin{vmatrix} 1 & 1 \\ 7 & 3 \end{vmatrix} = (6-28) - 5(3-7) = -2 \, ,$$

e dunque $\det A = 2$.

6.1. Determinante di matrici triangolari.

DEFINIZIONE 3.16. Sia A una matrice di ordine $n \times n$. La **diagonale principale** di A è la diagonale che va dall'angolo in alto a sinistra all'angolo in basso a destra.

Ad esempio nella matrice

$$A = \begin{pmatrix} 3 & 7 & 4 \\ 2 & 15 & -1 \\ 6 & 5 & 8 \end{pmatrix}$$

abbiamo messo in evidenza (in rosso) le entrate sulla diagonale principale. Esse sono: 3, 15, 8. Osserviamo che 3 occupa il posto $(1,1)$, 15 il posto $(2,2)$ e 8 il posto $(3,3)$.

Osserviamo che, in generale, in una matrice A di ordine $n \times n$, le entrate sulla diagonale principale occupano i posti (i,i), cioè sono le entrate:

$$\{a_{ii}\}, \quad i = 1, \dots, n.$$

DEFINIZIONE 3.17. Sia A una matrice di ordine $n \times n$.
• A si dice **triangolare superiore** se tutte le entrate sotto la diagonale principale sono nulle;
• A si dice **triangolare inferiore** se tutte le entrate sopra la diagonale principale sono nulle;
• A si dice **diagonale** se tutte le entrate fuori dalla diagonale principale sono nulle.

Ad esempio, la matrice seguente è una matrice triangolare superiore:

$$A = \begin{pmatrix} 8 & 7 & 4 \\ 0 & 5 & 6 \\ 0 & 0 & 2 \end{pmatrix} ,$$

la matrice seguente è una matrice triangolare inferiore:

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

mentre matrice seguente è una matrice diagonale:

$$D = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & 4 \end{pmatrix} .$$

Si osservi che una matrice diagonale è contemporaneamente triangolare inferiore e superiore.

Per le matrici triangolari (inferiori o superiori) il calcolo del determinante è piuttosto semplice.

Esempio 3.24.
Calcoliamo il determinante della seguente matrice riangolare:

$$A = \begin{pmatrix} 4 & 7 & 8 & 10 \\ 0 & 3 & 2 & 6 \\ 0 & 0 & 5 & 1 \\ 0 & 0 & 0 & 10 \end{pmatrix} .$$

Sviluppando lungo la prima colonna si ha

$$|A| = 4 \begin{vmatrix} 3 & 2 & 6 \\ 0 & 5 & 1 \\ 0 & 0 & 10 \end{vmatrix} = 4 \cdot 3 \begin{vmatrix} 5 & 1 \\ 0 & 10 \end{vmatrix} = 4 \cdot 3 \cdot 5 \cdot 10 \, ,$$

cioè il determinante di A è il prodotto degli elementi sulla diagonale principale.

Questa proprietà è vera in generale per le matrici triangolari (superiori, inferiori o diagonali) di ordine $n \times n$.

Proprietà 3.13. *Il determinante di una matrice triangolare è il prodotto degli elementi sulla diagonale principale.*

Osservazione 3.25. La matrice identità è una matrice diagonale e dunque il suo determinante è dato dal prodotto degli elementi sulla diagonale. Per cui si ha:

$$\det I_n = 1 \, .$$

6.2. Proprietà del determinante.

L'utilità del determinante è essenzialmente legata ad alcune sue proprietà, che andremo ora ad elencare.

Proprietà 3.14. Per ogni matrice A quadrata di ordine $n \times n$, valgono le seguenti proprietà:

(1) **Il determinante di A è uguale al determinante della sua trasposta A^{T}:**

$$\det A = \det(A^{\mathrm{T}}).$$

Osserviamo che poiché le righe di A^{T} coincidono con le colonne di A, sviluppando il determinante di A^{T} lungo la prima riga si ottiene esattamente lo sviluppo del determinante di A lungo la prima colonna.

(2) **Scambiando due colonne di A il determinante cambia segno.** Verifichiamo prima di tutto questa proprietà nel caso in cui le due colonne siano adiacenti. Ci limiteremo a fare la verifica in un caso particolare: sia $A = \begin{pmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{pmatrix}$ e sia $A' = \begin{pmatrix} 4 & 1 & 7 \\ 5 & 2 & 8 \\ 6 & 3 & 9 \end{pmatrix}$ la matrice ottenuta scambiando la prima colonna di A con la seconda. Dobbiamo confrontare il determinante di A con il determinante di A'. Sviluppando lungo la prima colonna il determinante di A si ottiene

$$\det A = \begin{vmatrix} 5 & 8 \\ 6 & 9 \end{vmatrix} - 2\begin{vmatrix} 4 & 7 \\ 6 & 9 \end{vmatrix} + 3\begin{vmatrix} 4 & 7 \\ 5 & 8 \end{vmatrix}\,.$$

Sviluppando lungo la seconda colonna il determinante di A' si ottiene

$$\det A' = -\begin{vmatrix} 5 & 8 \\ 6 & 9 \end{vmatrix} + 2\begin{vmatrix} 4 & 7 \\ 6 & 9 \end{vmatrix} - 3\begin{vmatrix} 4 & 7 \\ 5 & 8 \end{vmatrix}\,,$$

e dunque $\det A' = -\det A$.
Chiaramente la cosa dipende dal fatto che la matrice ottenuta cancellando la prima colonna da A è uguale alla matrice ottenuta cancellando la seconda colonna da A' e ciò rimane vero ogni volta che scambiamo due colonne **adiacenti**.
Se ora si vuole vedere come cambia il determinante scambiando la prima con la terza colonna di A, è sufficiente osservare che tale scambio può essere ottenuto attraverso tre scambi di colonne adiacenti come mostra lo schema qui sotto

$$\begin{aligned} A = (A^1|A^2|A^3) &\xrightarrow{A^1\leftrightarrow A^2} (A^2|A^1|A^3) \xrightarrow{A^1\leftrightarrow A^3} \\ (A^2|A^3|A^1) &\xrightarrow{A^2\leftrightarrow A^3} (A^3|A^2|A^1) \end{aligned}$$

Poiché ad ogni scambio il determinante cambia di segno alla fine avrò

$$\det(A^3|A^2|A^1) = (-1)^3 \det A = -\det A\,.$$

Ora in generale lo scambio di due qualsiasi colonne si può ottenere attraverso un numero **dispari** di scambi di colonne adiacenti. Ciò mostra che scambiando due colonne il determinante cambia **sempre** di segno.

(3) **Il determinante della matrice A' ottenuta moltiplicando una colonna di A per un numero λ, è uguale a $\lambda \det A$.**

Verifichiamo questa proprietà in un esempio particolare. Sia $A = \begin{pmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{pmatrix}$ e consideriamo la matrice $A' = \begin{pmatrix} 1 & 4\sqrt{3} & 7 \\ 2 & 5\sqrt{3} & 8 \\ 3 & 6\sqrt{3} & 9 \end{pmatrix}$ ottenuta moltiplicando la seconda colonna di A per il numero $\sqrt{3}$. Dobbiamo confrontare $\det A$ e $\det A'$.
Sviluppando $\det A$ lungo la seconda colonna si ottiene

$$\det A = -4 \begin{vmatrix} 2 & 8 \\ 3 & 9 \end{vmatrix} + 5 \begin{vmatrix} 1 & 7 \\ 3 & 9 \end{vmatrix} - 6 \begin{vmatrix} 1 & 7 \\ 2 & 8 \end{vmatrix} .$$

Del resto sviluppando lungo la seconda colonna $\det A'$ abbiamo

$$\det A' = -4\sqrt{3} \begin{vmatrix} 2 & 8 \\ 3 & 9 \end{vmatrix} + 5\sqrt{3} \begin{vmatrix} 1 & 7 \\ 3 & 9 \end{vmatrix} - 6\sqrt{3} \begin{vmatrix} 1 & 7 \\ 2 & 8 \end{vmatrix} .$$

Raccogliendo a fattor comune $\sqrt{3}$ risulta evidente che $\det A' = \sqrt{3} \det A$.

(4) **Se la colonna A^j è la somma di due vettori C_1 e C_2, il determinante di A è uguale alla somma di determinanti delle matrici A' e A'' ottenute sostituendo la colonna A^j rispettivamente con C_1 e C_2.**
Anche in questo caso verifichiamo questa proprietà su un esempio. Consideriamo la matrice $A = \begin{pmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{pmatrix}$.
Osserviamo che la seconda colonna di A può essere scritta come la seguente somma di vettori:

$$\begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} = \begin{pmatrix} 7 \\ 2 \\ -6 \end{pmatrix} + \begin{pmatrix} -3 \\ 3 \\ 12 \end{pmatrix}$$

Consideriamo le matrici A' e A'' ottenute sostitutendo la seconda colonna di A con i due addendi $\begin{pmatrix} 7 \\ 2 \\ -6 \end{pmatrix}$, $\begin{pmatrix} -3 \\ 3 \\ 12 \end{pmatrix}$.
Si ha

$$A' = \begin{pmatrix} 1 & 7 & 7 \\ 2 & 2 & 8 \\ 3 & -6 & 9 \end{pmatrix} \qquad A'' = \begin{pmatrix} 1 & -3 & 7 \\ 2 & 3 & 8 \\ 3 & 12 & 9 \end{pmatrix} .$$

Sviluppando lungo la seconda colonna sia $\det A'$ che $\det A''$ si ottiene

$$\begin{aligned} \det A' &= - \quad 7 \cdot \begin{vmatrix} 2 & 8 \\ 3 & 9 \end{vmatrix} + \quad 2 \cdot \begin{vmatrix} 1 & 7 \\ 3 & 9 \end{vmatrix} - \quad (-6) \cdot \begin{vmatrix} 1 & 7 \\ 2 & 8 \end{vmatrix} \\ \det A'' &= - \ (-3) \cdot \begin{vmatrix} 2 & 8 \\ 3 & 9 \end{vmatrix} + \ 3 \cdot \begin{vmatrix} 1 & 7 \\ 3 & 9 \end{vmatrix} - \ (12) \cdot \begin{vmatrix} 1 & 7 \\ 2 & 8 \end{vmatrix} . \end{aligned}$$

Da cui deriviamo che

$$\det A' + \det A'' = \tag{3.23}$$
$$(-7+3)\cdot\begin{vmatrix}2 & 8\\3 & 9\end{vmatrix} + (2+3)\cdot\begin{vmatrix}1 & 7\\3 & 9\end{vmatrix} + (+6-12)\cdot\begin{vmatrix}1 & 7\\2 & 8\end{vmatrix},$$

ovvero

$$\det A' + \det A'' = -4\cdot\begin{vmatrix}2 & 8\\3 & 9\end{vmatrix} + 5\cdot\begin{vmatrix}1 & 7\\3 & 9\end{vmatrix} - 6\cdot\begin{vmatrix}1 & 7\\2 & 8\end{vmatrix},$$

ma questa espressione coincide proprio con lo sviluppo di $\det A$ lungo la seconda colonna.

(5) **Se una colonna di A è composta da tutti 0 allora il determinante di A è uguale a 0.**
Infatti per verificarlo basta sviluppare lungo quella colonna.

(6) **Se due colonne della matrice A sono uguali allora il determinante di A è uguale a 0.**
Supponiamo che la prima colonna sia uguale alla terza. Abbiamo visto, proprietà (2), che scambiando la prima e la terza colonna il determinante cambia di segno, Del resto poiché queste colonne sono uguali effettuando questo scambio ritrovo la matrice di partenza. Ne deduciamo che $\det A = -\det A$ e ciò è possibile solo se $\det A = 0$.

(7) **Se una colonna di A è proporzionale ad un'altra colonna allora il determinante di A è uguale a 0.**
Supponiamo ad esempio che $A^2 = \lambda A^1$. Allora si ha

$$\begin{aligned}\det(A^1|A^2|A^3|\dots|A^n) &= \det(A^1|\lambda A^1|A^3|\dots|A^n)\\ &= \lambda\det(A^1|A^1|\dots|A^n),\end{aligned}$$

dove nell'ultimo passaggio abbiamo utilizzato la proprietà (3). Ora la matrice $(A^1|A^1|\dots|A^n)$ contiene due colonne uguali e dunque il suo determinante è 0.

(8) **Se le colonne di A sono linearmente dipendenti allora il determinante di A è uguale a 0.**
Verifichiamo che il determinante si annulla nel caso in cui la terza colonna sia combinazione lineare delle prime due. Ovvero, supponiamo $A^3 = \lambda A^1 + \mu A^2$ allora utilizzando la proprietà (4) si ottiene

$$\begin{aligned}&\det(A^1|A^2|A^3|\dots|A^n)\\ &= \det(A^1|A^2|\lambda A^1|\dots|A^n) + \det(A^1|A^2|\mu A^2|\dots|A^n).\end{aligned}$$

Ora in entrambi gli addendi a destra di questa espressione la terza colonna è proporzionale ad un altra colonna (rispettivamente alla prima e alla seconda) e dunque i determinanti di tali matrici sono entrambi nulli per la proprietà (7).

Osserviamo che, nell'elenco precedente:

- tutte le proprietà da (2) a (8), enunciate per le colonne di una matrice, possono essere enunciate anche per le righe;
- la proprietà (8) ci dice quindi che le matrici quadrate **non invertibile** hanno sempre **determinante nullo**;
- possiamo enunciare la proprietà (8) anche nel seguente modo: se il determinante di una matrice quadrata **non è nullo**, allora le colonne sono linearmente indipendenti.

Abbiamo quindi il seguente risultato:

PROPOSIZIONE 3.15. *Se una matrice quadrata A ha determinante* $\det A \neq 0$, *allora A è invertibile.*

Nella prossima sezione verificheremo che vale anche il viceversa: una matrice invertibili ha determinante diverso da 0. Dunque il determinante di una matrice ci fornisce un criterio semplice per stabilire se la matrice è invertibile o meno.

6.3. La formula di Binet.

TEOREMA 3.16 (Formula di Binet). *Siano A, B matrici quadrate, vale l'uguaglianza:*

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

Prima di dimostrare il teorema, vogliamo elencare alcune conseguenze importanti del teorema di Binet.

(1) **Il determinante della potenza k-esima di una matrice A è la potenza k-esima del determinante di A**:

$$\det(A^k) = (\det A)^k\,.$$

Questa proprietà si ottiene applicando k volte il teorema di Binet. Si noti che questa formula è utile in quanto ci permette di calcolare il determinante di A^k senza dover calcolare esplicitamente A^k (il calcolo della potenza di una matrice è in generale un'operazione alquanto dispendiosa), ma semplicemente elevando alla k il determinante di A.

(2) **Se A è invertibile allora** $\det A \neq 0$ **e** $\det A^{-1} = \frac{1}{\det A}$.
Essendo A invertibile, esiste la matrice inversa A^{-1}. Calcoliamo il determinante del prodotto AA^{-1} in due modi diversi. Da una parte, poiché sappiamo che $AA^{-1} = I_n$ si ha

$$\det(AA^{-1}) = \det I_n = 1\,.$$

Dall'altra applicando la regola di Binet, si ha

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

e dunque deduciamo che

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

Da questa uguaglianza verifichiamo subito che $\det A$ non può essere uguale a 0, e inoltre $\det A^{-1} = \frac{1}{\det A}$.

(3) **Se A e B sono matrici quadrate di ordine $n \times n$, allora si ha**:
$$\det(AB) = \det(BA)\,.$$
Infatti abbiamo
$$\det(AB) = \det A \det B \qquad\qquad \det(BA) = \det B \det A\,,$$
e poiché il determinante è una grandezza numerica il prodotto $\det A \det B$ è uguale al prodotto $\det B \det A$.

(4) **Se A è una matrice invertibile, allora si ha**:
$$\det(ABA^{-1}) = \det B\,.$$
Infatti si ha
$$\det(ABA^{-1}) = \det A \det B \det A^{-1} = \det A \det B \frac{1}{\det A} = \det B\,.$$

Come conseguenza del Teorema di Binet abbiamo quindi il seguente criterio operativo per stabilire quando una matrice quadrata è invertibile:

Proposizione 3.17. *Una matrice quadrata A di ordine $n \times n$ è invertibile se e solo se* $\det A \neq 0$.

Possiamo ora provare il teorema.

Dimostrazione del Teorema di Binet.

Fissata una matrice A di ordine $n \times n$, possiamo associare ad ogni matrice B la grandezza
$$d(B) = \det(AB)\,.$$
Per prima cosa dimostreremo che la grandezza $d(B)$ soddisfa proprietà simili a quelle del determinante. Utilizzando poi queste proprietà mostreremo che si ha l'uguaglianza $d(B) = \det A \det B$.

Lemma 3.18. *La grandezza $d(B)$ soddisfa le seguenti proprietà:*

- *Se B' è la matrice ottenuta scambiando due colonne di B, allora si ha: $d(B') = -d(B)$.*
- *Se una colonna B^j di B è una combinazione lineare di certi vettori, $B^j = \lambda_1 \boldsymbol{u}_1 + \lambda_2 \boldsymbol{u}_2$, allora*
$$d(B) = \lambda_1 d(B') + \lambda_2 d(B'')$$
dove B' e B'' sono le matrici ottenute sostituendo in B la colonna B^j rispettivamente con i vettori $\boldsymbol{u}_1, \boldsymbol{u}_2$.

Dimostrazione:

Verifichiamo la prima proprietà. Suppiamo per semplicità che la matrice B' sia ottenuta scambiando la prima e la seconda colonna di B, ovvero
$$B = (B^1|B^2|B^3|\dots|B^n) \qquad B' = (B^2|B^1|B^3|\dots|B^n)\,.$$

Allora abbiamo

$$AB = (AB^1|AB^2|AB^3|\dots|AB^n) \qquad AB' = (AB^2|AB^1|AB^3|\dots|AB^n)$$

ovvero la matrice AB' si ottiene scambiando la prima con la seconda colonna nella matrice AB. Per cui abbiamo

$$d(B') = \det(AB') = -\det(AB) = -d(B)\,.$$

Verifichiamo la seconda proprietà. Sempre per semplicità supponiamo

$$B^1 = \lambda_1 \boldsymbol{u}_1 + \lambda_2 \boldsymbol{u}_2$$

allora

$$AB^1 = A(\lambda_1 \boldsymbol{u}_1 + \lambda_2 \boldsymbol{u}_2) = \lambda_1 A\boldsymbol{u}_1 + \lambda_2 A\boldsymbol{u}_2\,.$$

Poiché $AB = (AB^1|AB^2|\dots|AB^n)$ risulta che

$$d(B) = \det(AB) = \lambda_1 \det(A\boldsymbol{u}_1|AB^2|\dots|AB^n) + \lambda_2 \det(A\boldsymbol{u}_2|AB^2|\dots|AB^n)\,.$$

Del resto abbiamo che $B' = (\boldsymbol{u}_1|B^2|\dots|B^n)$ per cui

$$d(B') = \det(AB') = \det(A\boldsymbol{u}_1|AB^2|\dots|AB^n)\,,$$

analogamente

$$d(B'') = \det(A\boldsymbol{u}_2|AB^2|\dots|AB^n)$$

da cui discende facilmente l'uguaglianza cercata.

Osserviamo che come nel caso del determinante abbiamo queste semplici conseguenze delle proprietà enunciate

- Se due colonne di B sono uguali allora $d(B) = 0$;
- Se una data colonna B^j è combinazione lineare di certi vettori

  $$B^j = \lambda_1 \boldsymbol{u}_1 + \lambda_2 \boldsymbol{u}_2 + \dots + \lambda_k \boldsymbol{u}_k$$

  allora abbiamo

  $$d(B) = \lambda_1 d(B_{(1)}) + \lambda_2 d(B_{(2)}) + \dots + \lambda_k d(B_{(k)})$$

  dove $B_{(1)}, B_{(2)}, \dots, B_{(k)}$ sono le matrici ottenute sostituendo al posto di B^j i vettori $\boldsymbol{u}_1, \dots, \boldsymbol{u}_k$.

Utilizzando ora le proprietà espresse nel Lemma si verifica abbastanza facilmente la formula di Binet. Per semplicità di notazione ci limiteremo a verificarla nel caso di matrici 3×3. Ciò non di meno il tipo di ragionamento che viene proposta può essere generalizzata facilmente nel caso di matrici di ordine arbitrario.

Dimostrazione della formula di Binet.
Verificheremo che

$$d(B) = \det A \det B \tag{3.24}$$

per passi successivi.

Caso 1: Verifichiamo la formula per $B = I_3$.

In questo caso $d(I_3) = \det(AI_3) = \det A$ mentre $\det A \det I_3 = \det A$ per cui l'uguaglianza è verificata.

Caso 2: Verifichiamo la formula (3.24) per le matrici B le cui colonne sono elementi della base canonica, ovvero B è della forma $(\boldsymbol{e}_i|\boldsymbol{e}_j|\boldsymbol{e}_k)$.

Abbiamo due casi: o una colonna di B è ripetuta oppure le colonne di B sono esattamente $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$ (ma non necessariamente in ordine). Nel primo caso abbiamo $d(B) = \det B = 0$ per cui la formula è verificata.

Nel secondo caso osserviamo che la matrice B può essere ottenuta scambiando ripetutamente le colonne della matrice identità. Ad esempio se B è la matrice $(\boldsymbol{e}_3, \boldsymbol{e}_1, \boldsymbol{e}_2)$, si osservi che la si ottiene partendo dall'identita con la seguente sequenza di scambi:

$$(\boldsymbol{e}_1|\boldsymbol{e}_2|\boldsymbol{e}_3) \to (\boldsymbol{e}_2|\boldsymbol{e}_1|\boldsymbol{e}_3) \to (\boldsymbol{e}_3|\boldsymbol{e}_1|\boldsymbol{e}_2)\,.$$

Ora ad ogni scambio sia d che det cambiano di segno per cui si ha

$$\det(B) = (-1)^\sigma \qquad d(B) = (-1)^\sigma \det A$$

dove σ è il numero di scambi necessari per ottenere la matrice B da I_n. Dunque in questo caso si ha ancora $d(B) = \det A \det B$.

Caso 3: Verifichiamo la formula (3.24) per le matrici B in cui le prime due colonne sono elementi della base canonica. In questo caso B è della forma $(\boldsymbol{e}_i|\boldsymbol{e}_j|\boldsymbol{u})$.

Poiché $\boldsymbol{u} = u_1\boldsymbol{e}_1 + u_2\boldsymbol{e}_2 + u_3\boldsymbol{e}_3$ si ha:

$$d(B) = u_1 d(\boldsymbol{e}_i|\boldsymbol{e}_j|\boldsymbol{e}_1) + u_2 d(\boldsymbol{e}_i|\boldsymbol{e}_j|\boldsymbol{e}_2) + u_3 d(\boldsymbol{e}_i|\boldsymbol{e}_j|\boldsymbol{e}_3)\,.$$

Osserviamo che le matrici che compaiono a destra sono tutte del tipo precedente per cui abbiamo

$$d(B) = u_1 \det A \det(\boldsymbol{e}_i|\boldsymbol{e}_j|\boldsymbol{e}_1) + u_2 \det A \det(\boldsymbol{e}_i|\boldsymbol{e}_j|\boldsymbol{e}_2) + u_3 \det A \det(\boldsymbol{e}_i|\boldsymbol{e}_j|\boldsymbol{e}_3)\,.$$

Raccogliendo $\det A$ si ottiene

$$d(B) = \det A\big(u_1 \det(\boldsymbol{e}_i|\boldsymbol{e}_j|\boldsymbol{e}_1) + u_2 \det(\boldsymbol{e}_i|\boldsymbol{e}_j|\boldsymbol{e}_2) + u_3 \det(\boldsymbol{e}_i|\boldsymbol{e}_j|\boldsymbol{e}_3)\big)\,.$$

Del resto

$$\det(B) = \det(\boldsymbol{e}_i|\boldsymbol{e}_j|\boldsymbol{u}) = u_1 \det(\boldsymbol{e}_i|\boldsymbol{e}_j|\boldsymbol{e}_1) + u_2 \det(\boldsymbol{e}_i|\boldsymbol{e}_j|\boldsymbol{e}_2) + u_3 \det(\boldsymbol{e}_i|\boldsymbol{e}_j|\boldsymbol{e}_3)$$

e dunque segue facilmente che $d(B) = \det A \det B$ per questo tipo di matrici.

Caso 4: Verifichiamo la formula (3.24) per le matrici B in cui la prima colonna è un elemento della base canonica, ovvero B della forma $(\boldsymbol{e}_i|\boldsymbol{w}|\boldsymbol{u})$.

In questo caso consideriamo la decomposizione della seconda colonna

$$\boldsymbol{w} = w_1\boldsymbol{e}_1 + w_2\boldsymbol{e}_2 + w_3\boldsymbol{e}_3$$

da cui deduciamo che

$$d(B) = w_1 d(\boldsymbol{e}_i|\boldsymbol{e}_1|\boldsymbol{u}) + w_2 d(\boldsymbol{e}_i|\boldsymbol{e}_2|\boldsymbol{u}) + w_3 d(\boldsymbol{e}_i|\boldsymbol{e}_3|\boldsymbol{u})\,.$$

Ora tutte le matrici al secondo membro sono del tipo precedente, per cui sappiamo che vale la formula e deduciamo che

$$d(B) = \det A(w_1 \det(\boldsymbol{e}_i|\boldsymbol{e}_1|\boldsymbol{u}) + w_2 \det(\boldsymbol{e}_i|\boldsymbol{e}_2|\boldsymbol{u}) + w_3 \det(\boldsymbol{e}_i|\boldsymbol{e}_3|\boldsymbol{u})\,.$$

Del resto si ha

$$\det B = \det(\boldsymbol{e}_i|\boldsymbol{w}|\boldsymbol{u}) = w_1 \det(\boldsymbol{e}_i|\boldsymbol{e}_1|\boldsymbol{u}) + w_2 \det(\boldsymbol{e}_i|\boldsymbol{e}_2|\boldsymbol{u}) + w_3 \det(\boldsymbol{e}_i|\boldsymbol{e}_3|\boldsymbol{u})\,,$$

per cui è ancora vero che $d(B) = \det A \det B$.

Caso 5: Verifichiamo la formula (3.24) per B matrice qualsiasi, ovvero della forma $(\boldsymbol{v}|\boldsymbol{w}|\boldsymbol{u})$.

In questo caso, considerando la decomposizione della prima colonna

$$\boldsymbol{v} = v_1\boldsymbol{e}_1 + v_2\boldsymbol{e}_2 + v_3\boldsymbol{e}_3$$

abbiamo

$$d(B) = v_1 d(\boldsymbol{e}_1|\boldsymbol{w}|\boldsymbol{u}) + v_2 d(\boldsymbol{e}_2|\boldsymbol{w}|\boldsymbol{u}) + v_3 d(\boldsymbol{e}_3|\boldsymbol{w}|\boldsymbol{u})\,.$$

Ora tutte le matrici al secondo membro sono del tipo precedente, per cui sappiamo che vale la formula e deduciamo che

$$d(B) = \det A\big(v_1 \det(\boldsymbol{e}_i|\boldsymbol{w}|\boldsymbol{u}) + v_2 \det(\boldsymbol{e}_2|\boldsymbol{w}|\boldsymbol{u}) + v_3 \det(\boldsymbol{e}_3|\boldsymbol{w}|\boldsymbol{u}))\,.$$

Del resto si ha

$$\det B = \det(\boldsymbol{v}|\boldsymbol{w}|\boldsymbol{u}) = v_1 \det(\boldsymbol{e}_1|\boldsymbol{w}|\boldsymbol{u}) + v_2 \det(\boldsymbol{e}_2|\boldsymbol{w}|\boldsymbol{u}) + w_3 \det(\boldsymbol{e}_3|\boldsymbol{w}|\boldsymbol{u})$$

per cui è ancora vero che $d(B) = \det A \det B$.

6.4. Calcolo dell'inversa con la formula di Cramer.

Abbiamo visto che il determinante permette di caratterizzare le matrici invertibili (sono quelle con determinante non nullo). La formula di Cramer ci permette di calcolare esplicitamente attraverso l'uso del determinante la matrice inversa di una matrice A invertibile.

PROPOSIZIONE 3.19 (Formula di Cramer per l'inversa). *Sia A una matrice quadrata di ordine $n \times n$ invertibile. Allora la matrice inversa di A è la seguente:*

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

dove $\alpha_{ij} = (-1)^{(i+j)}|A_{[j,i]}|$.

DIMOSTRAZIONE.
Ci limiteremo a dimostrare la formula per matrici di ordine 3. In tal caso la candidata inversa di A è la matrice

$$\frac{1}{\det A}\begin{pmatrix} |A_{[1,1]}| & -|A_{[2,1]}| & |A_{[3,1]}| \\ -|A_{[1,2]}| & |A_{[2,2]}| & -|A_{[3,2]}| \\ |A_{[1,3]}| & -|A_{[2,3]}| & |A_{[3,3]}| \end{pmatrix}.$$

Consideriamo la matrice

$$B = \begin{pmatrix} |A_{[1,1]}| & -|A_{[2,1]}| & |A_{[3,1]}| \\ -|A_{[1,2]}| & |A_{[2,2]}| & -|A_{[3,2]}| \\ |A_{[1,3]}| & -|A_{[2,3]}| & |A_{[3,3]}| \end{pmatrix}$$

Il seguente lemma è la chiave per la dimostrazione della Formula di Cramer

LEMMA 3.20. *Dato un qualsiasi $X \in \mathbb{R}^3$ e posto $Y = BX$ si ha che*

$$\begin{cases} y_1 = \det(X|A^2|A^3) \\ y_2 = \det(A^1|X|A^3) \\ y_3 = \det(A^1|A^2|X) \end{cases}$$

DIMOSTRAZIONE.
Osserviamo che

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} |A_{[1,1]}| & -|A_{[2,1]}| & |A_{[3,1]}| \\ -|A_{[1,2]}| & |A_{[2,2]}| & -|A_{[3,2]}| \\ |A_{[1,3]}| & -|A_{[2,3]}| & |A_{[3,3]}| \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

per cui si ha

$$y_1 = x_1|A_{[1,1]}| - x_2|A_{[2,1]}| + x_3|A_{[3,1]}|\ .$$

Del resto sviluppando lungo la prima colonna si ha

$$\det(X|A^2|A^3) = \begin{vmatrix} x_1 & a_{12} & a_{13} \\ x_2 & a_{22} & a_{23} \\ x_3 & a_{32} & a_{33} \end{vmatrix} = x_1|A_{[1,1]}| - x_2|A_{[2,1]}| + x_3|A_{[3,1]}|$$

e quindi chiaramente

$$y_1 = \det(X|A^2|A^3)\,.$$

Analogamente si verificano le altre uguaglianze cercate. □

Dimostriamo ora la Regola di Cramer enunciata sopra. Calcoliamo il prodotto BA. Si ha

$$BA = (BA^1|BA^2|BA^3)\,.$$

Del resto per il precedente lemma, posto $Y = BA^1$ si ha

$$y_1 = \det(A^1|A^2|A^3) = \det A \quad y_2 = \det(A^1|A^1|A^3) = 0 \quad y_3 = \det(A^1|A^2|A^1) = 0$$

e dunque $BA^1 = \begin{pmatrix} \det A \\ 0 \\ 0 \end{pmatrix}$. Analogamente si mostra che

$$BA^2 = \begin{pmatrix} 0 \\ \det A \\ 0 \end{pmatrix}, \quad BA^3 = \begin{pmatrix} 0 \\ 0 \\ \det A \end{pmatrix}$$

da cui segue che

$$BA = \begin{pmatrix} \det A & 0 & 0 \\ 0 & \det A & 0 \\ 0 & 0 & \det A \end{pmatrix} = \det A \cdot I_3$$

da cui si ricava facilmente che $A^{-1} = \frac{1}{\det A} B$. □

Osservazione 3.26. Nel caso di una matrice di ordine 2 la formula di Cramer dà un modo molto veloce per calcolare l'inversa. Infatti se $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ allora si ha

$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}\,.$$

Infatti in questo caso abbiamo

$$\begin{aligned} \alpha_{11} &= |A_{[1,1]}| = a_{22} = d \\ \alpha_{12} &= -|A_{[2,1]}| = -a_{12} = -b \\ \alpha_{21} &= -|A_{[1,2]}| = -a_{2,1} = -c \\ \alpha_{22} &= |A_{[2,2]}| = a_{11} = a \end{aligned}$$

Anche se apparentemente la formula di Cramer sembra indicare una via molto semplice per il calcolo dell'inversa, nella pratica questo metodo risulta computazionalmente molto dispendioso, in quanto è necessario calcolare il determinante di molte matrici. Già per una matrice 3×3 conviene calcolare l'inversa come esposto nella Sezione 3.

Ciò non di meno tale formula risulta essere particolarmente utile per il calcolo dell'inversa di matrici parametriche.

Esempio 3.27.
Consideriamo la seguente matrice A dipendente dal parametro h:

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

Si verifica facilmente che $\det A = 2h^2$ e dunque A è invertibile per ogni $h \neq 0$. L'inversa di A per $h \neq 0$ è la matrice:

$$A^{-1} = \frac{1}{2h^2} \begin{pmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{pmatrix}$$

dove i coefficienti α_{ij} sono dati dalla formula di Cramer. Abbiamo $\alpha_{11} = |A_{[1,1]}|$ e dunque mettendo in evidenza la sottomatrice $A_{[1,1]}$ in A si ha

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

da cui risulta facilmente $\alpha_{11} = h^2$.

Ora $\alpha_{12} = -|A_{[2,1]}|$ e dunque mettendo in evidenza la sottomatrice $A_{[2,1]}$ si ha

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

da cui $\alpha_{12} = h$. Procedendo in questo modo alla fine si ottiene

$$A^{-1} = \frac{1}{2h^2} \begin{pmatrix} h^2 & h & -(1+h) \\ 0 & 0 & 2h \\ h^2 & -h & (1-h) \end{pmatrix} .$$

6.5. Utilizzo delle proprietà del determinante per il calcolo.

Il calcolo del determinante utilizzando lo sviluppo per righe o per colonne è in generale un'operazione molto dispendiosa. Essa risulta semplice come abbiamo visto nel caso di matrici triangolari, ma nel caso generico il calcolo del determinante implica un gran numero di operazioni. Si può calcolare che per trovare il determinante di una matrice generica 10×10 utilizzando lo sviluppo lungo una riga bisogna effettuare più di 8.000.000 moltiplicazioni!

L'osservazione cruciale che ci permette di semplificare notevolmente il calcolo del determinante è la seguente.

Lemma 3.21. *Sia A una matrice $n \times n$ e una A' la matrice ottenuta sommando ad una colonna di A un multiplo di un'altra colonna. Allora, risulta:*

$$\det A' = \det A .$$

L'idea è allora quella di sommare in modo opportuno ad una colonna/riga di una matrice un multiplo di un'altra colonna/riga in modo da ottenere matrici più semplici. Siccome facendo questo tipo di operazioni il determinante non cambia, invece di calcolare il determinante di A si calcola il determinante della matrice semplificata A'.

Esempio 3.28.

(1) Calcoliamo il seguente determinante

$$d = \begin{vmatrix} 1 & 0 & 1 & 3 \\ 1 & 1 & 1 & 2 \\ 1 & 2 & 2 & 4 \\ 1 & 3 & 1 & 3 \end{vmatrix}.$$

Sottraendo alla terza colonna la prima colonna si ottiene

$$d = \begin{vmatrix} 1 & 0 & 0 & 3 \\ 1 & 1 & 0 & 2 \\ 1 & 2 & 1 & 4 \\ 1 & 3 & 0 & 3 \end{vmatrix}.$$

Sviluppando il determinante della matrice così ottenuta sulla terza colonna risulta

$$d = \begin{vmatrix} 1 & 0 & 3 \\ 1 & 1 & 2 \\ 1 & 3 & 3 \end{vmatrix}.$$

Sottraendo ora alla terza riga la prima risulta

$$d = \begin{vmatrix} 1 & 0 & 3 \\ 1 & 1 & 2 \\ 0 & 3 & 0 \end{vmatrix},$$

da cui sviluppando lungo l'ultima riga si ha

$$d = -3 \begin{vmatrix} 1 & 3 \\ 1 & 2 \end{vmatrix} = 3\,.$$

(2) Calcoliamo il seguente determinante:

$$d = \begin{vmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 3 \\ 3 & 4 & 3 & 2 \\ 4 & 3 & 2 & 1 \end{vmatrix}.$$

Sottraiamo alla seconda riga la prima riga moltiplicata per 2

$$d = \begin{vmatrix} 1 & 2 & 3 & 4 \\ 0 & -1 & -2 & -5 \\ 3 & 4 & 3 & 2 \\ 4 & 3 & 2 & 1 \end{vmatrix}.$$

Ora sottraiamo alla terza riga la prima riga moltiplicata per 3

$$d = \begin{vmatrix} 1 & 2 & 3 & 4 \\ 0 & -1 & -2 & -5 \\ 0 & -2 & -6 & -10 \\ 4 & 3 & 2 & 1 \end{vmatrix}.$$

Infine sottraiamo alla quarta riga la prima moltiplicata per 4

$$d = \begin{vmatrix} 1 & 2 & 3 & 4 \\ 0 & -1 & -2 & -5 \\ 0 & -2 & -6 & -10 \\ 0 & -5 & -10 & -15 \end{vmatrix}.$$

Dunque sviluppando lungo la prima colonna risulta

$$d = \begin{vmatrix} -1 & -2 & -5 \\ -2 & -6 & -10 \\ -5 & -10 & -15 \end{vmatrix},$$

portando fuori un -1 da ciascuna colonna si ottiene

$$d = (-1)(-1)(-1) \begin{vmatrix} 1 & 2 & 5 \\ 2 & 6 & 10 \\ 5 & 10 & 15 \end{vmatrix} = - \begin{vmatrix} 1 & 2 & 5 \\ 2 & 6 & 10 \\ 5 & 10 & 15 \end{vmatrix}.$$

Sottraendo alla terza riga la prima moltiplicata per 5 si ottiene infine

$$d = - \begin{vmatrix} 1 & 2 & 5 \\ 2 & 6 & 10 \\ 0 & 0 & -10 \end{vmatrix},$$

da cui sviluppando lungo la terza riga

$$d = 10 \begin{vmatrix} 1 & 2 \\ 2 & 6 \end{vmatrix} = 20\,.$$

6.6. Interpretazione geometrica della funzione determinante.

Nel caso di matrici 2×2 e di matrici 3×3 c'è un'importante interpretazione geometrica del determinante che ora vogliamo brevemente descrivere.

Consideriamo una matrice 2×2:

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

e indichiamo con $\boldsymbol{u}_1$ e $\boldsymbol{u}_2$ le sue colonne.

Possiamo considerare la rappresentazione polare dei vettori $\boldsymbol{u}_1, \boldsymbol{u}_2 \in \mathbb{E}^2_O$:

$$\begin{cases} a = \rho_1 \cos\theta_1 \\ c = \rho_1 \sin\theta_1 \end{cases} \qquad \begin{cases} b = \rho_2 \cos\theta_2 \\ d = \rho_2 \sin\theta_2 \end{cases}$$

dove $\rho_1 = |\boldsymbol{u}_1|$ e θ_1 è l'angolo che $\boldsymbol{u}_1$ forma con il vettore $\boldsymbol{e}_1$, mentre $\rho_2 = |\boldsymbol{u}_2|$ e θ_2 è l'angolo che $\boldsymbol{u}_2$ forma con il vettore $\boldsymbol{e}_1$.

Allora abbiamo:

$$\det A = ad - bc = \rho_1\rho_2(\cos\theta_1 \sin\theta_2 - \sin\theta_1 \cos\theta_2) = \rho_1\rho_2 \sin(\theta_2 - \theta_1)\,.$$

Osserviamo che $|\theta_1 - \theta_2|$ coincide con l'angolo φ compreso tra $\boldsymbol{u}_1$ e $\boldsymbol{u}_2$, ed il segno di $\theta_1 - \theta_2$ è positivo se $\boldsymbol{u}_2$ si ottiene ruotando $\boldsymbol{u}_1$ di φ in senso antiorario, è negativo altrimenti.

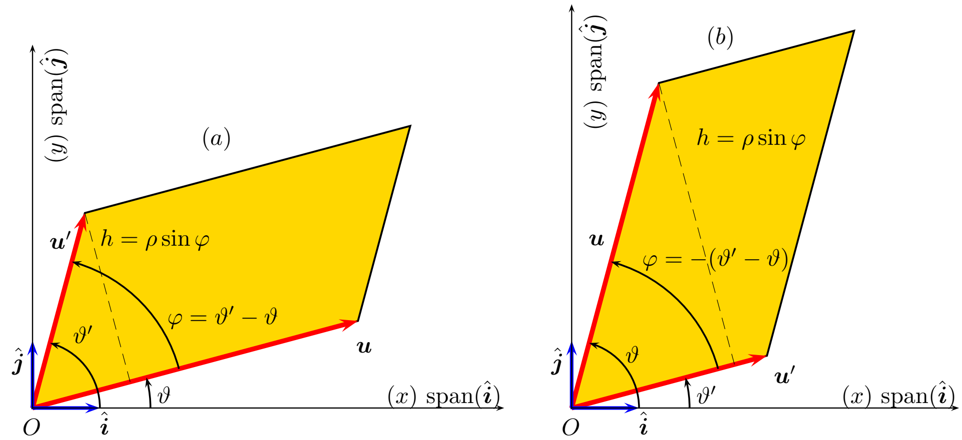

FIGURA 3.3. I due vettori $\boldsymbol{u}_1$ ed $\boldsymbol{u}_2$ nel piano cartesiano formano un angolo φ, positivo. A seconda dell'orientamento reciproco fra i vettori, si avrà $\varphi = \vartheta' - \vartheta$ (a), oppure $\varphi = -(\vartheta' - \vartheta)$ (b).

In particolare abbiamo che il determinante di A è positivo se $\boldsymbol{u}_2$ si ottiene ruotando $\boldsymbol{u}_1$ di angolo φ in senso antiorario, è nullo se $\varphi = k\pi$ (nel qual caso $\boldsymbol{u}_1$ e $\boldsymbol{u}_2$ hanno la stessa direzione), è negativo negli altri casi (quando cioè $\boldsymbol{u}_2$ è ottenuto ruotando $\boldsymbol{u}_1$ di φ in senso orario).

Inoltre il valore assoluto di $\det A$ è uguale a $\rho_1 \rho_2 \sin \varphi$.

Osservando che $|\rho_2 \sin \varphi|$ è l'altezza del parallelogramma individuato dai vettori $\boldsymbol{u}_1$ e $\boldsymbol{u}_2$, deduciamo che:
il valore assoluto di $\det A$ è uguale all'area del parallelogramma individuato dai vettori $\boldsymbol{u}_1$ e $\boldsymbol{u}_2$.

Si può verificare che vale un risultato simile per le matrici 3×3. Se indichiamo con $\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3$ le colonne di una matrice A di ordine 3, si ha che:
il valore assoluto di $\det A$ è uguale al volume del parallelepipedo individuato dai vettori $\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3$.

7. Il rango di una matrice

DEFINIZIONE 3.18. Sia A una matrice di odine $k \times n$. Il **rango** di A è la dimensione del sottospazio di $\mathbb{R}^k$ generato dalle colonne di A:

$$\mathrm{rg}(A) = \dim \ \mathrm{Span}(A^1, A^2, \dots, A^n)\,.$$

Nota Bene: $\mathrm{rg}(A) = 0$ se e solo se A è la matrice nulla di ordine $k \times n$!!

ESEMPIO 3.29.
Consideriamo la matrice

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

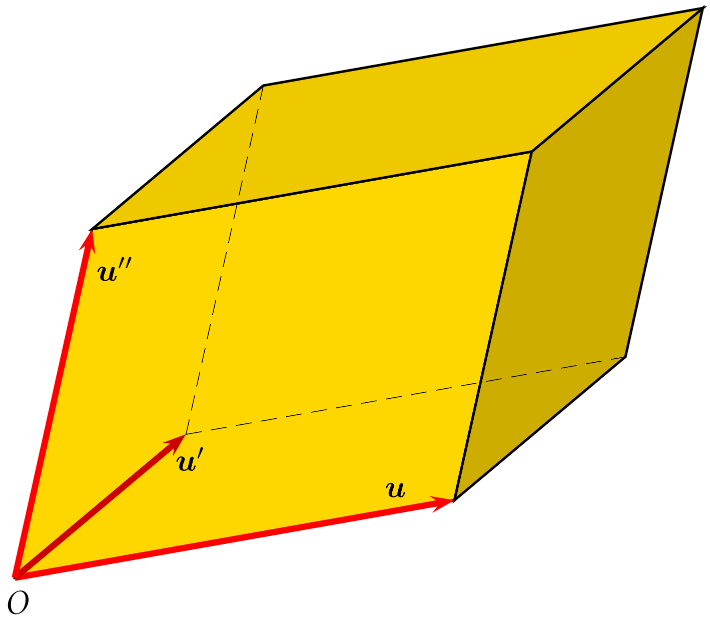

FIGURA 3.4. Tre vettori $\boldsymbol{u}_1$, $\boldsymbol{u}_2$ e $\boldsymbol{u}_3$ non complanari individuano un parallelepipedo; il volume del paralellepipedo si ottiene calcolando il determinante di una matrice di ordine 3 le cui colonne sono le componenti nella base canonica dei tre vettori, rispettivamente. Il volume ha un segno legato all'orientamento reciproco dei tre vettori. In accordo con le regole di calcolo del determinante, se i vettori sono complanari (ossia, se uno è nel sottospazio generato dagli altri due), il parallelepipedo degenera in una figura piana a volume nullo; corrispondentemente, il determinante si annulla, poiché una colonna è combinazione lineare delle altre due (punto (8) delle Proprietà 3.14).

Lo spazio generato dalle colonne di A è $W = \operatorname{Span}\left(A^1, A^2, A^3\right)$, dove:

$$A^1 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, A^2 = \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}, A^3 = \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix}.$$

Dunque per calcolare la dimensione di tale spazio è sufficiente applicare l'algoritmo di estrazione alla lista delle colonne di A.

Osserviamo che A^1 e A^2 sono indipendenti, mentre A^3 è combinazione lineare di A^1 e A^2:

$$A^3 = A^1 + A^2.$$

Dunque le colonne A^1, A^2 formano una base di W, per cui si ha

$$\operatorname{rg} A = \dim W = 2\,.$$

PROPRIETÀ 3.22. Osserviamo che, data una matrice A di ordine $k \times n$, si hanno le seguenti proprietà:

(1) **Il rango di A è il massimo numero di colonne linearmente indipendenti di A.**
Il rango di una matrice si può quindi calcolare applicando l'algoritmo di estrazione alle colonne della matrice A.

(2) **Il rango di A è minore o uguale al numero di colonne di A**:
$$\mathrm{rg}(A) \leq n.$$

(3) **Il rango di A è minore o uguale al numero di righe di A:**
$$\mathrm{rg}(A) \leq k.$$
Infatti le colonne di A sono elementi di $\mathbb{R}^k$, quindi $\mathrm{Span}(A^1, \ldots, A^n)$ è un sottospazio di $\mathbb{R}^k$, la sua dimensione è $\leq k$.

(4) **Il rango di A è minore o uguale al minimo tra gli interi n e k**:
$$\mathrm{rg}(A) \leq min(k, n)\,.$$

(5) **Il rango di A è uguale a k se e soltanto se le colonne di A sono un sistema di generatori di $\mathbb{R}^k$.**
Infatti si ha $\dim \mathrm{Span}(A^1, \ldots, A^n) = k$ se e solo se si ha:
$$\mathrm{Span}(A^1, \ldots, A^n) = \mathbb{R}^k.$$

(6) **Il rango di A è uguale a n se e soltanto se le colonne di A sono linearmente indipendenti**.
Infatti $\dim \mathrm{Span}\left(A^1, A^2, \cdots, A^n\right) = n$ se e solo se le colonne $A^1, \ldots, A^n$ sono indipendenti.

(7) **Se A' si ottiene da A scambiando due colonne**:
$$\mathrm{rg}(A') = \mathrm{rg}(A).$$
Infatti lo spazio vettoriale generato dalle colonne non dipende dall'ordine in cui le colonne appaiono.

Possiamo riassumere le osservazioni precedenti nello schema seguente:

Caso 1	$n \leq k$:	il rango massimo è n. $\mathrm{rg}(A) = n \quad \Rightarrow$ le colonne di A sono indipendenti (e generano un sottospazio di $\mathbb{R}^k$ di dimensione n.)
Caso 2:	$k \leq n$:	il rango massimo è k. $\mathrm{rg}(A) = k \quad \Rightarrow$ le colonne di A sono un sistema di generatori di $\mathbb{R}^k$.
Caso 3	$k = n$:	il rango massimo è n. $\mathrm{rg}(A) = n \quad \Rightarrow$ le colonne di A sono una base di $\mathbb{R}^n$.

7.1. Rango e minori di una matrice.

Osservazione 3.30. Sia A una matrice quadrata di ordine $n \times n$. Il rango di A è massimo (cioè $\operatorname{rg}(A) = n$) se e solo se le colonne di A sono linearmente indipendenti e quindi se e solo se la matrice A è invertibile (ovvero $\det A \neq 0$). Otteniamo quindi:

$$\operatorname{rg}(A) = n \iff \det A \neq 0.$$

Questa semplice osservazione mostra come vi sia una relazione tra determinante e rango. In questa sezione ci occuperemo di generalizzare e precisare questa relazione mostrando come il rango di una qualsiasi matrice A (non necessariamente quadrata) dipenda dai determinanti delle sottomatrici quadrate di A.

Definizione 3.19. Sia A una matrice di ordine $k \times n$.

- Una **sottomatrice** di A è una matrice A' che è possibile ottenere cancellando righe e colonne di A.
- Il determinante di ogni sottomatrice quadrata di A viene chiamato **minore** della matrice.
- Diremo che un minore Δ è di ordine h se è il determinante di una sottomatrice $h \times h$.

Notazione 3.20.
Utilizzeremo la notazione $A' \subset A$ per indicare che A' è una sottomatrice di A.

Esempio 3.31.

(1) Consideriamo la seguente matrice:

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

Le sottomatrici di ordine 1×1 sono semplicemente le entrate della matrice e i minori corrispondenti sono le entrate stesse. Dunque i minori di ordine 1 sono $1, 2, 3, 4, 5, 6$.
Le sottomatrici di ordine 2×2 si ottengono cancellando una colonna di A. Dunque ci sono esattamente 3 sottomatrici di ordine 2×2:

$$\begin{pmatrix} 2 & 3 \\ 5 & 6 \end{pmatrix}, \quad \begin{pmatrix} 1 & 3 \\ 4 & 6 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 \\ 4 & 5 \end{pmatrix}.$$

I minori di ordine 2 di A sono dunque

$$\begin{vmatrix} 2 & 3 \\ 5 & 6 \end{vmatrix} = -3, \quad \begin{vmatrix} 1 & 3 \\ 4 & 6 \end{vmatrix} = -6, \quad \begin{vmatrix} 1 & 2 \\ 4 & 5 \end{vmatrix} = -3.$$

(2) Consideriamo la seguente matrice:

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

Come prima i minori di ordine 1 sono semplicemente le entrate della matrice e dunque sono i numeri da 1 a 9.
Le sottomatrici 2×2 si ottengono cancellando una riga e una colonna da A. Quindi in totale sono 9 e in particolare sono:

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

Dunque i minori di ordine 2 sono:

$$\begin{vmatrix} 5 & 6 \\ 7 & 8 \end{vmatrix} = -2\,, \quad \begin{vmatrix} 4 & 6 \\ 9 & 8 \end{vmatrix} = -22\,, \quad \begin{vmatrix} 4 & 5 \\ 9 & 7 \end{vmatrix} = -17$$
$$\begin{vmatrix} 2 & 3 \\ 7 & 8 \end{vmatrix} = -5\,, \quad \begin{vmatrix} 1 & 3 \\ 9 & 8 \end{vmatrix} = -19\,, \quad \begin{vmatrix} 1 & 2 \\ 9 & 7 \end{vmatrix} = -11,$$
$$\begin{vmatrix} 2 & 3 \\ 5 & 6 \end{vmatrix} = -3\,, \quad \begin{vmatrix} 1 & 3 \\ 4 & 6 \end{vmatrix} = -6\,, \quad \begin{vmatrix} 1 & 2 \\ 4 & 5 \end{vmatrix} = -3\,.$$

L'unica sottomatrice 3×3 è la matrice A stessa e dunque l'unico minore di ordine 3 è $|A| = -9$.

(3) Consideriamo la seguente matrice

$$A = \begin{pmatrix} 1 & 15 & \sqrt{3} \\ 2 & 16 & 2\sqrt{3} \\ 3 & 13 & 3\sqrt{3} \\ 4 & 15 & 4\sqrt{3} \end{pmatrix}.$$

Osserviamo che $A^3 = \sqrt{3}A^1$. Ora se A' è una sottomatrice A di ordine 3 (ovvero se A' è ottenuta cancellando una riga di A), la prima e la terza colonna di A' sono ancora una multipla dell'altra per lo stesso fattore di proporzionalità (in questo esempio $\sqrt{3}$). In particolare il determinante di A' è 0. Segue che tutti i minori di ordine 3 sono nulli.

(4) Consideriamo il seguente esempio

$$A = \begin{pmatrix} 1 & \sqrt{2} & 1+\sqrt{2} \\ 2 & \sqrt{2} & 2+\sqrt{2} \\ 3 & \sqrt{2} & 3+\sqrt{2} \\ 4 & \sqrt{2} & 4+\sqrt{2} \end{pmatrix}.$$

Si ha $A^3 = A^1 + A^2$. In questo caso per le sottomatrici 3×3, $A' \subset A$, sarà vero che $(A')^1 + (A')^2 = (A')^3$ e dunque anche in questo caso si avrà $\det A' = 0$, per ogni sottomatrice $A' \subset A$ di ordine 3×3.

Chiaramente, gli esempi 3 e 4 sono casi particolari di una proprietà più generale. Supponiamo che la colonna n-esima di A sia combinazione lineare delle altre colonne

$$A^n = \lambda_1 A^1 + \lambda_2 A^2 + \ldots + \lambda_{n-1} A^{n-1}\,.$$

Data una sottomatrice A' di A ottenuta cancellando alcune righe di A (**ma nessuna colonna**) si avrà ancora

$$(A')^n = \lambda_1 (A')^1 + \lambda_2 (A')^2 + \ldots + \lambda_{n-1} (A')^{n-1} .$$

In particolare se A' è quadrata (cosa che può avvenire solo se il numero delle righe è maggiore del numero delle colonne) si avrà $\det A' = 0$.

Possiamo riassumere questa osservazione nella seguente relazione tra rango e minori che generalizza la relazione osservata all'inizio nel caso di matrici quadrate.

PROPRIETÀ 3.23. *Sia A una matrice $k \times n$ con $n \leq k$.*

- *Se le colonne di A sono linearmente dipendenti, allora tutti i minori di A di ordine n sono nulli.*
- *Se esiste un minore di A di ordine n non nullo, allora le colonne di A sono linearmente indipendenti e* $\operatorname{rg} A = n$.

La proprietà 3.23 che abbiamo ricavato sopra si applica **SOLO** nel caso in cui la matrice A contenga **più righe che colonne**. Inoltre, ci fornisce un criterio (una condizione necessaria e sufficiente) per stabilire quando il rango è massimo.

Vediamo ora come è possibile applicare tale criterio nel caso generale per ottenere informazioni sul rango dai minori della matrice. Come sempre iniziamo con un esempio.

ESEMPIO 3.32. Consideriamo la matrice di ordine 4×5:

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

Consideriamo la sottomatrice A' ottenuta scegliendo le colonne A^1, A^3, A^5 e le righe A_1, A_3, A_4 (e dunque cancellando le colonne A^2, A^4 e la riga A_2) Abbiamo

$$A = \begin{pmatrix} 1 & 2 & 3 & 1 & 1 \\ 1 & 1 & 2 & 7 & 2 \\ 0 & 1 & 4 & 1 & 1 \\ 0 & 1 & 4 & -1 & -1 \end{pmatrix} \quad \to \quad A' = \begin{pmatrix} 1 & 3 & 1 \\ 0 & 4 & 1 \\ 0 & 4 & -1 \end{pmatrix} .$$

Osserviamo che A' è anche una sottomatrice della matrice A'' formata dalle colonne A^1, A^3, A^5 (che sono le colonne che contengono A'):

$$A = \begin{pmatrix} 1 & 2 & 3 & 1 & 1 \\ 1 & 1 & 2 & 7 & 2 \\ 0 & 1 & 4 & 1 & 1 \\ 0 & 1 & 4 & -1 & -1 \end{pmatrix} \quad \to \quad A'' = \begin{pmatrix} 1 & 3 & 1 \\ 1 & 2 & 2 \\ 0 & 4 & 1 \\ 0 & 4 & -1 \end{pmatrix} .$$

Poiché il determinante di A' è $-8 \neq 0$ applicando la Proprietà 3.23 deduciamo che le colonne A^1, A^3, A^5 sono indipendenti.

Ovviamente questo ragionamento si può applicare ogni volta che ho un minore non nullo. Ne deduciamo che vale la seguente proprietà che generalizza quella già studiata.

Proprietà 3.24. *Sia A una matrice di ordine $k \times n$.*

- *Se $A' \subset A$ è una sottomatrice quadrata di ordine r con determinante non nullo, allora le colonne di A che contengono le colonne di A' sono linearmente indipendenti.*
- *In particolare,* $\operatorname{rg} A \geq r$.

Possiamo enunciare la proprietà nel seguente modo:
se esiste un minore di A di ordine r non nullo, allora $\operatorname{rg}(A) \geq r$.

ATTENZIONE! Nel caso in cui abbiamo un minore nullo, non abbiamo alcuna informazione sul rango.

Per ottimizzare l'informazione sul rango che possiamo ricavare guardando i minori, dobbiamo allora cercare il minore di ordine più grande non nullo.

Definizione 3.21. Sia A una matrice di ordine $k \times n$. Definiamo l' ordine massimo dei minori non nulli di A:

$$r_{MAX}(A) = \max\{r \mid \text{esiste un minore di } A \text{ di ordine } r \text{ non nullo}\}.$$

Nota bene: $r_{MAX}(A) = r$ significa che:

- **esiste un minore non nullo** di A di ordine r,
- **tutti** i minori di ordine $> r$ sono nulli.

Applicando la Proprietà 3.24, abbiamo la relazione: $\operatorname{rg}(A) \geq r_{MAX}(A)$. Mostreremo in effetti che vale l'uguaglianza.

Proposizione 3.25. *Sia A una matrice di ordine $k \times n$, vale:*

$$\operatorname{rg}(A) = r_{MAX}(A).$$

Dimostrazione.
Poniamo $r = r_{MAX}(A)$ e supponiamo per semplicità che il minore di ordine r non nullo sia quello ottenuto scegliendo le prime r righe e le prime r colonne (e cancellando tutto il resto). Sappiamo che le prime r colonne di A sono indipendenti, vogliamo mostrare che tutte le colonne $A^{r+1}, \ldots, A^n$ sono combinazioni lineari di $A^1, \ldots, A^r$. Ci limiteremo a verificare che A^{r+1} è combinazione lineare delle prime r colonne, in quanto lo stesso ragionamento può essere applicato alle altre colonne.

Sia A' la matrice ottenuta scegliendo le prime r righe e le prime r colonne e sia $\boldsymbol{u}$ il vettore di $\mathbb{R}^r$ ottenuto considerando le prime r entrate della colonna A^{r+1}. In tal modo abbiamo:

$$A = \left(\begin{array}{ccc|c|c} & & & & * \ldots * \\ & A' & & \boldsymbol{u} & \vdots \ddots \vdots \\ & & & & * \ldots * \\ \hline * & \ldots & * & * & * \ldots * \\ \vdots & \ddots & \vdots & \vdots & \vdots \ddots \vdots \\ * & \ldots & * & * & * \ldots * \end{array}\right).$$

Poiché $\det A' \neq 0$ sappiamo che $(A')^1, \ldots, (A')^r$ formano una base di $\mathbb{R}^r$. In particolare il vettore $\boldsymbol{u}$ è combinazione lineare di $(A')^1, \ldots, (A')^r$

$$\boldsymbol{u} = \lambda_1 (A')^1 + \lambda_2 (A')^2 + \ldots + \lambda_r (A')^r.$$

Sia ora C la corrispondente combinazione lineare dei vettori $A^1, \ldots, A^r$

$$C = \lambda_1 A^1 + \lambda_2 A^2 \ldots + \lambda_r A^r \,.$$

Osserviamo che le prime componenti di C coincidono con le componenti di $\boldsymbol{u}$ e dunque C è della forma

$$C = \begin{pmatrix} \boldsymbol{u} \\ \\ * \\ \vdots \\ * \end{pmatrix}$$

e in particolare le prime r componenti di C coincidono con le prime r componenti di A^{r+1}:

$$A^{r+1} - C = \begin{pmatrix} \mathbf{0}_r \\ \\ * \\ \vdots \\ * \end{pmatrix}$$

Ora se mostriamo che tutte le componenti di $A^{r+1} - C$ sono nulle (ovvero $A^{r+1} - C = \mathbf{0}_k$) allora avremmo $A^{r+1} = C$ e dunque A^{r+1} risulterebbe combinazione lineare delle prime r colonne come volevamo.

Posto $B = A^{r+1} - C$, dobbiamo dunque verificare che $b_{r+1} = b_{r+2} = \ldots = b_k = 0$. Osserviamo che:

- Applicando le proprietà del determinante si osserva facilmente che i minori di ordine $r+1$ della matrice $(A^1|A^2|\ldots A^r|A^{r+1})$ coincidono con i minori di ordine $r+1$ della matrice $(A^1|A^2|\ldots|A^r|A^{r+1} - C)$.
- Poiché per ipotesi la matrice A non contiene minori di ordine $r+1$ non nulli ricaviamo tutti i minori di ordine $r+1$ della matrice $(A^1|A^2|\ldots|A^r|A^{r+1})$ sono nulli.

Le due osservazioni implicano che la matrice

$$(A^1|A^2|\ldots|A^r|B)$$

non contiene minori di ordine $r+1$ non nulli. Ora prendiamo il minore corrispondente alla sottomatrice $(r+1) \times (r+1)$ ottenuta scegliendo le prime $r+1$ righe (e le prime $r+1$ colonne) si ha

$$\left| \begin{array}{c|c} A' & \mathbf{0}_r \\ \hline * \ldots * & b_{r+1} \end{array} \right| = b_{r+1} \det A' \,.$$

Ponendo dunque tale minore uguale a 0 otteniamo

$$b_{r+1} \det A' = 0 \,,$$

da cui deduciamo $b_{r+1} = 0$, visto che per ipotesi abbiamo $\det A' \neq 0$.

Consideriamo il minore corrispondente la sottomatrice ottenuta scegliendo le prime r righe e la riga $r+2$. In questo caso si ha

$$\left| \begin{array}{c|c} A' & \mathbf{0}_r \\ \hline * \ldots * & b_{r+2} \end{array} \right| = b_{r+2} \det A' \,,$$

da cui come prima ricaviamo $b_{r+2} = 0$ e così via. $\square$

Osservazione 3.33. Sia A una matrice di ordine $k \times n$.
Osserviamo che se A' è una sottomatrice di A allora $(A')^{\mathrm{T}}$ è una sottomatrice di A^{T}. Poiché $\det A' = \det(A')^{\mathrm{T}}$, ricaviamo che i minori di A coincidono con i minori di $(A)^{\mathrm{T}}$. In particolare, abbiamo:

$$r_{MAX}(A) = r_{MAX}(A^{\mathrm{T}}).$$

Dalla Proposizione 3.25 ricaviamo la seguente relazione non banale che lega il rango di A al rango della trasposta di A:

Corollario 3.26. *Sia A una matrice di ordine $k \times n$. Allora si ha:*

$$\mathrm{rg}(A) = \mathrm{rg}(A^{\mathrm{T}}).$$

Detto in altri termini, in una matrice A il numero massimo di colonne indipendenti, coincide con il numero massimo di righe indipendenti.

Nel prossimo capitolo vedremo un'altra dimostrazione più teorica dello stesso risultato.

Per concludere questa sezione descriviamo un metodo algoritmico per calcolare $r_{MAX}(A)$ detto **metodo degli orlati**.

Definizione 3.22. Sia A una matrice di ordine $k \times n$. Data una sottomatrice $A' \subset A$ di ordine r, un **orlato** di A' è una sottomatrice A'' di ordine $r+1$ tale che $A' \subset A'' \subset A$.

Esempio 3.34.
Consideriamo la matrice A di ordine 3×4:

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{pmatrix},$$

sia $A' \subset A$ la sottomatrice di ordine 2 ottenuta scegliendo la prima e la seconda riga e la prima e la terza colonna

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{pmatrix} \quad \longrightarrow \quad A' = \begin{pmatrix} 1 & 3 \\ 5 & 7 \end{pmatrix}.$$

Allora la sottomatrice A'' di ordine 3 ottenuta scegliendo la prima la seconda la terza riga, e la prima e la seconda e la terza colonna

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{pmatrix} \quad \longrightarrow \quad A'' = \begin{pmatrix} 1 & 2 & 3 \\ 5 & 6 & 7 \\ 9 & 10 & 11 \end{pmatrix}$$

è un orlato di A'.

Lo studente può verificare che la matrice A' ha esattamente 2 orlati.

Proposizione 3.27 (Regola degli orlati di Kronecker). *Sia A una matrice di ordine $k \times n$. Sia A' una sottomatrice di A di ordine r con determinante non nullo. Se tutti gli orlati di A hanno determinante nullo allora* $\mathrm{rg}(A) = r$.

Utilizzando tale criterio il calcolo del rango di una matrice A può avvenire in questo modo:

- Si fissa una sottomatrice $A_{(1)}$ di ordine 1 (ovvero un'entrata) non nulla.
- Si considerano tutti gli orlati di tale sottomatrice. Se i corrispondenti minori sono tutti nulli si deduce che il rango di A è 1. Altrimenti si fissa un'orlato $A_{(2)}$ con determinante diverso da 0.
- Si considerano tutti gli orlati di $A_{(2)}$. Se i corrispondenti minori sono tutti nulli, si deduce che il rango di A è 2. Altrimenti si fissa un orlato $A_{(3)}$ con determinante diverso da 0
- Si considerano tutti gli orlati di $A_{(3)}$. Se i corrispondenti minori sono tutti nulli si deduce che il rango di A è 3. Altrimenti si fissa un orlato $A_{(4)}$ con determinante non nullo.
- Si ripete ricorsivamente la procedura esposta sopra fino ad arrivare al calcolo del rango

Vediamo quache esempio di applicazione del metodo degli orlati appena descritto, che costituisce un approccio *"bottom-up"*.

Esercizio 7.1.

(1) Si determini il rango della seguente matrice:

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

Anzitutto, notiamo che la matrice è quadrata di ordine 3×3; quindi diversi approcci sono adottabili. Se vogliamo attenerci all'uso del metodo degli orlati, si dovrebbe partire da un minore di ordine 1 non nullo; questo c'è: $\Delta_{(1)} = (a_{1,2}) = 1 \neq 0$ va bene; il rango è almeno 1. Ora, cominciamo ad orlare: consideriamo la sottomatrice $A_{(2)}$ ottenuta orlando l'elemento considerato con la prima colonna e la seconda riga (ossia, $A_{(2)} = A_{[3,3]}$:

$$\begin{pmatrix} 0 & 1 & 2 \\ -1 & 0 & 1 \\ -2 & -1 & 0 \end{pmatrix} \Rightarrow \qquad A_{(2)} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Risulta ovvio che $\det(A_{(2)}) = 1$, quindi il rango è almeno 2. Ora, l'unica sottomatrice che possiamo costruire orlando $A_{(2)}$ è la matrice A stessa. Calcoliamone il determinante:

$$\det(A) = \begin{vmatrix} 0 & 1 & 2 \\ -1 & 0 & 1 \\ -2 & -1 & 0 \end{vmatrix} = -(1)\begin{vmatrix} -1 & 1 \\ -2 & 0 \end{vmatrix} + 2\begin{vmatrix} -1 & -1 \\ -2 & -1 \end{vmatrix} = -2 + 2 = 0.$$

Quindi, possiamo affermare che "tutti i minori di ordine 3 ottenuti orlando $A_{(2)}$ sono nulli", quindi $\text{rg}(A) = 2$. In effetti, notiamo che $2A^2 - A^1 = A^3$.

Si determini il rango della matrice di ordine 4×4:
$$B = \begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \\ 0 & -1 & -1 & 0 \\ 1 & 0 & 0 & -1 \end{pmatrix}$$
. Non tutte le entrate sono nulle, quindi cominciamo subito con la sottomatrice delle prime due righe e colonne:
$$B_{(2)} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} .$$
Questa ha chiaramente determinante $\det(B_{(2)}) = 1$, quindi $\mathrm{rg}(B) \geq 2$. Osserviamo che orlando la matrice $B_{(2)}$ otteniamo in tutto 4 sottomatrici; orlando con la terza riga e terza colonna si ha
$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & -1 & -1 \end{pmatrix}$$
che ha determinante nullo, perché ha due colonne uguali.

Orlando con la quarta riga e terza colonna si ha
$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \end{pmatrix} ,$$
ancora con determinante nullo, perché ha due righe uguali.

Le altre 2 sottomatrici usando la quarta colonna, con la terza e la quarta riga sono:
$$\begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & -1 & 0 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & -1 \end{pmatrix}$$
e hanno determinante nullo, perché entrambe hanno due colonne proporzionali.

Ossia, tutti le sottomatrici di ordine 3 ottenute orlando $B_{(2)}$ hanno determinante nullo, quindi, per la Proposizione (3.27) possiamo fermarci e dire che $\mathrm{rg}(B) = 2$. Il risultato, in realtà era ovvio (*perché?*).

7.2. Rango di matrici con parametri.

Il metodo degli orlati (3.27) visto sopra è meno efficace in presenza di uno o più parametri nella matrice. In questo caso, conviene un approccio *"top-down"*, ossia "dall'alto". Vediamo un esmepio (per ulteriori esempi, rimandiamo al Capitolo finale 9, dove riportiamo una selezione di esercizi proposti e, in parte, risolti.

Esercizio 7.2. Determinare il rango della matrice
$$M_h = \begin{pmatrix} h-1 & h & 1 & -h \\ 0 & 1 & h & -1 \\ 1-h & 2+h & 1 & -1 \end{pmatrix}$$

al variare del parametro h in $\mathbb{R}$.

La matrice ha ordine 3×4, quindi il suo rango massimo è 3. Consideriamo la sottomatrice formata dalle ultime 3 colonne, e calcoliamone il determinante (la scelta cade sulle colonne che sembrano più "semplici" dal punto di vista dei termini parametrici: si lascia l'esercizio di scegliere anche, per esempio le prime 3 colonne, secondo il procedimento che descriveremo):

$$\begin{aligned}\det(M'_h) &= \begin{vmatrix} h & 1 & -h \\ 1 & h & -1 \\ 2+h & 1 & -1 \end{vmatrix} = h^3 + h^2 - h - 1 = h^2(h+1) - (h+1) \\ &= (h^2-1)(h+1) = (h-1)(h+1)^2\,.\end{aligned}$$

Si vede subito che, per $h \neq -1, 1$, $\det(M'_h) \neq 0$ e, quindi, $\operatorname{rg}(M_h) = 3$.

Rimane da determinare $\operatorname{rg}(M_h)$ per i valori $h = 1$ e $h = -1$. Conviene studiare "a mano" la matrice M_h per ciascuno di tali valori. Per $h = -1$ otteniamo:

$$M_{-1} = \begin{pmatrix} -2 & -1 & 1 & 1 \\ 0 & 1 & -1 & -1 \\ 2 & 1 & 1 & -1 \end{pmatrix}\,.$$

Il rango di questa matrice è massimo; per verificare l'affermazione, basta calcolare il determinante della sottomatrice delle prime 3 colonne:

$$\begin{vmatrix} -2 & -1 & 1 \\ 0 & 1 & -1 \\ 2 & 1 & 1 \end{vmatrix} = \begin{vmatrix} -2 & -1 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 2 \end{vmatrix} = 2 \begin{vmatrix} -2 & -1 \\ 0 & 1 \end{vmatrix} = -4\,,$$

ossia $\operatorname{rg}(M_{-1}) = 3$.

Invece, per $h = 1$, la matrice M assume la forma:

$$M_{+1} = \begin{pmatrix} 0 & 1 & 1 & -1 \\ 0 & 1 & 1 & -1 \\ 0 & 3 & 1 & -1 \end{pmatrix}\,.$$

e, osservando che le prime due righe sono uguali, è facile convincersi che $\operatorname{rg}(M_{+1}) = 2$: la sottomatrice evidenziata in colore ha chiaramente determinante non nullo.

APPENDICE

8. Proprietà associativa per il prodotto fra matrici

Siano X e Y due matrici di ordini, rispettivamente, $k \times n$ e $n \times h$ (non necessariamente con $k \neq n$ e $n \neq h$), ossia due matrici per le quali possiamo calcolare il prodotto XY. Lasciamo allo studente il compito di verificare che quanto detto per il prodotto righe per colonne si riassume brevemente dicendo che, detta P la matrice di ordine $k \times h$ prodotto l'entrata generica della riga i-esima e della colonna j-esima di P si calcolano nella seguente maniera

$$(P)_{ij} = p_{ij} = (XY)_{ij} = \sum_{l=1}^{n} x_{il} y_{lj} \,. \tag{3.25}$$

Usando ripetutamente la scrittura 3.25 calcoliamo esplicitamente l'entrata della riga i-esima e della colonna j-esima del prodotto $A(BC)$ per le matrici aventi gli stessi ordini usati nella sezione 2.3:

$$[A(BC)]_{ij} = \sum_{l=1}^{n} a_{il} (BC)_{lj} = \sum_{l=1}^{n} a_{il} \left(\sum_{m=1}^{h} b_{lm} c_{mj} \right) . \tag{3.26}$$

Ora, le operazioni di addizione e moltiplicazione nel campo dei numeri reali –in cui sono prese le entrate delle matrici– godono entrambe della proprietà associativa; inoltre, ricordiamo che l'addizione gode di quella commutativa e la moltiplicazione di quella distributiva sull'addizione (la proprietà commutativa del prodotto qui non serve). Sfruttando tutto ciò, possiamo riordinare i termini nella (3.26) per ottenere:

$$\begin{aligned}
[A(BC)]_{ij} &= \sum_{l=1}^{n} \left(\sum_{m=1}^{h} a_{il} b_{lm} c_{mj} \right) && \text{(p. distributiva)} \\
&= \left(\sum_{l=1}^{n} \sum_{m=1}^{h} a_{il} b_{lm} c_{mj} \right) && \text{(p. associativa addiz.)} \\
&= \left(\sum_{m=1}^{h} \sum_{l=1}^{n} a_{il} b_{lm} c_{mj} \right) && \text{(p. commutativa addiz.)} \\
&= \sum_{m=1}^{h} \left(\sum_{l=1}^{n} a_{il} b_{lm} c_{mj} \right) && \text{(p. associativa addiz.)} \\
&= \sum_{m=1}^{h} \left(\sum_{l=1}^{n} a_{il} b_{lm} \right) c_{mj} && \text{(p. distributiva moltipl.)} \\
&= \sum_{m=1}^{h} (AB)_{im} c_{mj} && \text{(definiz. prdotto fra matrici)} \\
&= [(AB)C]_{ij} \,.
\end{aligned} \tag{3.27}$$

Procedendo in maniera analoga, si possono verificare anche tutte le altre proprietà del prodotto fra matrici. Si lascia allo studente l'esercizio di effettuare tali verifiche, per impratichirsi anche con le tecniche di utilizzo della notazione di sommatoria.

9. Teoremi sul calcolo del determinante

Abbiamo accennato, nella presentazione del Teorema di Laplace 3.12, al risultato preliminare circa il calcolo del determinante secondo la prima riga, con formula analoga a quella secondo la prima colonna. Lo enunciamo qui esplicitamente, dandone la dimostrazione: questa è un po' tecnica, ma vale la pena di vederla qui per chi vuole impratichirsi con l'uso della notazione di sommatoria.

TEOREMA 3.28. *Sia $A \in \mathcal{M}_{\mathbb{R}}(n)$ una matrice quadrata di ordine n (in generale, una matrice $A \in \mathcal{M}_{\mathbb{K}}(n)$ a entrate in un campo $\mathbb{K}$ qualunque); Il determinante di A si può calcolare anche secondo la formula seguente:*

$$\begin{aligned} \det(A) &= \sum_{i=1}^{n} (-1)^{i+1} a_{1,i} \det(A_{[1,i]}) \\ &= a_{1,1} \det(A_{[1,1]}) - a_{1,2} \det(A_{[1,2]}) + \dots (-1)^{n+1} a_{1,n} \det(A_{[1,n]}) \end{aligned} \tag{3.28}$$

DIMOSTRAZIONE. La dimostrazione verrà fatta utilizzando il principio di induzione.

Se $n = 1$, l'enunciato è vero; poiché la matrice contiene un solo elemento, si ha $\det(A) = |(a_{11})| = a_{11}$, e questo è vero sia che il calcolo si svolga secondo la prima riga o secondo la prima colonna: in entrambi i casi, tutto si riduce ad un solo elemento.

Supponiamo ora (*ipotesi induttiva*) di avere dimostrato il teorema per le matrici di ordine $n - 1$, con $n > 1$. Calcoliamo il determinante di A secondo la definizione (sviluppando secondo la prima colonna), e separiamo il primo termine della sommatoria dagli altri:

$$\begin{aligned} \det(A) &= \sum_{i=1}^{n} (-1)^{i+1} a_{i,1} \det(A_{[i,1]}) \\ &= a_{11} \det(A_{[1,1]}) + \sum_{i=2}^{n} (-1)^{i+1} a_{i,1} \det(A_{[i,1]}) \,. \end{aligned} \tag{3.29}$$

Osserviamo adesso che le matrici aggiunte $A_{[i,1]}$ sono di ordine $n-1$, quindi per l'ipotesi induttiva possiamo calcolare il loro determinante anche sviluppandolo secondo la loro prima riga.

Inoltre, osserviamo anche che l'elemento $a_{1,j}$ della matrice originaria A si trova nella prima riga e nella colonna $(j - 1)-$esima avendo cancellato la prima colonna. Analogamente, se si considera la matrice $A_{[1,j]}$, in questa l'elemento $a_{i,1}$ della matrice originaria A si trova nella prima colonna e nella riga $(i - 1)-$esima avendo cancellato la prima riga.

Possiamo scrivere, pertanto:

$$\begin{aligned}\det(A) &= a_{11}\det(A_{[1,1]}) + \sum_{i=2}^{n}(-1)^{i+1}a_{i,1}\,\det(A_{[i,1]})\\ &= a_{11}\det(A_{[1,1]}) + \sum_{i=2}^{n}(-1)^{i+1}a_{i,1}\sum_{k=1}^{n-1}(-1)^{1+k}a_{1,k+1}\det(A_{[1,i;\,1,(k+1)]})\\ &= a_{11}\det(A_{[1,1]}) + \sum_{i=2}^{n}(-1)^{i+1}a_{i,1}\sum_{j=2}^{n}(-1)^{j}a_{1,j}\det(A_{[1,i;\,1,j]})\\ &= a_{11}\det(A_{[1,1]}) + \sum_{i=2}^{n}\sum_{j=2}^{n}(-1)^{1+i+j}a_{i,1}a_{1,j}\det(A_{[1,i;\,1,j]})\,,\end{aligned}\tag{3.30}$$

dove abbiamo denotato con $A_{[1,i;\,1,j]}$ la matrice di ordine $(n-1)$ ottenuta cancellando le righe 1 e i e le colonne 1 e j.

Ma anche, sviluppando secondo la prima riga, e poi sviluppando secondo la prima colonna i determinanti delle matrici di ordine $(n-1)\times(n-1)$:

$$\begin{aligned}\det(A) &= a_{11}\det(A_{[1,1]}) + \sum_{j=2}^{n}(-1)^{j+1}a_{1,j}\,\det(A_{[1,j]})\\ &= a_{11}\det(A_{[1,1]}) + \sum_{j=2}^{n}(-1)^{j+1}a_{1,j}\sum_{k=1}^{n-1}(-1)^{1+k}a_{k+1,1}\det(A_{[1,(k+1);\,1,j]})\\ &= a_{11}\det(A_{[1,1]}) + \sum_{j=2}^{n}(-1)^{j+1}a_{1,j}\sum_{i=2}^{n}(-1)^{i}a_{i,1}\det(A_{[1,i;\,1,j]})\\ &= a_{11}\det(A_{[1,1]}) + \sum_{j=2}^{n}\sum_{i=2}^{n}(-1)^{1+i+j}a_{1,j}a_{i,1}\det(A_{[1,i;\,1,j]})\,.\end{aligned}\tag{3.31}$$

Nella seconda espressione, riordinando i termini si ottiene l'espressione dell'equazione (3.30), che è quanto volevamo dimostrare.

□

Ora, la proprietà del determinante della trasposta (punto (1) delle Proprietà 3.14 diventa una conseguenza abbastanza immediata.

COROLLARIO 3.29. *Sia $A \in \mathcal{M}_{\mathbb{R}}(n)$ una matrice quadrata di ordine n (in generale, una matrice $A \in \mathcal{M}_{\mathbb{K}}(n)$ a entrate in un campo $\mathbb{K}$ qualunque). Il determinante della matrice trasposta coincide con quello della matrice di partenza.*

DIMOSTRAZIONE. Sfruttando la definizione di determinante data in 3.14, possiamo procedere per induzione sull'ordine n della matrice.

Se $n = 1$, l'enunciato è vero; poiché la matrice contiene un solo elemento, si ha $A^{\mathrm{T}} = (a_{11})^{\mathrm{T}} = (a_{11}) = A$, e, quindi $\det(A) = \det(A^{\mathrm{T}}) = a_{11}$. Supponiamo adesso che il teorema sia vero per matrici di ordine $n-1$ (con $n > 1$).

Chiamiamo B la trasposta di A, cioè $B := A^{\mathrm{T}}$; sarà $b_{ij} = a_{ji}$. Scriviamo ora il determinante di B seguendo la definizione:

$$\det(B) = \sum_{i=1}^{n}(-1)^{i+1}b_{i,1}\,\det(B_{[i,1]}) = \sum_{i=1}^{n}(-1)^{i+1}b_{i,1}\,\det(A^{\mathrm{T}}_{[1,i]})\,,$$

dove abbiamo usato il fatto che B è la trasposta di A, quindi la matrice aggiunta $B_{i,1}$, che viene dalla cancellazione della riga $i-$esima e della prima colonna di B, si può ottenere cancellando la prima **riga** di A, l'$i-$esima **colonna** di A, e prendendo la trasposta di ciò che resta.

Ora, $A_{1,i}$ è di ordine $n-1$, quindi per *l'ipotesi induttiva* $\det(A_{[1,i]}) = \det(A^{\mathrm{T}}_{[1,i]})$; alllora,

$$\begin{aligned} \det(A^{\mathrm{T}}) = \det(B) &= \sum_{i=1}^{n} (-1)^{i+1} b_{i,1} \, \det(A_{[1,i]}) \\ &= \sum_{i=1}^{n} (-1)^{i+1} a_{1,i} \, \det(A_{[1,i]}) \\ &= \det(A)\,, \end{aligned} \tag{3.32}$$

sfruttando lo sviluppo del determinante secondo la prima riga riportato nell'Eq. (3.28) conicide con quello sviluppato secondo la prima colonna, secondo il Teorema 3.28; questo è proprio ciò che volevamo dimostrare. □

La formula generale per il calcolo del determinante, dovuta a Laplace, presenta una forma semplice una volta introdotta la seguente

Definizione 3.23. Sia $A \in \mathcal{M}_{\mathbb{K}}(n)$; viene chiamato *complemento algebrico* $A^*_{i,j}$ dell'elemento a_{ij} della matrice il determinante della matrice aggiunta $A_{[i,j]}$ se $i+j$ è pari, l'opposto del determinante della matrice aggiunta $A_{[i,j]}$ se $i+j$ è dispari. In altre parole:

$$A^*_{i,j} = (-1)^{i+j} \, \det(A_{[i,j]})\,. \tag{3.33}$$

A questo punto possiamo enunciare il

Teorema 3.30. *Il determinante di una matrice $A \in \mathcal{M}_{\mathbb{K}}(n)$ può essere calcolato secondo la $j-$esima colonna usando la formula*

$$\det(A) = \sum_{k=1}^{n} (-1)^{k+j} a_{k,j} \, \det(A_{[k,j]}) = \sum_{k=1}^{n} a_{k,j} \, A^*_{k,j}\,, \tag{3.34}$$

ossia, il determinante è dato dalla somma dei prodotti degli elementi della colonna per i relativi complementi algebrici.

Osservazione 3.35. Come detto precedentemente, il Teorema (3.12) ci permette di sviluppare il determinante anche secondo la $i-$esima riga, e, analogamente al caso precedente, il determinante di A è dato dalla somma dei prodotti degli elementi della riga per i relativi complementi algebrici.

Per calcolare il determinante siamo dunque liberi di scegliere la riga o la colonna che ci sembra più conveniente (pur di fare attenzione ai segni).

10. Prodotto vettoriale e prodotto misto in $\mathbb{E}^3_O$

Nello spazio $\mathbb{E}^3_O$ dei vettori applicati è possibile definire un particolare prodotto fra vettori, che si aggiunge al prodotto scalare (1.12) definito nel Capitolo 1.

Ricordiamo che, dati due vettori $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{E}^3_O$ il prodotto scalare è dato da

$$\langle \boldsymbol{u}, \boldsymbol{v} \rangle = |\boldsymbol{u}|\,|\boldsymbol{v}| \cos\vartheta =: u\, v \cos\vartheta\,,$$

dove ϑ è l'angolo convesso fra i vettori (Figura 1.14 nel Capitolo 1).

Riprendiamo in considerazione la configurazione geometrica che corrisponde alla somma fra due vettori generici $\boldsymbol{u}$ e $\boldsymbol{v}$. Usando la base canonica di $\mathbb{E}^3_O$,

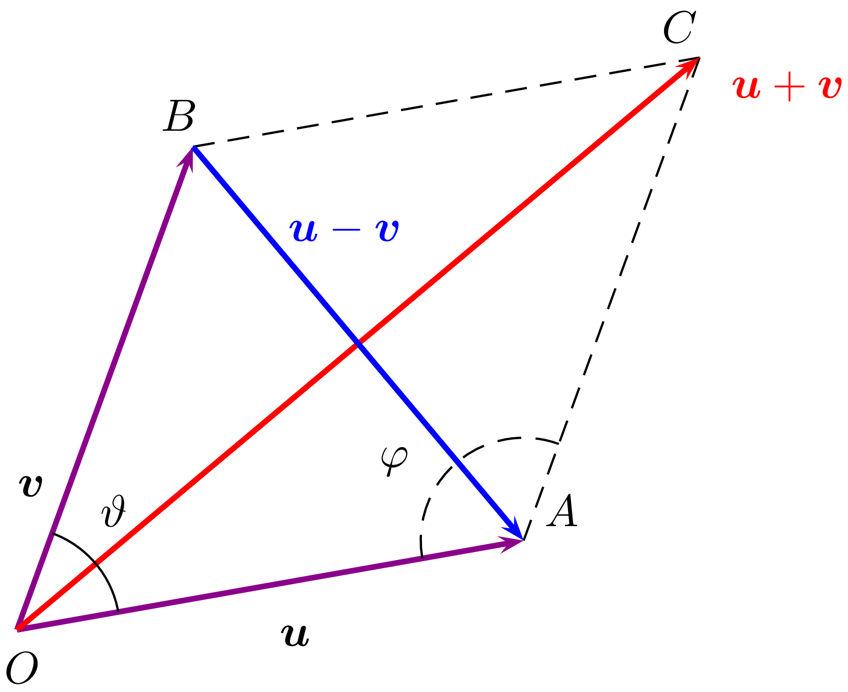

FIGURA 3.5. I vettori $\boldsymbol{u}$ e $\boldsymbol{v}$ si compongono secondo il parallelogramma $OACB$ per dare $\boldsymbol{u}+\boldsymbol{v}$; l'angolo φ fra i lati OA ed AC è il supplementare dell'angolo convesso ϑ fra i due vettori, ossia: $\varphi = \pi - \vartheta$.

$\mathcal{B} = \{\hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}, \hat{\boldsymbol{k}}\}$, rappresentiamo i due vettori mediante le loro coordinate:

$$\boldsymbol{u} = \begin{pmatrix} u_x \\ u_y \\ u_z \end{pmatrix} \qquad \boldsymbol{v} = \begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix}\,; \tag{3.35}$$

usando queste coordinate, i moduli dei vettori si calcolano facilmente: estendendo il Teorema di Pitagora alle tre dimensioni (diagonale del parallelepipedo), abbiamo

$$u^2 := |\boldsymbol{u}|^2 = u_x^2 + u_y^2 + u_z^2 \qquad v^2 := |\boldsymbol{v}| = v_x^2 + v_y^2 + v_z^2\,. \tag{3.36}$$

Il modulo del vettore somma è dato da:

$$\begin{aligned}|\boldsymbol{u}+\boldsymbol{v}|^2 &= (u_x+v_x)^2+(u_y+v_y)^2+(u_z+v_z)^2 \\ &= u_x^2+2u_xv_x+v_x^2+u_y^2+2u_yv_y+v_y^2+u_z^2+2u_zv_z+v_z^2 \\ &= u^2+v^2+2(u_xv_x+u_yv_y+u_zv_z)\,.\end{aligned}\tag{3.37}$$

D'altra parte, usando il teorema di Carnot, possiamo scrivere

$$\begin{aligned}|\boldsymbol{u}+\boldsymbol{v}|^2 &= \overline{OA}^2+\overline{AC}^2-2\overline{OA}\,\overline{AC}\cos\varphi \\ &= u^2+v^2-2|\boldsymbol{u}|\,|\boldsymbol{v}|\cos(\pi-\vartheta) \\ &= u^2+v^2+2|\boldsymbol{u}|\,|\boldsymbol{v}|\cos(\vartheta)=u^2+v^2+2\langle\boldsymbol{u},\boldsymbol{v}\rangle\,,\end{aligned}\tag{3.38}$$

dove abbiamo usato la definizione di prodotto scalare (1.12). Confrontando le scritture (3.37) e (3.38), riotteniamo l'espressione del prodotto scalare in termini delle componenti dei vettori.

Osservazione 3.36. Dalla considerazione geometrica appena fatta discendono quindi tutte le proprietà di bilinearità del prodotto scalare: esse sono ottenibili dall'espressione mediante le coordinate, sfruttando le proprietà delle operazioni di addizione e moltiplicazione nel campo reale.

Inoltre, dall'analisi della figura 3.5 otteniamo immediatamente la seguente

Osservazione 3.37. L'area del parallelogramma descritto dai vettori $\boldsymbol{u},\boldsymbol{v}$ è data da

$$|\boldsymbol{u}|\,|\boldsymbol{v}|\sin\vartheta\,.\tag{3.39}$$

Basta infatti ricordare che l'area si ottiene moltiplicando la base per l'altezza, e sfruttare i teoremi dei triangoli rettangoli.

Siamo pronti ora per dare la nostra nuova

Definizione 3.24. Siano $\boldsymbol{u},\boldsymbol{v}$ vettori di $\mathbb{E}^3_O$. Si chiama *prodotto vettoriale* fra i vettori $\boldsymbol{u}$ e $\boldsymbol{v}$ (nell'ordine) il vettore indicato con

$$\boldsymbol{w}=\boldsymbol{u}\wedge\boldsymbol{v}\,.\tag{3.40}$$

il vettore così ottenuto: se $\boldsymbol{u}$ o $\boldsymbol{v}$ sono il vettore nullo, oppure se $\boldsymbol{u}\in\mathrm{Span}(\boldsymbol{v})$, $\boldsymbol{w}=\boldsymbol{0}$; altrimenti $\boldsymbol{w}$ è il vettore ortogonale al piano $\alpha=\mathrm{Span}(\boldsymbol{u},\boldsymbol{v})$, con verso dato dalla regola della mano destra e di modulo

$$|\boldsymbol{w}|=|\boldsymbol{u}|\,|\boldsymbol{v}|\sin\vartheta\tag{3.41}$$

Osserviamo che:

- Il modulo di $\boldsymbol{u}\wedge\boldsymbol{v}$ è pari all'area del parallelogramma descritto dai due vettori.
- La regola della mano destra equivale a dire che: i tre vettori $\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}$ presi nell'ordine sono orientati reciprocamente come i tre vettori della terna della base canonica $\mathcal{B}=\{\hat{\boldsymbol{\imath}},\hat{\boldsymbol{\jmath}},\hat{\boldsymbol{k}}\}$, nell'ordine canonico. In altre parole: *un'osservatore posto come il terzo vettore $\boldsymbol{w}$ (con la testa dove "punta" la freccia) vedrebbe il primo vettore $\boldsymbol{u}$ sovrapporsi al secondo vettore $\boldsymbol{v}$ ruotando nel piano secondo l'angolo convesso ϑ in verso antiorario.*

La notazione $\boldsymbol{u}\wedge\boldsymbol{v}$ è quella più comune nella letteratura scientifica in italiano; nella letteratura anglosassone, invece, si incontra quasi universalmente la scrittura $\boldsymbol{u}\times\boldsymbol{v}$.

Anche il prodotto vettoriale può essere scritto sfruttando le coordinate dei vettori. Usando una notazione compatta, abbiamo la seguente

Proposizione 3.31.

$$\boldsymbol{u}\wedge\boldsymbol{v}=\begin{vmatrix}\hat{\boldsymbol{\imath}} & \hat{\boldsymbol{\jmath}} & \hat{\boldsymbol{k}}\\ u_x & u_y & u_z\\ v_x & v_y & v_z\,,\end{vmatrix}\tag{3.42}$$

con l'intesa che il determinante nell'equazione (3.42) vada interpretato in modo formale-mnemonico, e si intenda sviluppato unicamente secondo la prima riga.

Dimostrazione. Scriviamo cosa significa il calcolo indicato con la (3.42):

$$\boldsymbol{u}\wedge\boldsymbol{v}=(u_yv_z-u_zv_y)\hat{\boldsymbol{\imath}}+(u_zv_x-u_xv_z)\hat{\boldsymbol{\jmath}}+(u_xv_y-u_yv_x)\hat{\boldsymbol{k}}\,,\tag{3.43}$$

e calcoliamo esplicitamente il modulo quadro del secondo membro della (3.43):

$$\begin{aligned}|\boldsymbol{u}\wedge\boldsymbol{v}|^2 &= (u_yv_z-u_zv_y)^2+(u_zv_x-u_xv_z)^2+(u_xv_y-u_yv_x)^2\\ &=u_y^2v_z^2+u_z^2v_y^2+u_z^2v_x^2+u_x^2v_z^2+u_x^2v_y^2+u_y^2v_x^2\\ &\quad -2u_yv_zu_zv_y-2u_zv_xu_xv_z-2u_xv_yu_yv_x\\ &=u_y^2v_z^2+u_z^2v_y^2+u_z^2v_x^2+u_x^2v_z^2+u_x^2v_y^2+u_y^2v_x^2\\ &\quad +u_x^2v_x^2+u_y^2v_y^2+u_z^2v_z^2-(u_xv_x+u_yv_y+u_zv_z)^2\\ &=(u_x^2+u_y^2+u_z^2)(v_x^2+v_y^2+v_z^2)-(u_xv_x+u_yv_y+u_zv_z)^2\\ &=|\boldsymbol{u}|^2|\boldsymbol{v}|^2-(\langle\boldsymbol{u},\boldsymbol{v}\rangle)^2\\ &=|\boldsymbol{u}|^2|\boldsymbol{v}|^2-|\boldsymbol{u}|^2|\boldsymbol{v}|^2\cos^2\vartheta\\ &=|\boldsymbol{u}|^2|\boldsymbol{v}|^2(1-\cos^2\vartheta)=|\boldsymbol{u}|^2|\boldsymbol{v}|^2\sin^2\vartheta\,,\end{aligned}\tag{3.44}$$

dove i termini evidenziati in rosso sono aggiunti, e il loro opposto è inglobato nel quadrato evidenziato in blu, ed inoltre abbiamo sfruttato l'espressione del prodotto scalare mediante le componenti.

Quindi, il modulo del vettore calcolato con la (3.42) coincide con quanto previsto dalla definizione (3.41). Resta da vedere l'ortogonalità con i vettori e l'orientamento reciproco.

Calcoliamo il prodotto scalare di $\boldsymbol{w}$ con $\boldsymbol{u}$:

$$\begin{aligned}\langle\boldsymbol{u},\boldsymbol{w}\rangle &= u_x(u_yv_z-u_zv_y)+u_y(u_zv_x-u_xv_z)+u_z(u_xv_y-u_yv_x)\\ &=u_xu_yv_z-u_yu_xv_z+u_zu_xv_y-u_xu_zv_y+u_yu_zv_x-u_zu_yv_x\\ &=0\,.\end{aligned}$$

Analogamente, si ottiene $\langle\boldsymbol{v},\boldsymbol{w}\rangle=0$. Per il verso, occorre verificare che se si definisce una nuova base con $\hat{\boldsymbol{\imath}}'$ avente direzione e verso di $\boldsymbol{u}$, $\hat{\boldsymbol{\jmath}}'$ avente direzione e verso della proiezione ortonale di $\boldsymbol{v}$ rispetto a $\boldsymbol{u}$, $\boldsymbol{w}$ ha componente positiva rispetto al terzo vettore della nuova base $\hat{\boldsymbol{k}}'$, se questo ha verso dato dalla regola della mano destra: il calcolo, un po' noioso, è lasciato allo studente □

10.1. Prodotto misto.

Sfruttando le scritture riportate sopra, possiamo verificare la seguente

PROPOSIZIONE 3.32. *Il prodotto misto* $\langle(\boldsymbol{u}_1 \wedge \boldsymbol{u}_2), \boldsymbol{u}_3\rangle$ *si ottiene rappresentando i vettori nella base canonica di* $\mathbb{E}^3_O$ *e calcolando il determinante*

$$\langle(\boldsymbol{u}_1 \wedge \boldsymbol{u}_2), \boldsymbol{u}_3\rangle = \begin{vmatrix} (\boldsymbol{u}_1)_x & (\boldsymbol{u}_1)_y & (\boldsymbol{u}_1)_z \\ (\boldsymbol{u}_2)_x & (\boldsymbol{u}_2)_y & (\boldsymbol{u}_2)_z \\ (\boldsymbol{u}_3)_x & (\boldsymbol{u}_3)_y & (\boldsymbol{u}_3)_z \end{vmatrix} \tag{3.45}$$

OSSERVAZIONE 3.38. Si considerino i vettori $\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3$ disposti come in Figura 3.4; poiché il prodotto vettoriale $\boldsymbol{u}_1 \wedge \boldsymbol{u}_2$ ha modulo pari all'area del parallelogramma descritto dai due vettori, ed è un vettore ortogonale al loro piano, quando si moltiplica il terzo vettore scalarmente per questo, si proietta $\boldsymbol{u}_3$ sulla perpendicolare al piano, ossia si determina l'altezza del parallelepipedo descritto dai vettori. Il prodotto misto, quindi a meno del segno corrisponde al volume del parallelepipedo, come affermato in precedenza.

CAPITOLO 4

Applicazioni lineari

1. Definizioni ed esempi.

In generale, possiamo definire una funzione reale che dipende da più variabili: per esempio, prese le coordinate x, y, z che identificano un vettore di $\mathbb{E}^3_O$, possiamo considerare la funzione $d_0 \colon \mathbb{R}^3 \to \mathbb{R}$ che associa ad esse la distanza del punto dall'origine:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto d_0(x, y, z) = \sqrt{x^2 + y^2 + z^2}.$$

Estendendo l'idea introdotta, possiamo pensare ad un applicazione $\boldsymbol{f} \colon \mathbb{R}^n \to \mathbb{R}^m$ tra gli spazi vettoriali $\mathbb{R}^n$ e $\mathbb{R}^m$ come m funzioni reali $f_i(x_1, \dots x_n)$ (con $i = 1, \dots m$) delle n variabili che identificano un vettore del dominio; in altre parole potremo scrivere

$$\big(x_1, \dots x_n\big) \mapsto \boldsymbol{f}(x_1, x_2, \dots x_n) = \begin{pmatrix} f_1(x_1, \dots x_n) \\ f_2(x_1, \dots x_n) \\ \vdots \\ f_m(x_1, \dots x_n) \end{pmatrix}. \tag{4.1}$$

Ad esempio, la funzione $\boldsymbol{f} \colon \mathbb{R}^3 \to \mathbb{R}^2$ definita come

$$\boldsymbol{f} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x^2 - 3y + 5xz \\ 2x + y + 4z^3 \end{pmatrix}$$

In particolare, ci concentreremo sulle applicazioni tra gli spazi vettoriali $\mathbb{R}^n$ e $\mathbb{R}^m$ che saranno esprimibili mediante espressioni molto semplici fra le coordinate del vettore sulle quali agiscono; saranno funzioni che esprimono legami che chiameremo *lineari*. Scopriremo che questo tipo di applicazione "rispetta" la struttura algebrica fissata in ciascun spazio, cioè "conserva" le operazioni di somma di vettori e di moltiplicazione per uno scalare. Verificheremo infine che tali funzioni trasformano sottospazi in sottospazi.
Cerchiamo ora di capire meglio quale significato daremo a questo termine.

Definizione 4.1. Sia $A \in \mathcal{M}_{\mathbb{R}}(m,n)$ una matrice a entrate reali di ordine $m \times n$, scriviamo

$$A = (A^1|A^2|\dots|A^n) \quad A^i \in \mathbb{R}^m, \quad A^i = \begin{pmatrix} a_{1i} \\ a_{2i} \\ \dots \\ a_{mi} \end{pmatrix}, \quad i = 1,\dots,n.$$

Osserviamo che per ogni vettore $X \in \mathbb{R}^n$, possiamo eseguire la moltiplicazione $A \cdot X$ ed ottenere un vettore di $\mathbb{R}^m$:

$$AX = (A^1|A^2|\dots|A^n) \cdot \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix} = x_1 A^1 + \cdots + x_n A^n \in \mathbb{R}^m.$$

La matrice $A \in \mathcal{M}_{\mathbb{R}}(m,n)$ definisce quindi in modo naturale un'applicazione

$$L_A \colon \mathbb{R}^n \to \mathbb{R}^m, \quad X \mapsto AX. \tag{4.2}$$

Esempio 4.1.
Consideriamo, per esempio, la matrice

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

Questa definisce l'applicazione $L_A \colon \mathbb{R}^3 \to \mathbb{R}^2$ data da

$$L_A \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 3x - y + 2z \\ x + y + 4z \end{pmatrix},$$

ossia

$$\begin{cases} x' = 3x - y + 2z \\ y' = x + y + 4z \end{cases} \tag{4.3}$$

Osservazione 4.2. Notiamo alcune conseguenze immediate della definizione 4.1:

- se A è la matrice nulla, L_A è l'applicazione identicamente nulla, cioè $\forall X \in \mathbb{R}^n$ si ha $L_A(X) = \mathbf{0}_{\mathbb{R}^m}$;
- se $m = n$ e $A = I_n$, L_A è l'identità di $\mathbb{R}^n$, cioè $L_A(X) = X$, per ogni $X \in \mathbb{R}^n$;
- se $m = n = 1$, e $A = (a)$, allora $L_A \colon \mathbb{R} \to \mathbb{R}$ è la funzione $L_A(x) = ax$, $a \in \mathbb{R}$;
- se $m = 1$, $A = (A_1) = (a_1|\dots|a_n)$, $L_A \colon \mathbb{R}^n \to \mathbb{R}$ è l'applicazione seguente: $\forall X \in \mathbb{R}^n$, $L_A(X) = a_1x_1 + \cdots + a_nx_n$.

Lasciamo allo studente la verifica dettagliata delle osservazioni elencate.

Proposizione 4.1. *Sia $(A^1|A^2|\dots|A^n)$, l'applicazione $L_A \colon \mathbb{R}^n \to \mathbb{R}^m$ definita dalla matrice A ha le seguenti proprietà:*

(1) $L_A(\mathbf{0}_{\mathbb{R}^n}) = \mathbf{0}_{\mathbb{R}^m}$*;*
(2) *siano $\{\boldsymbol{e}_1,\dots,\boldsymbol{e}_n\}$ i vettori della base canonica di $\mathbb{R}^n$, si ha:*

$$L_A(\boldsymbol{e}_i) = A^i \; \forall\, i = 1,\dots,n\,;$$

(3) *date due matrici* $A, B \in \mathcal{M}_{\mathbb{R}}(m,n)$*:*

$$L_A = L_B \ \text{ se e solo se } \ A = B\,;$$

(4) *dato un vettore* $Y \in \mathbb{R}^m$*, esiste* $X \in \mathbb{R}^n$ *tale che* $L_A(X) = Y$ *se e solo se* $Y \in \mathrm{Span}(A^1, \dots, A^n)$.

Le proprietà 4.1 (1) e (2) si ottengono applicando la definizione:

$$L_A(\mathbf{0}_{\mathbb{R}^n}) = 0A^1 + 0A^2 + \cdots + 0A^n = \mathbf{0}_{\mathbb{R}^m} + \dots \mathbf{0}_{\mathbb{R}^m} = \mathbf{0}_{\mathbb{R}^m};$$

per ogni $i = 1, 2, \dots, n$ si ha

$$L_A(\boldsymbol{e}_i) = 0A^1 + 0A^2 + \cdots + 1A^i + \cdots + 0A^n = A^i.$$

Verifichiamo la proprietà (3). Ovviamente se $A = B$ allora le applicazioni L_A e L_B coincidono. Supponiamo ora che $L_A = L_B$, ciò significa che

$$L_A(X) = L_B(X) \quad \forall X \in \mathbb{R}^n.$$

In particolare si ha: $L_A(\boldsymbol{e}_i) = L_B(\boldsymbol{e}_i)$, $\forall i = 1, \dots, n$, dalla proprietà (2) segue che le colonne di A coincidono con le colonne di B e quindi $A = B$.
Infine, verifichiamo la proprietà (4). Sia $Y \in \mathbb{R}^m$, esiste $X \in \mathbb{R}^n$ tale che $L_A(X) = Y$ se e solo se esistono n numeri reali $x_1, x_2 \dots, x_n$ tali che

$$x_1A^1 + x_2A^2 + \cdots + x_nA^n = AX = Y,$$

cioè Y è combinazione lineare dei vettori $A^1, \dots, A^n$.

Le applicazioni L_A sono caratterizzate dalle seguenti proprietà:

Proprietà 4.2. *Le applicazioni definite dalle matrici di* $\mathcal{M}_{\mathbb{R}}(m,n)$ *sono tutte e sole le applicazioni* $L\colon \mathbb{R}^n \to \mathbb{R}^m$ *che soddisfano le seguenti* ***proprietà di linearità****:*

$\mathcal{L}_1$) $L(X_1 + X_2) = L(X_1) + L(X_2), \quad \forall X_1, X_2 \in \mathbb{R}^n$*;*
$\mathcal{L}_2$) $L(\alpha X) = \alpha L(X)$*,* $\forall X \in \mathbb{R}^n$*,* $\forall \alpha \in \mathbb{R}$*.*

Dimostrazione. Osserviamo che ogni applicazione L_A soddisfa tali proprietà, che risultano essere conseguenze delle analoghe proprietà di linearità (si vedano le Proprietà 3.4) del prodotto di matrici, infatti:

$$L_A(X_1 + X_2) = A \cdot (X_1 + X_2) = A \cdot X_1 + A \cdot X_2 = L_A(X_1) + L_A(X_2),$$

$$L_A(\alpha X) = A \cdot (\alpha X) = \alpha(A \cdot X) = \alpha L_A(X),$$

$\forall X, X_1, X_2 \in \mathbb{R}^n$, $\forall \alpha \in \mathbb{R}$.
Sia ora $L\colon \mathbb{R}^n \to \mathbb{R}^m$ un'applicazione che soddisfa le proprietà di linearità, e sia $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$ la base canonica di $\mathbb{R}^n$: consideriamo i vettori

$$B^1 = L(\boldsymbol{e}_1), B^2 = L(\boldsymbol{e}_2), \dots \ (B^i = L(\boldsymbol{e}_i), \quad i = 1, \dots, n),$$

e sia $B \in \mathcal{M}_{\mathbb{R}}(m,n)$ la matrice formata da queste colonne: $B = (B^1|\dots|B^n)$.
Sia $X \in \mathbb{R}^n$, applicando le proprietà di linearità di L abbiamo:

$$\begin{aligned} L(X) &= L(x_1\boldsymbol{e}_1 + \cdots + x_n\boldsymbol{e}_n) \\ &= x_1L(\boldsymbol{e}_1) + \cdots + x_nL(\boldsymbol{e}_n) = x_1B^1 + \cdots + x_nB^n \\ &= B \cdot X\,, \end{aligned}$$

$\forall X \in \mathbb{R}^n$. Possiamo concludere che L coincide con l'applicazione L_B definita dalla matrice B. □

Le proprietà di linearità ci garantiscono che l'applicazione L "conserva" le operazioni di somma di vettori e di moltiplicazione per un numero reale fissate negli spazi $\mathbb{R}^n$ e $\mathbb{R}^m$: infatti la prima proprietà ci dice che l'immagine della somma di due vettori di $\mathbb{R}^n$ è la somma in $\mathbb{R}^m$ delle loro immagini, la seconda che l'immagine del prodotto di un vettore di $\mathbb{R}^n$ per uno scalare è il prodotto in $\mathbb{R}^m$ dell'immagine per lo stesso scalare. Ciò significa che fare le operazioni in $\mathbb{R}^n$ e poi trasformare con l'applicazione L è equivalente a trasformare prima con L e poi fare le operazioni. Ciò giustifica la seguente definizione:

DEFINIZIONE 4.2. Un'applicazione $L \colon V \to W$ tra due spazi vettoriali reali è detta **applicazione lineare** o **operatore lineare** di V in W se verifica le seguenti condizioni di linearità:

(1) $L(\boldsymbol{v}_1 + \boldsymbol{v}_2) = L(\boldsymbol{v}_1) + L(\boldsymbol{v}_2), \quad \forall \boldsymbol{v}_1, \boldsymbol{v}_2 \in V$;
(2) $L(\alpha \boldsymbol{v}) = \alpha L(\boldsymbol{v}), \quad \forall \boldsymbol{v} \in V, \forall \alpha \in \mathbb{R}$.

Un'applicazione lineare $L \colon V \to V$ è detto **operatore lineare** (o **endomorfismo**) di V.

OSSERVAZIONE 4.3.

- Le applicazioni definite dalle matrici di $\mathcal{M}_{\mathbb{R}}(m, n)$ sono **tutte e sole** le applicazioni lineari di $\mathbb{R}^n$ in $\mathbb{R}^m$.
- Ogni applicazione lineare $L \colon V \to W$ trasforma il vettore nullo di V nel vettore nullo di W, cioè si ha:

$$L(\mathbf{0}_V) = \mathbf{0}_W. \tag{4.4}$$

Infatti, preso un qualunque vettore $\boldsymbol{v} \in V$, dall'uguglianza $\mathbf{0}_V = 0.\boldsymbol{v}$ (si veda la Proposizione 2.1 (2)), applicando la seconda proprietà di linearità abbiamo

$$L(\mathbf{0}_V) = L(0.\boldsymbol{v}) = 0L(\boldsymbol{v}) = \mathbf{0}_W.$$

- In meccanica, gli operatori lineari di $\mathbb{E}^3_O$ sono detti *tensori del secondo ordine (o di rango 2).*

ESEMPIO 4.4. Consideriamo i seguenti esempi:

(1) Sia V uno spazio vettoriale reale, consideriamo l'applicazione identità $I_V \colon V \to V$ data da

$$I_V(\boldsymbol{v}) = \boldsymbol{v}, \quad \forall \boldsymbol{v} \in V.$$

Verifichiamo che è un'applicazione lineare. Si ha:

$$I_V(\boldsymbol{v}_1 + \boldsymbol{v}_2) = \boldsymbol{v}_1 + \boldsymbol{v}_2 = I_V(\boldsymbol{v}_1) + I_V(\boldsymbol{v}_2),\ \forall \boldsymbol{v}_1, \boldsymbol{v}_2 \in V;$$

$$I_V(\alpha \boldsymbol{v}) = \alpha \boldsymbol{v} = \alpha I_V(\boldsymbol{v}),\ \forall \boldsymbol{v} \in V,\ \forall \alpha \in \mathbb{R}.$$

(2) Siano V e W due spazi vettoriali reali, consideriamo un'applicazione costante $L\colon V \to W$ data da
$$L(\boldsymbol{v}) = \boldsymbol{w}_0 \quad \forall \boldsymbol{v} \in V, \quad \boldsymbol{w}_0 \in W.$$
Notiamo che affinché L sia lineare si deve avere (c.n.)
$$L(\mathbf{0}_V) = \mathbf{0}_W,$$
possiamo quindi concludere che se $\boldsymbol{w}_0 \neq \mathbf{0}_W$, L non è lineare. Supponiamo ora che $\boldsymbol{w}_0 = \mathbf{0}_W$, allora L è l'applicazione nulla:
$$L(\boldsymbol{v}) = \mathbf{0}_W, \quad \forall \boldsymbol{v} \in V.$$
Verifichiamo infine che l'applicazione nulla è lineare:
$$L(\boldsymbol{v}_1 + \boldsymbol{v}_2) = \mathbf{0}_W + \mathbf{0}_W = L(\boldsymbol{v}_1) + L(\boldsymbol{v}_2),\ \forall \boldsymbol{v}_1, \boldsymbol{v}_2 \in V;$$
$$L(\alpha \boldsymbol{v}) = \mathbf{0}_W = \alpha.\mathbf{0}_W = \alpha L(\boldsymbol{v}),\ \forall \boldsymbol{v} \in V,\ \forall \alpha \in \mathbb{R}.$$

(3) Sia $L\colon \mathbb{R}^2 \to \mathbb{R}^2$ l'applicazione data da
$$L\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x^2 \\ x - y \end{pmatrix}.$$
Vogliamo stabilire se L è un'applicazione lineare. Osserviamo che risulta soddisfatta la condizione necessaria per la linearità:
$$L\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix};$$
ma ciò non è sufficiente. Proviamo, prendendo alcuni vettori di $\mathbb{R}^2$, se L verifica la prima condizione di linearità. Scegliamo i vettori:
$$\boldsymbol{v}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \boldsymbol{v}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$
abbiamo
$$L(\boldsymbol{v}_1 + \boldsymbol{v}_2) = L\begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 1 \end{pmatrix}.$$
Risulta tuttavia:
$$L(\boldsymbol{v}_1) + L(\boldsymbol{v}_2) = \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix},$$
essendo
$$L(\boldsymbol{v}_1 + \boldsymbol{v}_2) \neq L(\boldsymbol{v}_1) + L(\boldsymbol{v}_2),$$
possiamo concludere che L non è lineare.

(4) Consideriamo lo spazio vettoriale $\mathcal{M}_{\mathbb{R}}(2)$ delle matrici reali quadrate di ordine 2. Sia $L\colon \mathcal{M}_{\mathbb{R}}(2) \to \mathbb{R}$ l'applicazione data da
$$L\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11} + a_{22}.$$
(dove $a_{11}+a_{22}$ è la *traccia* della matrice, Definizione 4.6). Verifichiamo che L soddisfa la prima condizione di linearità. Siano $A, B \in \mathcal{M}_{\mathbb{R}}(2)$ due matrici reali quadrate di ordine 2:
$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \quad B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix},$$

si ha

$$L(A) = a_{11} + a_{22} \quad L(B) = b_{11} + b_{22}.$$

La somma delle matrici A e B è la seguente matrice:

$$A + B = \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} \\ a_{21} + b_{21} & a_{22} + b_{22} \end{pmatrix},$$

abbiamo, quindi:

$$\begin{aligned} L(A + B) &= (a_{11} + b_{11}) + (a_{22} + b_{22}) \\ &= (a_{11} + a_{22}) + (b_{11} + b_{22}) \\ &= L(A) + L(B). \end{aligned} \tag{4.5}$$

Verifichiamo ora la seconda proprietà di linearità. Siano $\lambda \in \mathbb{R}$ e $A \in \mathcal{M}_{\mathbb{R}}(2)$:

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

il prodotto della matrice A per il numero reale λ è la seguente matrice:

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

Abbiamo quindi:

$$L(\lambda A) = \lambda a_{11} + \lambda a_{22} = \lambda(a_{11} + a_{22}) = \lambda L(A).$$

(5) Consideriamo lo spazio vettoriale $\mathcal{M}_{\mathbb{R}}(n)$ delle matrici reali quadrate di ordine $n \geq 2$. Sia $d\colon : \mathcal{M}_{\mathbb{R}}(n) \to \mathbb{R}$ l'applicazione che associa ad ogni matrice il determinante: $d(A) = |A|$. Verifichiamo che d non è un'applicazione lineare. Osserviamo che dalle proprietà del determinante abbiamo:

$$|\alpha A| = (\alpha)^n |A| \quad \forall \alpha \in \mathbb{R},$$

per cui se $\alpha \neq 0$ e $\alpha \neq \pm 1$ risulta:

$$d(\alpha A) \neq \alpha\, d(A),$$

possiamo quindi concludere che d non è un'applicazione lineare.

(6) Consideriamo lo spazio vettoriale $\mathbb{R}[x]$ dei polinomi $p(x)$ in x di grado qualunque a coefficienti reali. Sia $D\colon \mathbb{R}[x] \to \mathbb{R}[x]$ l'applicazione che associa al polinomio $p(x)$ il polinomio $D(p(x))$ ottenuto derivando $p(x)$. Posto

$$p(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n,$$

si ha

$$D(p(x)) = a_1 + 2a_2 x + \cdots + n a_n x^{n-1}.$$

D è un operatore lineare, infatti dalle proprietà della derivazione delle funzioni reali si ha:

$$D(p_1(x) + p_2(x)) = D(p_1(x)) + D(p_2(x)), \quad \forall p_1(x), p_2(x) \in \mathbb{R}_n[x],$$

e

$$D(\alpha p(x)) = \alpha D(p(x)), \forall \alpha \in \mathbb{R}, \forall p(x) \in \mathbb{R}_n[x].$$

Concludiamo questa parte verificando la proprietà che avevamo annunciato: *ogni applicazione lineare trasforma sottospazi in sottospazi.*

PROPRIETÀ 4.3. *Siano $L\colon V \to W$ un'applicazione lineare tra spazi vettoriali reali e $U \subseteq V$ un sottospazio di V. L'immagine di U secondo L (Definizione 0.9):*

$$L(U) = \{L(\boldsymbol{u}),\ \boldsymbol{u} \in U\} = \{\boldsymbol{w} \in W |\ \ \exists \boldsymbol{u} \in U\colon L(\boldsymbol{u}) = \boldsymbol{w}\}$$

è un sottospazio di W.

Per verificare che l'insieme $L(U) \subseteq W$ è un sottospazio di W basta provare che:

(1) $\forall \boldsymbol{w}_1, \boldsymbol{w}_2 \in L(U)$, $\boldsymbol{w}_1 + \boldsymbol{w}_2 \in L(U)$;
(2) $\forall \boldsymbol{w} \in L(U)$, $\forall \alpha \in \mathbb{R}$, $\alpha \boldsymbol{w} \in L(U)$.

DIMOSTRAZIONE. Siano $\boldsymbol{w}_1, \boldsymbol{w}_2 \in L(U)$: esistono allora due vettori $\boldsymbol{u}_1, \boldsymbol{u}_2 \in U$ tali che

$$L(\boldsymbol{u}_1) = \boldsymbol{w}_1 \quad L(\boldsymbol{u}_2) = \boldsymbol{w}_2.$$

Per la linearità di L abbiamo:

$$L(\boldsymbol{u}_1 + \boldsymbol{u}_2) = L(\boldsymbol{u}_1) + L(\boldsymbol{u}_2) = \boldsymbol{w}_1 + \boldsymbol{w}_2,$$

poiché U è un sottospazio di V si ha $\boldsymbol{u}_1 + \boldsymbol{u}_2 = \boldsymbol{u} \in U$ e l'uguaglianza

$$\boldsymbol{w}_1 + \boldsymbol{w}_2 = L(\boldsymbol{u}), \quad \boldsymbol{u} \in U,$$

da cui otteniamo $\boldsymbol{w}_1 + \boldsymbol{w}_2 \in L(U)$.

Siano, ora, $\boldsymbol{w} \in L(U)$ e $\alpha \in \mathbb{R}$: esiste un vettore $\boldsymbol{u} \in U$ tale che $L(\boldsymbol{u}) = \boldsymbol{w}$. Per la linearità di L abbiamo

$$L(\alpha \boldsymbol{u}) = \alpha L(\boldsymbol{u}) = \alpha \boldsymbol{w},$$

poiché U è un sottospazio di V si ha $\alpha \boldsymbol{u} = \boldsymbol{u}' \in U$ e $L(\boldsymbol{u}') = \alpha \boldsymbol{w}$, per cui possiamo concludere che $\alpha \boldsymbol{w} \in L(U)$. □

2. Nucleo ed Immagine di un'applicazione lineare.

Associamo ad ogni applicazione lineare tra spazi vettoriali reali due sottoinsiemi che sono utili nello studio dell'applicazione.

DEFINIZIONE 4.3. Sia $L\colon V \to W$ un'applicazione lineare tra spazi vettoriali reali, associamo a L i seguenti insiemi:

- **Nucleo** di L:
$$\operatorname{Ker} L = \{\boldsymbol{v} \in V \mid L(\boldsymbol{v}) = \boldsymbol{0}_W\};$$
- **Immagine** di L:
$$\operatorname{Im} L = \{L(\boldsymbol{v}),\ \boldsymbol{v} \in V\} = \{\boldsymbol{w} \in W \mid \exists\ \boldsymbol{v} \in V\colon\ L(\boldsymbol{v}) = \boldsymbol{w}\}.$$

Osservazione 4.5. Osserviamo che:

- $\operatorname{Ker} L \subseteq V$ e $\operatorname{Ker} L \neq \varnothing$: infatti poiché $L(\mathbf{0}_V) = \mathbf{0}_W$ abbiamo
$$\mathbf{0}_V \in \operatorname{Ker} L,$$
quindi contiene almeno il vettore nullo di V;
- analogamente, $\operatorname{Im} L \subseteq W$ e $\operatorname{Im} L \neq \varnothing$: infatti poiché $\mathbf{0}_W = L(\mathbf{0}_V)$, abbiamo
$$\mathbf{0}_W \in \operatorname{Im} L,$$
quindi $\operatorname{Im} L$ contiene almeno il vettore nullo di W.

Proprietà 4.4. *Sia $L\colon V \to W$ un'applicazione lineare tra spazi vettoriali reali:*

(1) $\operatorname{Ker} L$ *è un sottospazio di* V*;*
(2) $\operatorname{Im} L$ *è un sottospazio di* W*.*

Dimostrazione. Ricordiamo che $\operatorname{Ker} L$ è un sottospazio di V se è chiuso rispetto alla somma di vettori e chiuso rispetto alla moltiplicazione per uno scalare reale. Verifichiamo che
$$\forall \boldsymbol{v}_1, \boldsymbol{v}_2 \in \operatorname{Ker} L \ \text{ si ha } \ \boldsymbol{v}_1 + \boldsymbol{v}_2 \in \operatorname{Ker} L,$$
$$\forall \boldsymbol{v} \in \operatorname{Ker} L, \quad \forall \alpha \in \mathbb{R} \ \text{ si ha } \ \alpha \boldsymbol{v} \in \operatorname{Ker} L.$$
Per la prima proprietà di linearità di L abbiamo:
$$L(\boldsymbol{v}_1 + \boldsymbol{v}_2) = L(\boldsymbol{v}_1) + L(\boldsymbol{v}_2),$$
poiché $\boldsymbol{v}_i \in \operatorname{Ker} L$ risulta $L(\boldsymbol{v}_i) = \mathbf{0}_W$, $i = 1, 2$, quindi
$$L(\boldsymbol{v}_1 + \boldsymbol{v}_2) = L(\boldsymbol{v}_1) + L(\boldsymbol{v}_2) = \mathbf{0}_W + \mathbf{0}_W = \mathbf{0}_W,$$
cioè $L(\boldsymbol{v}_1 + \boldsymbol{v}_2) = \mathbf{0}_W$. Possiamo concludere che $\boldsymbol{v}_1 + \boldsymbol{v}_2 \in \operatorname{Ker} L$.

Analogamente per la seconda proprietà di linearità di L per ogni vettore $\boldsymbol{v} \in \operatorname{Ker} L$ si ha:
$$L(\alpha \boldsymbol{v}) = \alpha L(\boldsymbol{v}) = \alpha\, \mathbf{0}_W = \mathbf{0}_W,$$
da cui ricaviamo che $\alpha \boldsymbol{v} \in \operatorname{Ker} L$.

Osserviamo che per definizione risulta
$$\operatorname{Im} L = \{\boldsymbol{w} \in W \mid \exists \boldsymbol{v} \in V \colon L(\boldsymbol{v}) = \boldsymbol{w}\} = L(V),$$
poiché abbiamo verificato che ogni applicazione lineare trasforma sottospazi in sottospazi, anche $L(V)$ risulta essere un sottospazio di W (infatti, V è un sottospazio banale di V stesso). □

I sottospazi nucleo e immagine di una applicazione lineare ne caratterizzano l'iniettività e la suriettività. Ricordiamo le nozioni di applicazione iniettiva e suriettiva (Definizioni 0.12).

Sia $L\colon V \to W$ un'applicazione tra due insiemi.

(1) L è **iniettiva** se $\forall \boldsymbol{v}_1, \boldsymbol{v}_2 \in V$:
$$\boldsymbol{v}_1 \neq \boldsymbol{v}_2 \implies L(\boldsymbol{v}_1) \neq L(\boldsymbol{v}_2);$$

(2) L è **suriettiva** se ogni elemento di W è immagine di almeno un elemento di V, i.e.

$$\forall\, \boldsymbol{w} \in W\ \exists\ \boldsymbol{v} \in V\ :\ L(\boldsymbol{v}) = \boldsymbol{w} \iff \operatorname{Im} L = W.$$

Proprietà 4.5. *Sia $L\colon V \to W$ un'applicazione lineare tra spazi vettoriali reali.*

(1) *L è iniettiva* $\iff \operatorname{Ker} L = \{\mathbf{0}_V\}$*;*
(2) *L è suriettiva* $\iff \operatorname{Im} L = W$*.*

Dimostrazione. Verifichiamo la prima proprietà. Mostriamo che se L è iniettiva il nucleo di L contiene solo il vettore nullo. Sia, per assurdo, $\boldsymbol{v} \in \operatorname{Ker} L$ un vettore del nucleo non nullo: $\boldsymbol{v} \neq \mathbf{0}_V$ e $L(\boldsymbol{v}) = \mathbf{0}_W$. Ricordiamo che poiché L è lineare si ha anche: $\mathbf{0}_W = L(\mathbf{0}_V)$, otteniamo quindi l'uguaglianza

$$L(\boldsymbol{v}) = L(\mathbf{0}_V),$$

poiché $\boldsymbol{v} \neq \mathbf{0}_V$ ciò contraddice l'iniettività di L.
Viceversa, supponiamo ora che il nucleo di L contenga solo il vettore nullo. Per provare che L è iniettiva proviamo che, $\forall \boldsymbol{v}_1, \boldsymbol{v}_2 \in V$:

$$L(\boldsymbol{v}_1) = L(\boldsymbol{v}_2) \implies \boldsymbol{v}_1 = \boldsymbol{v}_2.$$

Sia $L(\boldsymbol{v}_1) = L(\boldsymbol{v}_2)$, per la linearità di L abbiamo:

$$L(\boldsymbol{v}_1 - \boldsymbol{v}_2) = L(\boldsymbol{v}_1) - L(\boldsymbol{v}_2) = \mathbf{0}_W,$$

e quindi $\boldsymbol{v}_1 - \boldsymbol{v}_2 \in \operatorname{Ker} L$, poiché il nucleo contiene solo il vettore nullo, risulta necessariamente $\boldsymbol{v}_1 - \boldsymbol{v}_2 = \mathbf{0}_V$ e quindi $\boldsymbol{v}_1 = \boldsymbol{v}_2$.
La seconda proprietà segue immediatamente dalla definizione di suriettività. □

Esempio 4.6 (**Fondamentale**). Sia $L_A\colon \mathbb{R}^n \to \mathbb{R}^m$ l'applicazione lineare definita dalla matrice reale $A \in \mathcal{M}_{\mathbb{R}}(m,n)$:

$$L_A(X) = AX, \quad \forall X \in \mathbb{R}^n.$$

Vogliamo determinare i sottospazi nucleo ed immagine di L_A. Dalla definizione di nucleo e immagine abbiamo:

$$\operatorname{Ker} L_A = \{X \in \mathbb{R}^n \mid AX = \mathbf{0}_{\mathbb{R}^m}\ \},$$

$$\operatorname{Im} L_A = \{AX,\ X \in \mathbb{R}^n\} = \{Y \in \mathbb{R}^m \mid \exists X \in \mathbb{R}^n\ :\ AX = Y\ \}.$$

Siano $A^1, A^2, \dots, A^n$ i vettori colonna della matrice A, ricordiamo che moltiplicando la matrice A per il vettore $X \in \mathbb{R}^n$ si ottiene la combinazione lineare delle colonne di A con vettore dei coefficienti X:

$$A.X = x_1A^1 + x_2A^2 + \cdots + x_nA^n,$$

per cui possiamo concludere che:

- $\operatorname{Im} L_A = \operatorname{Span}(A^1, \dots, A^n)$;
- $\dim(\operatorname{Im} L_A) = \dim(\operatorname{Span}(A^1, \dots, A^n)) = \operatorname{rg}(A)$;
- L_A è suriettiva $\iff \mathbb{R}^m = \operatorname{Span}(A^1, \dots, A^n) \iff \operatorname{rg}(A) = m$;
- esiste un vettore non nullo $X \in \operatorname{Ker} L_A \iff$ i vettori colonna $A^1, A^2, \dots, A^n$ sono linearmente dipendenti;

- L_A è iniettiva $\iff$ i vettori colonna $A^1, \dots, A^n$ sono linearmente indipendenti $\iff$ $\mathrm{rg}(A) = n$.

Ci limitiamo ora a considerare spazi vettoriali finitamente generati, abbiamo allora la seguente descrizione dello spazio immagine che generalizza l'esempio precedente:

PROPRIETÀ 4.6. *Sia $L\colon V \to W$ un'applicazione lineare tra spazi vettoriali reali di dimensione finita. Sia $\mathcal{B}_V = \{\boldsymbol{v}_1, \dots, \boldsymbol{v}_n\}$ una base di V: i vettori $L(\boldsymbol{v}_1), L(\boldsymbol{v}_2), \dots, L(\boldsymbol{v}_n)$ sono un* sistema di generatori *dello spazio* $\operatorname{Im} L$,*ossia*

$$\operatorname{Im} L = \operatorname{Span}(L(\boldsymbol{v}_1), L(\boldsymbol{v}_2), \dots, L(\boldsymbol{v}_n)),$$

quindi $\dim(\operatorname{Im} L) \le \dim V$.

DIMOSTRAZIONE. Verifichiamo la proprietà enunciata. Sia $\boldsymbol{w} \in \operatorname{Im} L$, esiste allora un vettore $\boldsymbol{v} \in V$ tale che $L(\boldsymbol{v}) = \boldsymbol{w}$. Poiché $V = \operatorname{Span}(\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n)$, abbiamo:

$$\boldsymbol{v} = x_1\boldsymbol{v}_1 + x_2\boldsymbol{v}_2 + \cdots + x_n\boldsymbol{v}_n,$$

con $x_i \in \mathbb{R}$. Sostituendo al posto di $\boldsymbol{v}$ tale scrittura e applicando la linearità di L otteniamo:

$$\boldsymbol{w} = L(\boldsymbol{v}) = L(x_1\boldsymbol{v}_1 + x_2\boldsymbol{v}_2 + \cdots + x_n\boldsymbol{v}_n) = x_1L(\boldsymbol{v}_1) + x_2L(\boldsymbol{v}_2) + \cdots + x_nL(\boldsymbol{v}_n),$$

cioè $\boldsymbol{w}$ è combinazione lineare dei vettori $L(\boldsymbol{v}_1), L(\boldsymbol{v}_2), \dots, L(\boldsymbol{v}_n)$. Possiamo quindi concludere che $L(\boldsymbol{v}_1), L(\boldsymbol{v}_2), \dots, L(\boldsymbol{v}_n)$ sono generatori di $\operatorname{Im} L$.

NOTA BENE: Non è detto che $\{L(\boldsymbol{v}_1), L(\boldsymbol{v}_2), \dots, L(\boldsymbol{v}_n)\}$ siano linearmente indipendenti, e quindi una base di $\operatorname{Im} L$!
Poiché $\operatorname{Im} L \subseteq W$, ovviamente si ha $\dim(\operatorname{Im} L) \le \dim W$. Supponiamo ora che

$$n = \dim V > \dim W,$$

allora i vettori $\{L(\boldsymbol{v}_1), L(\boldsymbol{v}_2), \dots, L(\boldsymbol{v}_n)\}$ sono necessariamente linearmente dipendenti in W, quindi non sono una base di $\operatorname{Im} L$. $\square$

Se l'applicazione lineare è **iniettiva**, i vettori $\{L(\boldsymbol{v}_1), L(\boldsymbol{v}_2), \dots, L(\boldsymbol{v}_n)\}$ sono linearmente indipendenti e quindi costituiscono una base di $\operatorname{Im} L$:

PROPOSIZIONE 4.7. *Sia $L\colon V \to W$ un'applicazione lineare* **iniettiva***: l'immagine di una base di V è una base di* $\operatorname{Im} L$.

DIMOSTRAZIONE. Sia $\mathcal{B}_V = \{\boldsymbol{v}_1, \dots, \boldsymbol{v}_n\}$ una base di V. Poiché i vettori $\{L(\boldsymbol{v}_1), \dots, L(\boldsymbol{v}_n)\}$ sono un sistema di generatori di $\operatorname{Im} L$ è sufficiente provare che sono linearmente indipendenti. Supponiamo, per assurdo, che tali vettori siano linearmente dipendenti: esisterebbero, allora, n scalari non tutti nulli tali che

$$\lambda_1L(\boldsymbol{v}_1) + \cdots + \lambda_nL(\boldsymbol{v}_n) = \mathbf{0}_W.$$

Osserviamo che per la linearità risulta

$$L(\lambda_1\boldsymbol{v}_1 + \cdots + \lambda_n\boldsymbol{v}_n) = \lambda_1L(\boldsymbol{v}_1) + \cdots + \lambda_nL(\boldsymbol{v}_n) = \mathbf{0}_W,$$

cioè il vettore $\lambda_1 \boldsymbol{v}_1 + \cdots + \lambda_n \boldsymbol{v}_n \in \operatorname{Ker} L$. Poiché per ipotesi L è iniettiva, si ha: $\operatorname{Ker} L = \{\mathbf{0}_V\}$. Possiamo quindi concludere che

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

cioè i vettori $\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_n\}$ sarebbero linearmente dipendenti in V, che è assurdo poiché per ipotesi sono una base di V. $\square$

Esempio 4.7. Sia $L\colon \mathbb{R}^3 \to \mathbb{R}^3$ l'operatore lineare dato da:

$$L \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

Fissata la base canonica $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ di $\mathbb{R}^3$, determiniamo le immagini dei vettori della base:

$$L(\boldsymbol{e}_1) = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = A^1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix},$$

$$L(\boldsymbol{e}_2) = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = A^2 = \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix},$$

$$L(\boldsymbol{e}_3) = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = A^3 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}.$$

I vettori $L(\boldsymbol{e}_1), L(\boldsymbol{e}_2), L(\boldsymbol{e}_3)$ sono generatori dello spazio $\operatorname{Im} L$, tuttavia non sono linearmente indipendenti: infatti $L(\boldsymbol{e}_3) = L(\boldsymbol{e}_1) + L(\boldsymbol{e}_2)$. Osserviamo tuttavia che $L(\boldsymbol{e}_1)$ ed $L(\boldsymbol{e}_2)$ sono linearmente indipendenti, quindi abbiamo:

$$\operatorname{Im} L = \operatorname{Span}(L(\boldsymbol{e}_1), L(\boldsymbol{e}_2))$$

ossia, $\{L(\boldsymbol{e}_1), L(\boldsymbol{e}_2)\}$ è una base di $\operatorname{Im} L$.

Risulta allora $\dim(\operatorname{Im} L) = 2$, per cui $\operatorname{Im} L$ è il piano per l'origine generato dai vettori $L(\boldsymbol{e}_1), L(\boldsymbol{e}_2)$.

Vogliamo infine determinare il nucleo di L:

$$\operatorname{Ker} L = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \mid \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\},$$

i vettori di $\operatorname{Ker} L$ sono tutti e soli i vettori le cui coordinate x, y, z sono le soluzioni del seguente sistema lineare:

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

Osserviamo che l'ultima equazione è superflua, e le prime due sono linearmente indipendenti, quindi $\operatorname{Ker} L$ è l'intersezione di due piani in $\mathbb{R}^3$. Precisamente risolvendo abbiamo:

$$\operatorname{Ker} L = \left\{ \begin{pmatrix} x \\ x \\ -x \end{pmatrix}, x \in \mathbb{R} \right\} = \operatorname{Span}\begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}.$$

Osserviamo che $\dim \operatorname{Im} L + \dim \operatorname{Ker} L = 3 = \dim \mathbb{R}^3$. Questo risultato fondamentale è vero in generale, ed è espresso nel **Teorema delle dimensioni** 4.8, che fornisce una relazione tra le dimensioni degli spazi nucleo ed immagine:

TEOREMA 4.8 (delle dimensioni). *Sia $L\colon V \to W$ un'applicazione lineare tra spazi vettoriali reali di dimensione finita. Vale la seguente relazione:*

$$\dim(V) = \dim(\operatorname{Ker} L) + \dim(\operatorname{Im} L).$$

DIMOSTRAZIONE. Osserviamo che se $\dim(\operatorname{Ker} L) = 0$, allora $\operatorname{Ker} L = \{\mathbf{0}_V\}$ e quindi L è iniettiva. Abbiamo provato, (prop. 4.7), che L trasforma una base di V in una base di $\operatorname{Im} L$. Quindi risulta $\dim V = \dim(\operatorname{Im} L)$.

Supponiamo ora $\dim(\operatorname{Ker} L) = h > 0$, verifichiamo che $\dim \operatorname{Im} L = n - h$. Sia $U \subset V$ un sottospazio complementare di $\operatorname{Ker} L$, tale che

$$U \cap \operatorname{Ker} L = \{\mathbf{0}_V\} \quad U \oplus \operatorname{Ker} L = V.$$

Siano $\{\boldsymbol{u}_1, \dots, \boldsymbol{u}_h\}$ una base di $\operatorname{Ker} L$ e $\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_{n-h}\}$ una base di U, abbiamo visto che l'unione di tali basi è una base $\mathcal{B}_V$ di V. Consideriamo le immagini dei vettori di $\mathcal{B}_V$:

$$\{L(\boldsymbol{u}_1), \dots, L(\boldsymbol{u}_h), L(\boldsymbol{v}_1), \dots, L(\boldsymbol{v}_{n-h})\},$$

sono un sistema di generatori di $\operatorname{Im} L$. Osserviamo però che risulta

$$L(\boldsymbol{u}_1) = \dots = L(\boldsymbol{u}_h) = \mathbf{0}_W,$$

poiché $\boldsymbol{u}_i \in \operatorname{Ker} L$, $i = 1, \dots, h$, quindi possiamo concludere che anche

$$L(\boldsymbol{v}_1), \dots, L(\boldsymbol{v}_{n-h})$$

sono generatori di $\operatorname{Im} L$. Rimane da provare che tali vettori sono linearmente indipendenti in W. Supponiamo che tali vettori siano linearmente dipendenti: esistono $n - h$ scalari non tutti nulli tali che

$$\lambda_1 L(\boldsymbol{v}_1) + \dots + \lambda_{n-h} L(\boldsymbol{v}_{n-h}) = \mathbf{0}_W,$$

osserviamo che per la linearità risulta

$$L(\lambda_1 \boldsymbol{v}_1 + \dots + \lambda_{n-h} \boldsymbol{v}_{n-h}) = \mathbf{0}_W,$$

cioè il vettore $\lambda_1 \boldsymbol{v}_1 + \dots + \lambda_{n-h} \boldsymbol{v}_{n-h} \in \operatorname{Ker} L$. Poiché per ipotesi

$$U = \operatorname{Span}(\boldsymbol{v}_1, \dots, \boldsymbol{v}_{n-h}) \qquad \text{e} \qquad U \cap \operatorname{Ker} L = \{\mathbf{0}_V\},$$

possiamo concludere che

$$\lambda_1 \boldsymbol{v}_1 + \dots + \lambda_{n-h} \boldsymbol{v}_{n-h} = \mathbf{0}_V,$$

cioè i vettori sono linearmente dipendenti in V, che è assurdo poiché per ipotesi i vettori sono una base di U. □

OSSERVAZIONE 4.8. Osserviamo che risulta

$$\dim(\operatorname{Im} L) = \dim V \iff \dim(\operatorname{Ker} L) = 0 \iff \operatorname{Ker} L = \{\mathbf{0}_V\} \iff L \text{ è iniettiva.}$$

Il teorema ha un interessante corollario:

Corollario 4.9. *Sia* $L_A\colon \mathbb{R}^n \to \mathbb{R}^m$ *l'applicazione lineare definita dalla matrice* $A \in \mathcal{M}_{\mathbb{R}}(m,n)$*:*

$$L_A(X) = AX, \qquad \forall X \in \mathbb{R}^n.$$

Abbiamo:

$$\operatorname{Ker} L_A = \{X \in \mathbb{R}^n \mid AX = \mathbf{0}_{\mathbb{R}^m}\ \},$$
$$\operatorname{Im} L_A = \{Y \in \mathbb{R}^m \mid \exists X \in \mathbb{R}^n\ : AX = Y\ \},$$

inoltre abbiamo precedentemente osservato che risulta

$$\operatorname{Im} L_A = \operatorname{Span}(A^1, A^2, \dots, A^n),$$

dove $A^1, A^2, \dots, A^n$ *sono i vettori colonna della matrice* A*. Quindi* $\dim(\operatorname{Im} L_A) = \operatorname{rg}(A)$*, applicando il teorema delle dimensioni otteniamo:*

$$\dim(\operatorname{Ker} L_A) = \dim(\mathbb{R}^n) - \dim(\operatorname{Im} L_A) = n - \operatorname{rg}(A).$$

Esempio 4.9. Vediamo alcuni esempi di applicazione del Teorema delle dimensioni

(1) Nell'esempio precedente 4.7 abbiamo:

$$\dim(\operatorname{Im} L) = 2 \quad \dim(\mathbb{R}^3) = 3,$$

quindi applicando il teorema delle dimensioni risulta

$$\dim(\operatorname{Ker} L) = \dim(\mathbb{R}^3) - \dim(\operatorname{Im} L) = 3 - 2 = 1,$$

infatti come abbiamo verificato $\operatorname{Ker} L$ è una retta per l'origine in $\mathbb{R}^3$.

(2) Sia $L\colon \mathbb{R}^3 \to \mathbb{R}$ l'applicazione data da:

$$L\begin{pmatrix} x \\ y \\ z \end{pmatrix} = x + y - z.$$

Verifichiamo che L è un'applicazione lineare: basta osservare che risulta

$$L\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 & 1 & -1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix},$$

L è lineare essendo l'applicazione associata alla matrice

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

Abbiamo:

$$\operatorname{Ker} L = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \,|\, x + y - z = 0 \right\},$$

$\operatorname{Ker} L$ è un piano passante per l'origine di $\mathbb{R}^3$, quindi $\dim(\operatorname{Ker} L) = 2$. Applicando il teorema delle dimensioni abbiamo:

$$\dim(\mathbb{R}^3) = 3 = \dim(\operatorname{Ker} L) + \dim(\operatorname{Im} L) = 2 + \dim(\operatorname{Im} L),$$

da cui ricaviamo che $\dim(\operatorname{Im} L) = 1$. Poiché $\operatorname{Im} L \subseteq \mathbb{R}$, abbiamo necessariamente

$$\operatorname{Im} L = \mathbb{R},$$

e quindi L è suriettiva.

Come applicazione del Teorema delle dimensioni possiamo provare il seguente risultato, che in precedenza ci siamo limitati ad enunciare e giustificare (Osservazione 3.33, Corollario 3.26):

PROPOSIZIONE 4.10. *Ogni matrice* ***reale*** $A \in \mathcal{M}_{\mathbb{R}}(m,n)$ *soddisfa la relazione:*

$$\operatorname{rg}(A) = \operatorname{rg}(A^{\mathrm{T}}),$$

dove A^{T} *è la matrice trasposta di* A.

DIMOSTRAZIONE. La matrice $A \in \mathcal{M}_{\mathbb{R}}(m,n)$ definisce l'applicazione lineare

$$L_A \colon \mathbb{R}^n \to \mathbb{R}^m, \quad L_A(X) = AX, \quad \forall X \in \mathbb{R}^n;$$

mentre la matrice trasposta A^{T} definisce l'applicazione lineare

$$L_{(A^{\mathrm{T}})} \colon \mathbb{R}^m \to \mathbb{R}^n, \quad L_{(A^{\mathrm{T}})}(Y) = (A^{\mathrm{T}})Y, \ \forall Y \in \mathbb{R}^m.$$

Consideriamo i seguenti sottospazi di $\mathbb{R}^m$: $\operatorname{Im} L_A$ e $\operatorname{Ker} L_{(A^{\mathrm{T}})}$. Verifichiamo che risulta:

$$\operatorname{Im} L_A \cap \operatorname{Ker} L_{(A^{\mathrm{T}})} = \{\mathbf{0}_{\mathbb{R}^m}\}.$$

Sia Y un vettore colonna di $\mathbb{R}^m$, Y è un elemento di $\mathcal{M}_{\mathbb{R}}(m,1)$, indichiamo con $Y^{\mathrm{T}} \in \mathcal{M}_{\mathbb{R}}(1,m)$ il vettore riga di $\mathbb{R}^m$ trasposto di Y. Osserviamo che il prodotto delle matrici Y^{T} e Y è il seguente numero reale

$$Y^{\mathrm{T}} \cdot Y = y_1^2 + y_2^2 + \cdots + y_m^2,$$

che risulta essere zero se e solo se $y_i = 0 \ \forall i = 1, \dots, m$, cioè $Y = \mathbf{0}_{\mathbb{R}^m}$.

Sia ora $Y \in \operatorname{Im} L_A \cap \operatorname{Ker} L_{(A^{\mathrm{T}})}$: poiché $Y \in \operatorname{Im} L_A$, esiste un vettore $X \in \mathbb{R}^n$ tale che $Y = AX$. Inoltre poiché $Y \in \operatorname{Ker} L_{(A^{\mathrm{T}})}$ verifica la condizione $(A^{\mathrm{T}})Y = \mathbf{0}_{\mathbb{R}^m}$. Calcoliamo il prodotto $Y^{\mathrm{T}} \cdot Y$, sostituendo AX al posto di Y e ricordando che $Y \in \operatorname{Ker} L_{(A^{\mathrm{T}})}$ si ha:

$$Y^{\mathrm{T}} \cdot Y = (AX)^{\mathrm{T}} Y = (X^{\mathrm{T}} A^{\mathrm{T}})Y = X^{\mathrm{T}}(A^{\mathrm{T}} Y) = X^{\mathrm{T}} \cdot \mathbf{0}_{\mathbb{R}^m} = 0.$$

Possiamo quindi concludere che Y è il vettore nullo di $\mathbb{R}^m$ e la somma dei sottospazi $\operatorname{Im} L_A$ e $\operatorname{Ker} L_{(A^{\mathrm{T}})}$ è diretta.

Poiché $\operatorname{Im} L_A \oplus \operatorname{Ker} L_{(A^{\mathrm{T}})}$ è un sottospazio di $\mathbb{R}^m$ si ha:

$$\dim(\operatorname{Im} L_A \oplus \operatorname{Ker} L_{(A^{\mathrm{T}})}) = \dim(\operatorname{Im} L_A) + \dim(\operatorname{Ker} L_{(A^{\mathrm{T}})}) \leq m.$$

Poiché $\dim(\operatorname{Im} L_A) = \operatorname{rg}(A)$ e applicando il teorema delle dimensioni (4.8)

$$\dim(\operatorname{Ker} L_{(A^{\mathrm{T}})}) = m - \operatorname{rg}(A^{\mathrm{T}}),$$

sostituendo ottenimamo quindi la disuguaglianza

$$\operatorname{rg}(A) \leq \operatorname{rg}(A^{\mathrm{T}}).$$

Scambiando i ruoli delle matrici A e A^{T} si ottiene in modo analogo:

$$\operatorname{rg}(A^{\mathrm{T}}) \leq \operatorname{rg}(A)$$

da cui ricaviamo l'uguaglianza tra i ranghi delle due matrici. □

Vale la pena di aggiungere anche la seguente

Osservazione 4.10. Sia $L\colon V \to W$ un'applicazione lineare tra spazi vettoriali reali di dimensione finita. Come conseguenza del Teorema delle dimensioni (4.8) abbiamo le seguenti proprietà:

(1) L è iniettiva $\iff$ $\dim(\operatorname{Im} L) = \dim V$;
(2) L è suriettiva $\iff$ $\dim(\operatorname{Im} L) = \dim W$;
(3) se $\dim V = \dim W$,

$$L \text{ è suriettiva } \iff L \text{ è iniettiva } \iff L \text{ è biunivoca.}$$

Dimostrazione. Infatti, ricordiamo che:
L è iniettiva $\iff$ $\operatorname{Ker} L = \{0_V\}$ $\iff$ $\dim(\operatorname{Ker} L) = 0$, dal teorema delle dimensioni (4.8) ciò succede se e solo se $\dim(\operatorname{Im} L) = \dim V$.
Ricordiamo che:
L è suriettiva $\iff$ $\operatorname{Im} L = W$ $\iff$ $\dim(\operatorname{Im} L) = \dim W$.
Se $\dim V = \dim W$, le condizioni scritte sopra coincidono, per cui le due proprietà sono equivalenti. □

3. Spazi vettoriali isomorfi.

Abbiamo visto che, fissando in uno spazio vettoriale reale V di dimensione n una base $\mathcal{B}_V$, possiamo associare in modo univoco ad ogni vettore $\boldsymbol{v} \in V$ una $n-$upla di numeri reali e quindi un vettore di $\mathbb{R}^n$: il vettore delle coordinate di v rispetto alla base fissata $\mathcal{B}_V$

$$X = [\boldsymbol{v}]_{\mathcal{B}_V}.$$

Abbiamo visto che possiamo in un certo senso "dimenticarci" della natura degli elementi di V e lavorare con i vettori delle coordinate. Siamo ora in grado di formalizzare questo discorso, introducendo la nozione di isomorfismo tra spazi vettoriali.

Definizione 4.4. Sia $L\colon V \to W$ un'applicazione tra due spazi vettoriali reali.

- L è un **isomorfismo** di V in W se è un'applicazione lineare biunivoca;
- V e W sono detti **spazi vettoriali isomorfi** se esiste un isomorfismo di V in W, scriviamo: $V \simeq W$.

Proposizione 4.11. *Sia $L\colon V \to W$ un'applicazione lineare tra spazi vettoriali reali di dimensioni finite.*

Se L è un isomorfismo, allora $\dim V = \dim W$.

Dimostrazione. Infatti se L è un isomorfismo, allora L è iniettiva e quindi $\dim(\operatorname{Im} L) = \dim V$; ma L è anche suriettiva, quindi $\dim(\operatorname{Im} L) = \dim W$, per cui possiamo concludere che i due spazi hanno la stessa dimensione. □

Ovviamente, la condizione è solo necessaria; però vale la seguente

Osservazione 4.11.

- Sia $L\colon V \to W$ un'applicazione lineare tra spazi vettoriali reali con la stessa dimensione: $\dim V = \dim W = n$. Dalle considerazioni precedenti, possiamo concludere che:
$$L \text{ isomorfismo} \iff L \text{ iniettiva} \iff L \text{ suriettiva}\ .$$
- Ricordando la Prop. 4.7, possiamo dire che un isomorfismo $L\colon V \to W$ *trasforma una base di V in una base di W*.

Esempio 4.12. Consideriamo lo spazio vettoriale $\mathcal{M}_{\mathbb{R}}(m,n)$ delle matrici reali con m righe ed n colonne. Sia $T\colon \mathcal{M}_{\mathbb{R}}(m,n) \to \mathcal{M}_{\mathbb{R}}(n,m)$ l'applicazione che associa ad ogni matrice A la sua trasposta:
$$T(A) = A^{\mathrm{T}}.$$
Ricordiamo che i vettori colonna di A sono i vettori riga della matrice A^{T}, abbiamo quindi:
$$(A^{\mathrm{T}})_{ij} = (A)_{ji} \quad \forall i = 1,\dots,n \;\; j = 1,\dots,m.$$
Verifichiamo che T è un isomorfismo di $\mathcal{M}_{\mathbb{R}}(m,n)$ in $\mathcal{M}_{\mathbb{R}}(n,m)$. Verifichiamo dapprima che T è un'applicazione lineare, cioè che valgono le seguenti proprietà:
$$T(A+B) = T(A) + T(B), \quad \forall A, B \in \mathcal{M}_{\mathbb{R}}(m,n),$$
$$T(\lambda A) = \lambda T(A), \quad \forall A \in \mathcal{M}_{\mathbb{R}}(m,n), \quad \forall \lambda \in \mathbb{R}.$$
Tali proprietà seguono immediatamente dalla definizione di trasposta, infatti ricordiamo che:
$$(A+B)^{\mathrm{T}} = A^{\mathrm{T}} + B^{\mathrm{T}}, \quad \forall A, B \in \mathcal{M}_{\mathbb{R}}(m,n);$$
$$(\lambda A)^{\mathrm{T}} = \lambda A^{\mathrm{T}}, \quad \forall A \in \mathcal{M}_{\mathbb{R}}(m,n), \quad \forall \lambda \in \mathbb{R}.$$
Verifichiamo infine che T è un isomorfismo. Poiché si ha
$$\dim(\mathcal{M}_{\mathbb{R}}(m,n)) = mn = \dim(\mathcal{M}_{\mathbb{R}}(n,m)),$$
basta verificare, per esempio, che T è iniettiva.
Sia $A \in \mathcal{M}_{\mathbb{R}}(m,n)$ una matrice tale che:
$$T(A) = A^{\mathrm{T}} = O_{n\times m},$$
allora risulta
$$(A^{\mathrm{T}})_{ij} = 0 \quad \forall i = 1,\dots,n, j = 1,\dots,m;$$
da cui deduciamo che
$$(A)_{ji} = 0 \quad \forall i = 1,\dots,n, j = 1,\dots,m,$$
cioè A è la matrice nulla, vale a dire $\operatorname{Ker} T = \{O_{m\times n}\}$, quindi T è iniettiva.
Possiamo quindi concludere che:
$$\mathcal{M}_{\mathbb{R}}(m,n) \simeq \mathcal{M}_{\mathbb{R}}(n,m).$$
In particolare, per $m = 1$, otteniamo che lo spazio vettoriale dei vettori *colonna* di $\mathbb{R}^n$ è isomorfo allo spazio vettoriale dei vettori *riga* di $\mathbb{R}^n$.

Proviamo ora che fissare una base in uno spazio vettoriale reale V di dimensione n, permette di definire in modo naturale un isomorfismo tra V e lo spazio vettoriale $\mathbb{R}^n$.

PROPOSIZIONE 4.12 (**Isomorfismo di rappresentazione**). *Siano V uno spazio vettoriale reale di dimensione n e $\mathcal{B} = \{\boldsymbol{v}_1, \dots, \boldsymbol{v}_n\}$ una base di V. Sia*

$$L_{\mathcal{B}} \colon V \to \mathbb{R}^n, \quad \boldsymbol{v} \to [\boldsymbol{v}]_{\mathcal{B}},$$

l'applicazione che associa ad ogni vettore di V la $n-$upla delle coordinate rispetto alla base $\mathcal{B}$.

$L_{\mathcal{B}}$ è un isomorfismo di V in $\mathbb{R}^n$.

DIMOSTRAZIONE. Consideriamo l'applicazione $L_{\mathcal{B}} \colon V \to \mathbb{R}^n$:

$$\boldsymbol{v} \to X = [\boldsymbol{v}]_{\mathcal{B}} \in \mathbb{R}^n,$$

dove X è il vettore delle coordinate di $\boldsymbol{v}$ nella base fissata. Ricordiamo che le coordinate di $\boldsymbol{v}$ sono i coefficienti dei vettori della base nella combinazione lineare seguente:

$$\boldsymbol{v} = x_1\boldsymbol{v}_1 + \dots + x_n\boldsymbol{v}_n.$$

Proviamo che $L_{\mathcal{B}}$ è un'applicazione lineare.

Basta provare che:

(1) $\forall \boldsymbol{v}, \boldsymbol{u} \in V$: $\quad [\boldsymbol{v} + \boldsymbol{u}]_{\mathcal{B}} = [\boldsymbol{v}]_{\mathcal{B}} + [\boldsymbol{u}]_{\mathcal{B}}$;
(2) $\forall \boldsymbol{v} \in V, \ \forall \alpha \in \mathbb{R}$: $\quad [\alpha\boldsymbol{v}]_{\mathcal{B}} = \alpha[\boldsymbol{v}]_{\mathcal{B}}$.

Verifichiamo la (1). Le coordinate di $\boldsymbol{v}$ e di $\boldsymbol{u}$ sono rispettivamente i coefficienti dei vettori della base nelle combinazioni lineari seguenti:

$$\boldsymbol{v} = x_1\boldsymbol{v}_1 + \dots + x_n\boldsymbol{v}_n,$$

$$\boldsymbol{u} = y_1\boldsymbol{v}_1 + \dots + y_n\boldsymbol{v}_n.$$

Sommando i vettori $\boldsymbol{v}$ e $\boldsymbol{u}$ e applicando le proprietà degli spazi vettoriali, otteniamo:

$$\begin{aligned} \boldsymbol{v} + \boldsymbol{u} &= (x_1\boldsymbol{v}_1 + \dots + x_n\boldsymbol{v}_n) + (y_1\boldsymbol{v}_1 + \dots + y_n\boldsymbol{v}_n), \\ &= (x_1 + y_1)\boldsymbol{v}_1 + \dots + (x_n + y_n)\boldsymbol{v}_n, \end{aligned}$$

da cui deduciamo che le coordinate di $\boldsymbol{v}+\boldsymbol{u}$ si ottengono sommando le coordinate di $\boldsymbol{v}$ e le coordinate di $\boldsymbol{u}$.

Verifichiamo la (2). Le coordinate di $\boldsymbol{v}$ sono i coefficienti dei vettori della base nella combinazione lineare seguente:

$$\boldsymbol{v} = x_1\boldsymbol{v}_1 + \dots + x_n\boldsymbol{v}_n,$$

moltiplicando $\boldsymbol{v}$ per il numero reale α e applicando le proprietà degli spazi vettoriali, otteniamo:

$$\alpha\boldsymbol{v} = \alpha(x_1\boldsymbol{v}_1 + \dots + x_n\boldsymbol{v}_n) = (\alpha x_1)\boldsymbol{v}_1 + \dots + (\alpha x_n)\boldsymbol{v}_n,$$

da cui deduciamo che le coordinate di $\alpha\boldsymbol{v}$ si ottengono moltiplicando per il numero reale α le coordinate di $\boldsymbol{v}$.

Poiché $\dim V = \dim \mathbb{R}^n = n$, per provare che $L_{\mathcal{B}}$ è un isomorfismo basta provare che $L_{\mathcal{B}}$ è iniettiva, cioè che risulta:

$$\operatorname{Ker} L_{\mathcal{B}} = \{\mathbf{0}_V\}.$$

Sia $\boldsymbol{v} \in \operatorname{Ker} L_{\mathcal{B}}$, si ha: $[\boldsymbol{v}]_{\mathcal{B}} = \mathbf{0}_{\mathbb{R}^n}$, cioè $\boldsymbol{v}$ è la combinazione banale dei vettori della base:

$$\boldsymbol{v} = 0\boldsymbol{v}_1 + \cdots + 0\boldsymbol{v}_n = \mathbf{0}_V,$$

ciò implica che $\boldsymbol{v} = \mathbf{0}_V$ e possiamo concludere che $L_{\mathcal{B}}$ è iniettiva (in effetti, è immediato dalla definizione di L che è iniettiva). □

Grazie all'isomorfismo $L_{\mathcal{B}}$ di V in $\mathbb{R}^n$, definito dalla base $\mathcal{B}$, possiamo effettivamente tradurre le proprietà algebriche dei vettori di V in analoghe proprietà delle corrispondenti $n-$uple di rappresentazione. Una di queste proprietà è la dipendenza lineare: s **vettori sono linearmente dipendenti in** V **se e solo se sono linearmente dipendenti i vettori (**$n-$**uple) delle loro coordinate in** $\mathbb{R}^n$.

Proprietà 4.13. *Siano V uno spazio vettoriale reale e $\mathcal{B}$ una base di V. I vettori $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_s\}$ sono linearmente dipendenti in V se e solo se i vettori delle loro coordinate $\{[\boldsymbol{u}_1]_{\mathcal{B}}, \ldots, [\boldsymbol{u}_s]_{\mathcal{B}}\}$ sono linearmente dipendenti in $\mathbb{R}^n$.*

Dimostrazione. Ricordiamo che i vettori $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_s\}$ sono linearmente dipendenti in V se e solo se esiste una loro combinazione lineare non banale che dà il vettore nullo di V:

$$\lambda_1\boldsymbol{u}_1 + \cdots + \lambda_s\boldsymbol{u}_s = \mathbf{0}_V. \tag{4.6}$$

Osserviamo che, per la linearità di $L_{\mathcal{B}}$ risulta:

$$[\lambda_1\boldsymbol{u}_1 + \cdots + \lambda_s\boldsymbol{u}_s]_{\mathcal{B}} = \lambda_1[\boldsymbol{u}_1]_{\mathcal{B}} + \cdots + \lambda_s[\boldsymbol{u}_s]_{\mathcal{B}}. \tag{4.7}$$

Ora, se esiste una combinazione lineare non banale dei vettori $\boldsymbol{u}_1, \ldots \boldsymbol{u}_k$ che produce il vettore nullo (4.6), l'uguaglianza 4.7 mostra che allora esiste una combinazione lineare non banale delle $n-$uple di coordinate dei vettori:

$$\lambda_1[\boldsymbol{u}_1]_{\mathcal{B}} + \cdots + \lambda_s[\boldsymbol{u}_s]_{\mathcal{B}} = [\mathbf{0}_V]_{\mathcal{B}} = \mathbf{0}_{\mathbb{R}^n}, \tag{4.8}$$

da cui deduciamo che i vettori $\{[\boldsymbol{u}_1]_{\mathcal{B}}, \ldots, [\boldsymbol{u}_s]_{\mathcal{B}}\}$ sono linearmente dipendenti in $\mathbb{R}^n$.

Viceversa, partendo dalla combinazione lineare non banale delle $n-$uple che produe la $n-$upla nulla, sempre tramite l'identità 4.7 arriviamo a trovare una combinazione lineare non banale dei vettori che produce il vettore nullo (Eq. 4.6), cioè i vettori $\boldsymbol{u}_1, \ldots \boldsymbol{u}_k$ sono linearmente dipendenti. □

Osservazione 4.13. Dato che la condizione di dipendenza lineare fra vettori e fra $n-$uple sono equivalenti, lo è anche l'indipendenza lineare.

Concludiamo con il seguente risultato:

Corollario 4.14. ***Ogni*** *spazio vettoriale reale V di dimensione n è isomorfo allo spazio $\mathbb{R}^n$.*

Esempio 4.14.

- Lo spazio vettoriale $\mathbb{E}^3_O$ è isomorfo a $\mathbb{R}^3$.
- Lo spazio vettoriale $\mathcal{M}_{\mathbb{R}}(2)$ è isomorfo a $\mathbb{R}^4$.
- Lo spazio vettoriale $\mathcal{M}_{\mathbb{R}}(m, n)$ è isomorfo a $\mathbb{R}^{mn}$.

- Proveremo che due spazi vettoriali della stessa dimensione sono isomorfi.

4. Applicazioni lineari e matrici.

D'ora in poi consideriamo spazi vettoriali di dimensione finita e fissiamo una base in ciascun spazio. Dimostriamo la seguente proprietà fondamentale: *ogni applicazione lineare $L\colon V \to W$ è completamente determinata dalle immagini dei vettori della base di V.* Più precisamente: sia $n = \dim V$, proviamo che fissato un sistema S_W di n vettori arbitrari in W esiste un'unica applicazione lineare che trasforma la base fissata $\mathcal{B}_V$ di V in S_W.

PROPOSIZIONE 4.15. *Siano V e W due spazi vettoriali reali di dimensione finita. Siano $\mathcal{B}_V = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n\}$ una base di V e $\{\boldsymbol{w}_1, \dots, \boldsymbol{w}_n\}$ un insieme di n vettori di W. Esiste un'unica applicazione lineare $L\colon V \to W$ tale che $L(\boldsymbol{v}_i) = \boldsymbol{w}_i$, $i = 1, \dots, n$.*

DIMOSTRAZIONE. Definiamo l'applicazione $L\colon V \to W$ nel seguente modo:

$$\forall \boldsymbol{v} \in V \quad L(\boldsymbol{v}) = x_1\boldsymbol{w}_1 + \cdots + x_n\boldsymbol{w}_n,$$

dove $x_1, x_2, \dots, x_n$ sono le coordinate di $\boldsymbol{v}$ nella base fissata $\mathcal{B}_V = \{\boldsymbol{v}_1, \dots, \boldsymbol{v}_n\}$, cioè $x_1, x_2, \dots, x_n$ sono gli scalari che compaiono nella scrittura:

$$\boldsymbol{v} = x_1\boldsymbol{v}_1 + \cdots + x_n\boldsymbol{v}_n.$$

Proviamo che L è lineare. Siano $\boldsymbol{v} = x_1\boldsymbol{v}_1 + \cdots + x_n\boldsymbol{v}_n$ e $\boldsymbol{u} = y_1\boldsymbol{v}_1 + \cdots + y_n\boldsymbol{v}_n$, allora:

$$\boldsymbol{v} + \boldsymbol{u} = (x_1 + y_1)\boldsymbol{v}_1 + \cdots + (x_n + y_n)\boldsymbol{v}_n,$$

quindi abbiamo

$$L(\boldsymbol{v} + \boldsymbol{u}) = (x_1 + y_1)\boldsymbol{w}_1 + \cdots + (x_n + y_n)\boldsymbol{w}_n =$$
$$= (x_1\boldsymbol{w}_1 + \cdots + x_n\boldsymbol{w}_n) + (y_1\boldsymbol{w}_1 + \cdots + y_n\boldsymbol{w}_n) = L(\boldsymbol{v}) + L(\boldsymbol{u}).$$

Siano ora $\boldsymbol{v} = x_1\boldsymbol{v}_1 + \cdots + x_n\boldsymbol{v}_n$ e $\alpha \in \mathbb{R}$, allora:

$$\alpha\boldsymbol{v} = (\alpha x_1)\boldsymbol{v}_1 + \cdots + (\alpha x_n)\boldsymbol{v}_n,$$

per cui abbiamo:

$$L(\alpha\boldsymbol{v}) = (\alpha x_1)\boldsymbol{w}_1 + \cdots + (\alpha x_n)\boldsymbol{w}_n = \alpha(x_1\boldsymbol{w}_1 + \cdots + x_n\boldsymbol{w}_n) = \alpha L(\boldsymbol{v}).$$

Inoltre, $\forall i = 1, \dots, n$ risulta:

$$L(\boldsymbol{v}_i) = 0\boldsymbol{w}_1 + 0\boldsymbol{w}_2 + \cdots + 1\boldsymbol{w}_i + \cdots + 0\boldsymbol{w}_n = \boldsymbol{w}_i.$$

Proviamo infine l'unicità: verifichiamo che se $T\colon V \to W$ è un'applicazione lineare che verifica le condizioni

$$T(\boldsymbol{v}_i) = \boldsymbol{w}_i, \quad i = 1, \dots n,$$

allora T coincide con l'applicazione L. Infatti, $\forall \boldsymbol{v} \in V$ abbiamo per la linearità di T:

$$T(\boldsymbol{v}) = T(x_1\boldsymbol{v}_1 + \cdots + x_n\boldsymbol{v}_n) = x_1T(\boldsymbol{v}_1) + \dots x_nT(\boldsymbol{v}_n) =$$

$$= x_1\boldsymbol{w}_1 + \dots x_n\boldsymbol{w}_n = L(\boldsymbol{v}),$$

poiché $T(\boldsymbol{v}) = L(\boldsymbol{v})\ \forall \boldsymbol{v} \in V$, le applicazioni coincidono. $\square$

NOTA BENE: In altre parole, un'applicazione lineare $L\colon V \to W$ è univocamente determinata conoscendo le immagini dei vettori della base scelta in V!

Esempio 4.15. Fissata la base canonica $\{\boldsymbol{e}_1, \boldsymbol{e}_2\}$ di $\mathbb{R}^2$, determinare l'operatore lineare $L\colon \mathbb{R}^2 \to \mathbb{R}^2$ tale che:

$$L(\boldsymbol{e}_1) = \boldsymbol{e}_1 + 2\boldsymbol{e}_2 \quad L(\boldsymbol{e}_2) = -\boldsymbol{e}_1.$$

Ricordiamo che ogni vettore $\boldsymbol{v} \in \mathbb{R}^2$ si scrive in uno e un solo modo come combinazione lineare di $\boldsymbol{e}_1$ e $\boldsymbol{e}_2$:

$$\boldsymbol{v} = x\boldsymbol{e}_1 + y\boldsymbol{e}_2.$$

L'unico operatore lineare L che soddisfa le condizioni precedenti è il seguente:

$$L(\boldsymbol{v}) = xL(\boldsymbol{e}_1) + yL(\boldsymbol{e}_2) = x(\boldsymbol{e}_1 + 2\boldsymbol{e}_2) + y(-\boldsymbol{e}_1) = (x - y)\boldsymbol{e}_1 + 2x\boldsymbol{e}_2,$$

da cui ricaviamo l'spressione di L

$$L\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x - y \\ 2x \end{pmatrix}.$$

Siamo ora in grado di provare il seguente risultato precedentemente enunciato:

Proprietà 4.16. *Due spazi vettoriali reali V e W con* $\dim V = \dim W$ *sono isomorfi.*

Dimostrazione. Basta provare che esiste un isomorfismo $L\colon V \to W$. Siano $\mathcal{B}_V = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n\}$ una base di V e $\{\boldsymbol{w}_1, \dots, \boldsymbol{w}_n\}$ una base di W. Per la proposizione precedente esiste un'unica applicazione lineare $L\colon V \to W$ tale che $L(\boldsymbol{v}_i) = \boldsymbol{w}_i$, $i = 1, \dots, n$. Si ha:

$$\operatorname{Im} L = \operatorname{Span}(L(\boldsymbol{v}_1), \dots, L(\boldsymbol{v}_n)) = \operatorname{Span}(\boldsymbol{w}_1, \dots, \boldsymbol{w}_n) = W,$$

da cui ricaviamo che L è suriettiva. Poiché $\dim V = \dim W$, L è anche iniettiva, (oss. 4.11), possiamo quindi concludere che L è un isomorfismo. $\square$

Ricordiamo che fissare una base $\mathcal{B}$ in uno spazio vettoriale reale di dimensione n significa associare ad ogni vettore una n-upla di scalari, *le coordinate del vettore* nella base fissata, quindi un vettore di $\mathbb{R}^n$. Il vantaggio di usare le coordinate è che ogni applicazione lineare può essere scritta in *forma matriciale*, si ha infatti il fondamentale risultato:

Teorema 4.17 (**Fondamentale di Rappresentazione**). *Siano V e W due spazi vettoriali reali di dimensione finita. Siano $\mathcal{B}_V = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n\}$ una base di V e $\mathcal{B}_W = \{\boldsymbol{w}_1, \boldsymbol{w}_2, \dots, \boldsymbol{w}_m\}$ una base di W. Data un'applicazione lineare $L\colon V \to W$ esiste un'unica matrice*

$$A \in \mathcal{M}_{\mathbb{R}}(m, n)$$

tale che l'espressione di L in coordinate è la seguente:

$$Y = AX,$$

dove X è il vettore delle coordinate di $\boldsymbol{v}$ nella base $\mathcal{B}_V$ ed Y è il vettore delle coordinate di $L(\boldsymbol{v})$ nella base $\mathcal{B}_W$:

$$X = [\boldsymbol{v}]_{\mathcal{B}_V} \quad Y = [L(\boldsymbol{v})]_{\mathcal{B}_W}.$$

La matrice A è detta ***matrice associata*** *a L o* ***matrice di rappresentazione***, *o anche* ***matrice rappresentativa*** *di L nelle basi $\mathcal{B}_V$ e $\mathcal{B}_W$.*
Sia A^i la i-esima colonna della matrice A, A^i è costituita dalle coordinate del vettore $L(\boldsymbol{v}_i)$ nella base $\mathcal{B}_W$, $i = 1, 2 \dots, n$:

$$A^i = [L(\boldsymbol{v}_i)]_{\mathcal{B}_W}, \quad i = 1, \dots, n.$$

DIMOSTRAZIONE. Siano $\boldsymbol{v} \in V$ e $X \in \mathbb{R}^n$ il vettore delle coordinate di $\boldsymbol{v}$ rispetto alla base $\mathcal{B}_V$, possiamo scrivere:

$$\boldsymbol{v} = x_1\boldsymbol{v}_1 + x_2\boldsymbol{v}_2 + \cdots + x_n\boldsymbol{v}_n.$$

Per la linearità, applicando L otteniamo la relazione seguente in W:

$$L(\boldsymbol{v}) = x_1L(\boldsymbol{v}_1) + x_2L(\boldsymbol{v}_2) + \cdots + x_nL(\boldsymbol{v}_n).$$

Consideriamo ora le coordinate rispetto alla base $\mathcal{B}_W$, per l'isomorfismo di rappresentazione, otteniamo la relazione seguente:

$$[L(\boldsymbol{v})]_{\mathcal{B}_W} = x_1[L(\boldsymbol{v}_1)]_{\mathcal{B}_W} + x_2[L(\boldsymbol{v}_2)]_{\mathcal{B}_W} + \cdots + x_n[L(\boldsymbol{v}_n)]_{\mathcal{B}_W}.$$

Posto $A^i = [L(\boldsymbol{v}_i)]_{\mathcal{B}_W}$, $\quad i = 1, \dots, n$, e $Y = [L(\boldsymbol{v})]_{\mathcal{B}_W}$ la relazione precedente diventa:

$$Y = x_1A^1 + x_2A^2 + \cdots + x_nA^n = AX,$$

dove A è la matrice di colonne $A^1, \dots, A^n$. $\square$

NOTA BENE: La matrice A **dipende dalla scelta delle basi** in V e in W!

ESEMPIO 4.16.

(1) Sia $L \colon \mathbb{R}^2 \to \mathbb{R}^3$ l'applicazione lineare data da:

$$L\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x - 2y \\ -x + y \\ -4x \end{pmatrix}.$$

Scriviamo la matrice associata a L nelle basi canoniche di $\mathbb{R}^2$ e $\mathbb{R}^3$. Fissata $\mathcal{B} = \{\boldsymbol{e}_1, \boldsymbol{e}_2\}$ la base standard di $\mathbb{R}^2$, abbiamo:

$$L(\boldsymbol{e}_1) = L\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ -4 \end{pmatrix} \quad L(\boldsymbol{e}_2) = L\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}.$$

La matrice associata a L è la matrice $A \in \mathcal{M}_{\mathbb{R}}(3, 2)$ le cui colonne sono le coordinate, nella base canonica di $\mathbb{R}^3$, di $L(\boldsymbol{e}_1)$ e $L(\boldsymbol{e}_2)$:

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

(2) Sia $I_V : V \to V$ l'operatore identità di V. Fissata una stessa base $\mathcal{B}_V = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n\}$ nel dominio e codominio, scriviamo la matrice associata a I_V. Abbiamo:

$$I_V(\boldsymbol{v}_1) = \boldsymbol{v}_1, I_V(\boldsymbol{v}_2) = \boldsymbol{v}_2 \quad \dots I_V(\boldsymbol{v}_n) = \boldsymbol{v}_n,$$

da cui ricaviamo il vettore delle coordinate di $I_V(\boldsymbol{v}_i)$, per ogni $i = 1, \dots, n$:

$$[I_V(\boldsymbol{v}_i)]_{\mathcal{B}_V} = [\boldsymbol{v}_i]_{\mathcal{B}_V} = \boldsymbol{e}_i,$$

dove $\boldsymbol{e}_i$ indica l'i-esimo vettore della base canonica di $\mathbb{R}^n$. Possiamo concludere che la matrice associata a I_V è la matrice identità di ordine n:

$$A = I_n.$$

In altre parole, l'operatore identità in V è sempre rappresentato dalla matrice

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

purché la base usata sia la stessa nel dominio e nel codominio.

(3) Sia $L : \mathbb{E}^2_O \to \mathbb{E}^2_O$ la rotazione in senso antiorario attorno all'origine dell'angolo $\theta = \frac{\pi}{2}$. Verifichiamo che L è un operatore lineare di $\mathbb{E}^2_O$ e fissata la base canonica di $\mathbb{E}^2_O$ scriviamo la rappresentazione matriciale di L.

Verifichiamo che L è lineare. Siano $\boldsymbol{v}_1, \boldsymbol{v}_2 \in \mathbb{E}^2_O$, consideriamo il parallelogramma $\mathcal{P}$ di lati $\boldsymbol{v}_1$ e $\boldsymbol{v}_2$. Sia $L(\mathcal{P})$ l'immagine di $\mathcal{P}$ secondo la L: osserviamo che questo è il parallelogramma di lati $L(\boldsymbol{v}_1)$ e $L(\boldsymbol{v}_2)$. Osserviamo che la diagonale per O del parallelogramma $L(\mathcal{P})$ è la rotazione della diagonale per O del parallelogramma $\mathcal{P}$, quindi abbiamo:

$$L(\boldsymbol{v}_1 + \boldsymbol{v}_2) = L(\boldsymbol{v}_1) + L(\boldsymbol{v}_2), \quad \forall \boldsymbol{v}_1, \boldsymbol{v}_2 \in \mathbb{E}^2_O$$

(Figura 4.1). Siano ora $\alpha \in \mathbb{R}$ e $\boldsymbol{v} \in \mathbb{E}^2_O$, osserviamo che la rotazione del vettore $\alpha\boldsymbol{v}$ ha la direzione della rotazione di $\boldsymbol{v}$, in particolare si verifica che risulta:

$$L(\alpha\boldsymbol{v}) = \alpha L(\boldsymbol{v}), \quad \forall \alpha \in \mathbb{R},\ \forall \boldsymbol{v} \in \mathbb{E}^2_O.$$

Sia $\mathcal{B} = \{\boldsymbol{e}_1, \boldsymbol{e}_2\}$ la base canonica di $\mathbb{E}^2_O$. Calcoliamo le immagini dei vettori della base:

$$L(\boldsymbol{e}_1) = \boldsymbol{e}_2 \quad L(\boldsymbol{e}_2) = -\boldsymbol{e}_1.$$

Sia $A \in \mathcal{M}_{\mathbb{R}}(2)$ la matrice associata a L nella base fissata $\mathcal{B}$, le colonne di A sono rispettivamente i vettori delle coordinate di $L(\boldsymbol{e}_1)$ e $L(\boldsymbol{e}_2)$ nella base $\mathcal{B}$:

$$A^1 = [L(\boldsymbol{e}_1)]_{\mathcal{B}} = [\boldsymbol{e}_2]_{\mathcal{B}},$$
$$A^2 = [L(\boldsymbol{e}_2)]_{\mathcal{B}} = [-\boldsymbol{e}_1]_{\mathcal{B}},$$

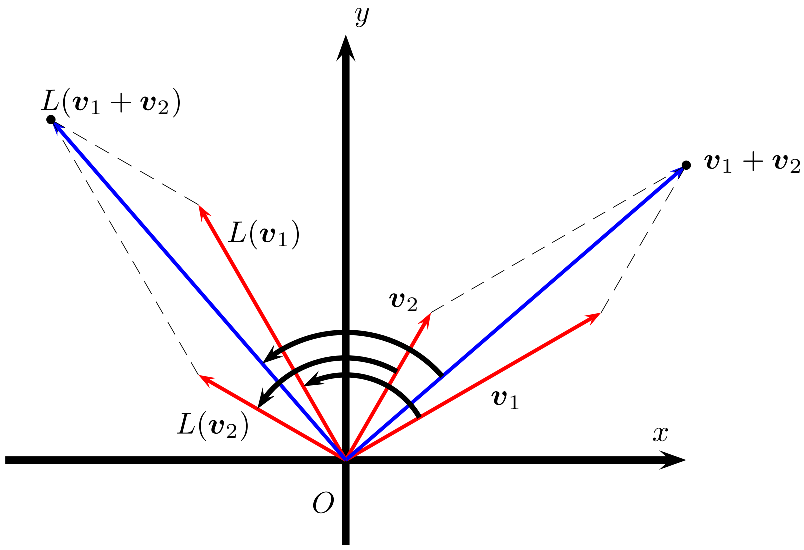

FIGURA 4.1. I vettori $\boldsymbol{v}_1$ e $\boldsymbol{v}_2$ definiscono un parallelogramma $\mathcal{P}$. Applicando la L che impone una rotazione oraria di $\pi/2$, si osserva subito che $L(\boldsymbol{v}_1 + \boldsymbol{v}_2) = L(\boldsymbol{v}_1) + L(\boldsymbol{v}_2)$ per ogni scelta di $\boldsymbol{v}_1$ e $\boldsymbol{v}_2$.

da cui ricaviamo

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

Nella base fissata $\mathcal{B}$ l'operatore L ha la seguente espressione matriciale:

$$Y = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} X, \quad \forall X \in \mathbb{R}^2.$$

La rappresentazione matriciale di un'applicazione lineare ci consente di lavorare negli spazi $\mathbb{R}^n$, tale rappresentazione ci consente di ricavare le seguenti proprietà degli spazi nucleo e immagine:

PROPRIETÀ 4.18. *Sia $L\colon V \to W$ un'applicazione lineare tra due spazi vettoriali reali di dimensione finita. Siano $\mathcal{B}_V = \{\boldsymbol{v}_1, \dots, \boldsymbol{v}_n\}$ e $\mathcal{B}_W = \{\boldsymbol{w}_1, \dots, \boldsymbol{w}_m\}$ rispettivamente una base di V e una base di W e $A \in \mathcal{M}_{\mathbb{R}}(m, n)$ la matrice associata a L nelle basi fissate. Risulta allora:*

(1) $\operatorname{Ker} L = \{\boldsymbol{v} \in V \mid [\boldsymbol{v}]_{\mathcal{B}_V} = X \quad A \cdot X = \mathbf{0}_{\mathbb{R}^m}\}$;
(2) $\operatorname{Im} L = \{\boldsymbol{w} \in W \mid [\boldsymbol{w}]_{\mathcal{B}_W} = Y \quad Y = AX, \quad X \in \mathbb{R}^n\}$;
(3) $\dim(\operatorname{Im} L) = \operatorname{rg}(A)$.

DIMOSTRAZIONE. Indicati con X e Y rispettivamente i vettori delle coordinate di $\boldsymbol{v}$ e $L(\boldsymbol{v})$ nelle basi fissate in V e W, si ha:

$$Y = AX.$$

Ricordiamo la definizione del sottospazio nucleo di L (Definizione 4.3):

$$\operatorname{Ker} L = \{\boldsymbol{v} \in V \mid \ L(\boldsymbol{v}) = \mathbf{0}_V \ \},$$

abbiamo quindi:

$\boldsymbol{v} \in \operatorname{Ker} L \iff L(\boldsymbol{v}) = \mathbf{0}_V \iff Y = \mathbf{0}_{\mathbb{R}^m} \iff A \cdot X = \mathbf{0}_{\mathbb{R}^m}.$

Ricordiamo la definizione del sottospazio immagine di L (Definizione 0.9):

$$\operatorname{Im} L = \{L(\boldsymbol{v}),\, \boldsymbol{v} \in V\} = \{\boldsymbol{w} \in W \mid \exists \boldsymbol{v} \in V\colon L(\boldsymbol{v}) = \boldsymbol{w}\},$$

poniamo $X = [\boldsymbol{v}]_{\mathcal{B}_V}$ e $Y = [\boldsymbol{w}]_{\mathcal{B}_W}$, abbiamo che $w \in \operatorname{Im} L$ se e solo se $Y = AX$, con $X \in \mathbb{R}^n$.

Ricordiamo infine che

$$\operatorname{Im} L = \operatorname{Span}(L(\boldsymbol{v}_1), \dots, L(\boldsymbol{v}_n)),$$

inoltre per ogni $i = 1, \dots, n$, la colonna A^i della matrice A è il vettore delle coordinate di $L(\boldsymbol{v}_i)$ nella base $\mathcal{B}_W$:

$$A^i = [L(\boldsymbol{v}_i)]_{\mathcal{B}_W}, \quad \forall i = 1, \dots, n.$$

Per concludere la dimostrazione basta provare che gli spazi vettoriali

$$\operatorname{Span}(L(\boldsymbol{v}_1), \dots, L(\boldsymbol{v}_n)) \;\text{ e }\; \operatorname{Span}(A^1, \dots, A^n)$$

hanno la stessa dimensione:

$$\dim(\operatorname{Span}(L(\boldsymbol{v}_1), \dots, L(\boldsymbol{v}_n))) = \dim(\operatorname{Span}(A^1, \dots, A^n)) = \operatorname{rg}(A).$$

Ma ciò segue dalla ben nota proprietà (4.13):

• *i vettori* $\{\boldsymbol{u}_1, \dots, \boldsymbol{u}_s\}$ *sono linearmente dipendenti in* W *se e solo se i vettori delle loro coordinate* $\{[\boldsymbol{u}_1]_{\mathcal{B}_W}, \dots, [\boldsymbol{u}_s]_{\mathcal{B}_W}\}$ *sono linearmente dipendenti in* $\mathbb{R}^m$. □

Esempio 4.17.

Sia $L\colon \mathbb{R}^3 \to \mathbb{R}^3$ l'operatore lineare dato da

$$L\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2x - y + z \\ y + z \\ x + z \end{pmatrix}.$$

(1) Scrivere la matrice associata a L rispetto alla base canonica di $\mathbb{R}^3$;
(2) calcolare le dimensioni degli spazi $\operatorname{Ker} L$ e $\operatorname{Im} L$;
(3) scrivere le equazioni degli spazi $\operatorname{Ker} L$ e $\operatorname{Im} L$.

Sia $\mathcal{B} = \{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ la base standard di $\mathbb{R}^3$. Determiniamo le immagini dei vettori della base:

$$L(\boldsymbol{e}_1) = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} \quad L(\boldsymbol{e}_2) = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \quad L(\boldsymbol{e}_3) = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix},$$

la matrice associata a L è quindi la seguente matrice $A \in \mathcal{M}_{\mathbb{R}}(3)$:

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

L'espressione matriciale di L è la seguente:

$$L\begin{pmatrix} x \\ y \\ z \end{pmatrix} = A\begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

Osserviamo che risulta $|A| = 0$, inoltre le colonne A^1 e A^2 sono linearmente indipendenti. Possiamo quindi concludere che $\dim(\operatorname{Im} L) = \operatorname{rg}(A) = 2$ e quindi $\operatorname{Im} L$ è un piano per l'origine in $\mathbb{R}^3$, più precisamente

$$\operatorname{Im} L = \operatorname{Span}(A^1, A^2) = \operatorname{Span}(L(\boldsymbol{e}_1), L(\boldsymbol{e}_2)).$$

Abbiamo quindi la seguente rappresentazione parametrica di $\operatorname{Im} L$:

$$\begin{cases} x = 2t - s \\ y = s & \qquad s, t \in \mathbb{R} \\ z = t, \end{cases}$$

da cui ricaviamo l'equazione cartesiana di $\operatorname{Im} L$:

$$x + y - 2z = 0.$$

Applicando il teorema delle dimensioni (4.8) abbiamo:

$$\dim \mathbb{R}^3 = \dim(\operatorname{Ker} L) + \dim(\operatorname{Im} L),$$

da cui ricaviamo:

$$\dim(\operatorname{Ker} L) = 3 - \dim(\operatorname{Im} L) = 1.$$

Possiamo concludere che $\operatorname{Ker} L$ è una retta per l'origine in $\mathbb{R}^3$. Per ottenere le equazioni cartesiane di $\operatorname{Ker} L$, osserviamo che i punti di $\operatorname{Ker} L$ sono le soluzioni del sistema lineare omogeneo:

$$A \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

dato dalle 3 equazioni seguenti:

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

Osserviamo che le prime due equazioni sono indipendenti e la terza equazione è combinazione lineare delle altre, quindi è superflua. Le equazioni cartesiane di $\operatorname{Im} L$ sono quindi:

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

5. Matrici associate alla stessa applicazione lineare

Abbiamo visto che ad ogni applicazione lineare tra due spazi vettoriali di dimensione finita, fissate una base in ciascuno spazio, possiamo associare una matrice e scrivere l'applicazione in forma matriciale. La matrice associata, quindi, **dipende dalla scelta della base** in ciascun spazio, in generale coppie di basi diverse danno matrici diverse. Vediamo in dettaglio un semplice esempio.

Esempio 4.18. Fissato $\mathcal{R}(O, \hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}})$, sistema di riferimento cartesiano ortogonale nel piano, sia $L\colon \mathbb{E}^2_O \to \mathbb{E}^2_O$ l'operatore lineare che associa ad un vettore applicato del piano $\boldsymbol{v} = \overrightarrow{OP}$ il vettore $L(\boldsymbol{v}) = \overrightarrow{OP'}$ dove P' è il simmetrico di P rispetto alla bisettrice del I e III quadrante. Fissiamo la base canonica di $\mathbb{E}^2_O$: $\mathcal{B}_1 = \{\hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}\}$. Abbiamo:

$$L(\hat{\boldsymbol{\imath}}) = \hat{\boldsymbol{\jmath}} \quad L(\hat{\boldsymbol{\jmath}}) = \hat{\boldsymbol{\imath}},$$

otteniamo quindi che la matrice associata a L nella base fissata $\mathcal{B}_1$ è la seguente:

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

Consideriamo ora una base $\mathcal{B}_2 = \{\boldsymbol{v}_1, \boldsymbol{v}_2\}$ di $\mathbb{E}^2_O$ data da due vettori indipendenti:

$$\boldsymbol{v}_1 = a_1\hat{\boldsymbol{\imath}} + b_1\hat{\boldsymbol{\jmath}} \quad \boldsymbol{v}_2 = a_2\hat{\boldsymbol{\imath}} + b_2\hat{\boldsymbol{\jmath}}.$$

Abbiamo:

$$L(\boldsymbol{v}_1) = L(a_1\hat{\boldsymbol{\imath}} + b_1\hat{\boldsymbol{\jmath}}) = a_1L(\hat{\boldsymbol{\imath}}) + b_1L(\hat{\boldsymbol{\jmath}}) = a_1\hat{\boldsymbol{\jmath}} + b_1\hat{\boldsymbol{\imath}},$$

$$L(\boldsymbol{v}_2) = L(a_2\hat{\boldsymbol{\imath}} + b_2\hat{\boldsymbol{\jmath}}) = a_2L(\hat{\boldsymbol{\imath}}) + b_2L(\hat{\boldsymbol{\jmath}}) = a_2\hat{\boldsymbol{\jmath}} + b_2\hat{\boldsymbol{\imath}}.$$

La matrice associata a L nella base $\mathcal{B}_2$ è la seguente:

$$B = \begin{pmatrix} b_1 & b_2 \\ a_1 & a_2 \end{pmatrix}.$$

Osserviamo che, in generale, le matrici A e B non coincidono: basta scegliere, per esempio $b_1 \neq 0$ o $a_2 \neq 0$, .

Sia $L\colon V \to W$ un'applicazione lineare tra due spazi vettoriali di dimensione finita. Siano $A \in \mathcal{M}_{\mathbb{R}}(m, n)$ la matrice associata a L rispetto alle basi $\mathcal{B}_V$ e $\mathcal{B}_W$ e $B \in \mathcal{M}_{\mathbb{R}}(m, n)$ la matrice associata a L rispetto alle basi $\mathcal{B}'_V$ e $\mathcal{B}'_W$. Ci proponiamo di ricavare una relazione tra le matrici A e B.
Per ogni vettore $\boldsymbol{v} \in V$, consideriamo i vettori delle coordinate di $\boldsymbol{v}$ nelle basi fissate in V:

$$X = [\boldsymbol{v}]_{\mathcal{B}_V} \quad X' = [\boldsymbol{v}]_{\mathcal{B}'_V}.$$

Indicata con $N \in \mathcal{GL}(n, \mathbb{R})$ la matrice invertibile che realizza il cambiamento di basi in V, i vettori delle coordinate nelle due basi sono legati dalla relazione:

$$X = NX'.$$

Analogamente, per ogni vettore $L(\boldsymbol{v}) \in \operatorname{Im} L$, consideriamo i vettori delle coordinate di $L(\boldsymbol{v})$ nelle basi fissate in W:

$$Y = [L(\boldsymbol{v})]_{\mathcal{B}_W} \quad Y' = [L(\boldsymbol{v})]_{\mathcal{B}'_W}.$$

Indicata con $M \in \mathcal{GL}(m, \mathbb{R})$ la matrice invertibile che realizza il cambiamento di basi in W, i vettori delle coordinate nelle due basi sono legati dalla relazione:

$$Y = MY'.$$

Ricordiamo infine che, fissate rispettivamente in V e W le basi $\mathcal{B}_V$ e $\mathcal{B}_W$, l'espressione matriciale di L è la seguente:

$$Y = AX,$$

mentre fissate le basi $\mathcal{B}'_V$ e $\mathcal{B}'_W$ l'espressione matriciale di L è la seguente:
$$Y' = BX'.$$
Sostituendo nella prima espressione matriciale di L le relazioni $X = NX'$ e $Y = MY'$ otteniamo:
$$NY' = A(MX'),$$
moltiplicando entrambi i membri a sinistra per la matrice N^{-1} otteniamo:
$$Y' = (N^{-1}AM)X.$$
Poiché la matrice associata a L nelle basi $\mathcal{B}'_V$ e $\mathcal{B}'_W$ è univocamente determinata, abbiamo necessariamente l'uguaglianza:
$$B = N^{-1}AM.$$
Abbiamo quindi provato il seguente risultato:

PROPOSIZIONE 4.19. *Sia $L \colon V \to W$ un'applicazione lineare tra spazi vettoriali reali di dimensione finita. Siano $A \in \mathcal{M}_{\mathbb{R}}(m,n)$ la matrice associata a L rispetto alle basi $\mathcal{B}_V$ e $\mathcal{B}_W$ e $B \in \mathcal{M}_{\mathbb{R}}(m,n)$ la matrice associata a L rispetto alle basi $\mathcal{B}'_V$ e $\mathcal{B}'_W$. La relazione tra le matrici A e B è la seguente:*
$$B = N^{-1}AM\,, \tag{4.9}$$
dove $N \in \mathcal{GL}(n,\mathbb{R})$ è la matrice invertibile che realizza il cambiamento di basi in V e $M \in \mathcal{GL}(m,\mathbb{R})$ la matrice invertibile che realizza il cambiamento di basi in W.

COROLLARIO 4.20. *Sia $L \colon V \to V$ un operatore lineare di uno spazio vettoriale V di dimensione n. Siano $A \in \mathcal{M}_{\mathbb{R}}(n)$ la matrice associata a L rispetto alla base $\mathcal{B}_V$ e $B \in \mathcal{M}_{\mathbb{R}}(n)$ la matrice associata a L rispetto alla base $\mathcal{B}'_V$. La relazione tra le matrici A e B risulta essere la seguente:*
$$B = N^{-1}AN,$$
dove $N \in \mathcal{GL}(n,\mathbb{R})$ è la matrice invertibile che realizza il cambiamento di basi in V.

Nel seguito studieremo in dettaglio le proprietà di un operatore lineare e saremo particolarmente interessati a cercare cambiamenti di base che ne possano semplificare l'espressione matriciale. Per questo motivo ci soffermiamo sulla relazione che abbiamo ora ricavato.

DEFINIZIONE 4.5. Siano $A, B \in \mathcal{M}_{\mathbb{R}}(n)$ due matrice quadrate reali di ordine n, B è **simile** ad A se esiste una matrice invertibile $N \in \mathcal{GL}(n,\mathbb{R})$ tale che
$$B = N^{-1}AN.$$

OSSERVAZIONE 4.19. Nell'anello $\mathcal{M}_{\mathbb{R}}(n)$ (Proposizione 3.5) la relazione di similitudine di matrici ora introdotta gode delle seguenti proprietà:

(1) *Proprietà riflessiva*: ogni matrice A è simile a se stessa.
Basta prendere $N = I_n$.

(2) *Proprietà simmetrica*: se B è simile ad A, anche A è simile a B.
Infatti dall'uguaglianza $B = N^{-1}AN$, moltiplicando a destra per la matrice N^{-1} e a sinistra per la matrice N, otteniamo: $NBN^{-1} = A$.

(3) *Proprietà transitiva*: se B è simile ad A e A è simile a C, allora B è simile a C.
Abbiamo per definizione:

$$B = N^{-1}AN \quad A = M^{-1}CM, \quad N, M \in \mathcal{GL}(n, \mathbb{R}).$$

Sostituendo la seconda espressione nella prima otteniamo:

$$B = N^{-1}(M^{-1}CM)N = (N^{-1}M^{-1})C(MN) = (MN)^{-1}C(MN),$$

da cui ricaviamo che B è simile a C.

Grazie alla proprietà di simmetria, d'ora in poi diremo semplicemente che le matrici A e B *sono simili*.

Possiamo quindi affermare che la relazione di similitudine tra matrici quadrate è una *relazione di equivalenza* (Definizione 0.6).

Osservazione 4.20. Siano $A, B \in \mathcal{M}_{\mathbb{R}}(n)$, ossiamo che *$A$ e B sono simili se e solo se rappresentano lo stesso operatore lineare in basi diverse.*
Abbiamo verificato che se A e B sono le matrici associate ad un operatore in basi diverse allora A e B sono simili.
Viceversa, siano ora A e B due matrici simili. Consideriamo l'operatore

$$L_A \colon \mathbb{R}^n \to \mathbb{R}^n$$

dato da $L_A(X) = AX$. La matrice A è la matrice associata a L_A nella base canonica $\mathcal{B}$ di $\mathbb{R}^n$. Consideriamo ora il cambiamento di coordinate in $\mathbb{R}^n$ di equazione $X = NX'$, che corrisponde a fissare la nuova base $\mathcal{B}' = \{N^1, \dots, N^n\}$, dove $N^1, \dots, N^n \in \mathbb{R}^n$ sono i vettori colonna della matrice N. La matrice associata all'operatore L_A nella base $\mathcal{B}'$ è la matrice $B = N^{-1}AN$. Le matrici rappresentano quindi lo stesso operatore lineare in basi diverse.

Ricordando l'esempio 4 dell'elenco 4.4, definiamo la traccia di una matrice.

Definizione 4.6. Sia $A \in \mathcal{M}_{\mathbb{R}}(n)$ una matrice reale quadrata; chiameremo **traccia** della matrice $\operatorname{tr} A$ la somma dei suoi elementi diagonali, ossia:

$$\operatorname{tr} A = a_{11} + a_{22} + \dots a_{nn} = \sum_{i=1}^{n} a_{ii}. \tag{4.10}$$

Per le prossime considerazioni, ci serve un risultato molto utile che sintetizziamo nel seguente

Lemma 4.21. *Siano $A, B \in \mathcal{M}_{\mathbb{R}}(n)$ matrici quadrate di ordine n reali. Allora la traccia del prodotto fra le matrici non dipende dall'ordine delle matrici nel prodotto, ossia*

$$\operatorname{tr}(A \cdot B) = \operatorname{tr}(B \cdot A)\,. \tag{4.11}$$

Dimostrazione. Ricordando la definizione di traccia 4.6, e la definizione del prodotto righe per colonne fra matrici, possiamo scrivere

$$\begin{aligned}
\operatorname{tr}(A \cdot B) &= \sum_{i=1}^{n} (A \cdot B)_{ii} = \sum_{i=1}^{n} (\sum_{j=1}^{n} A_{ij} B_{ji}) \\
&= \sum_{j=1}^{n} \sum_{i=1}^{n} A_{ij} B_{ji} = \sum_{j=1}^{n} (\sum_{i=1}^{n} B_{ji} A_{ij}) \\
&= \sum_{j=1}^{n} (B \cdot A)_{jj} \\
&= \operatorname{tr}(B \cdot A)\,,
\end{aligned}$$

dove abbiamo usato le proprietà della sommatoria (indice muto, Osservazione 0.19) e, a più riprese, le proprietà associativa e commutativa dell'addizione. □

Studiamo ora le proprietà comuni a due matrici simili:

Proprietà 4.22 (**Invarianti per similitudine**).
Siano $A, B \in \mathcal{M}_{\mathbb{R}}(n)$ due matrici simili, allora si ha:

(1) $\operatorname{rg}(A) = \operatorname{rg}(B)$*;*
(2) $|A| = |B|$*;*
(3) $\operatorname{tr} A = \operatorname{tr} B$*.*

Dimostrazione. Verifichiamo immediatamente le prime due proprietà.
(1) Poiché le matrici sono simili, rappresentano lo stesso operatore lineare $L\colon \mathbb{R}^n \to \mathbb{R}^n$ in due basi diverse. Ricordiamo che, se L si rappresenta con una matrice A, la dimensione dello spazio immagine $\operatorname{Im} L$ è il rango della matrice A. Quindi otteniamo:

$$\operatorname{rg}(A) = \dim(\operatorname{Im} L) = \operatorname{rg}(B),$$

da cui deduciamo che le due matrici hanno lo stesso rango.
(2) Poiché le matrici sono simili, esiste una matrice invertibile $M \in \mathcal{GL}(n, \mathbb{R})$ che verifica l'uguaglianza

$$B = M^{-1}AM.$$

Calcolando il determinante a sinistra e a destra nell'uguaglianza e applicando la regola di Binet otteniamo:

$$|B| = |M^{-1}AM| = |M^{-1}||A||M|.$$

Ricordiamo ora che il determinante di una matrice quadrata è un numero reale, per cui applicando la proprietà commutativa del prodotto si ha:

$$|B| = |A||M^{-1}||M|,$$

ricordando ora che $|M^{-1}| = |M|^{-1}$, si ottiene che $|A| = |B|$.
Per dimostrare la proprietà (3), ricordiamo il Lemma (4.21); con la matrice M come al punto precedente, possiamo scrivere

$$\operatorname{tr} B = \operatorname{tr}(M^{-1}AM) = \operatorname{tr}(AMM^{-1}) = \operatorname{tr}(AI(n)) = \operatorname{tr} A\,.$$

□

Osservazione 4.21. Le proprietà che abbiamo enunciato sono **condizioni necessarie ma non sufficienti** per garantire la similitudine tra due matrici, come mostreremo nei seguenti esempi.

Esempio 4.22.

(1) Consideriamo le seguenti matrici reali quadrate di ordine 2:
$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad B = I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$
Osserviamo che risulta:
$$|A| = |B| = 1,$$
$$\operatorname{rg}(A) = \operatorname{rg}(B) = 2,$$
$$\operatorname{tr} A = \operatorname{tr} B = 2.$$
Tuttavia le matrici non sono simili. Infatti, supponiamo per assurdo che esista una matrice invertibile $M \in \mathcal{GL}(2, \mathbb{R})$ tale che
$$A = M^{-1}BM,$$
poiché $B = I_2$, si avrebbe
$$A = M^{-1}I_2M = M^{-1}M = I_2,$$
che è falso. L'esempio ci mostra che la validità di tutte le condizioni enunciate sopra non implica la similitudine tra le due matrici.

(2) Consideriamo le seguenti matrici reali quadrate di ordine 2:
$$A = \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix} \quad B = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$
Osserviamo che risulta:
$$|A| = |B| = -1,$$
$$\operatorname{rg}(A) = \operatorname{rg}(B) = 2,$$
$$\operatorname{tr} A = 2 \neq \operatorname{tr} B = 0,$$
per cui possiamo concludere che le matrici non sono simili. L'esempio mostra che basta che non sia verificata una di queste proprietà per concludere che le matrici non sono simili.

(3) Verifichiamo che le seguenti matrici reali quadate di ordine 2 sono simili:
$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \quad B = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$
Osserviamo che risulta:
$$|A| = |B| = 1,$$
$$\operatorname{rg}(A) = \operatorname{rg}(B) = 2,$$
$$\operatorname{tr} A = \operatorname{tr} B = 0,$$
sono quindi verificate le condizioni necessarie per la similitudine. Per verificare che le matrcici sono simili proviamo che rappresentano lo

stesso operatore lineare in basi diverse. La matrice A definisce l'operatore lineare $L_A\colon \mathbb{R}^2 \to \mathbb{R}^2$:

$$L_A \begin{pmatrix} x \\ y \end{pmatrix} = A \begin{pmatrix} x \\ y \end{pmatrix}.$$

Sia $\{\boldsymbol{e}_1, \boldsymbol{e}_2\}$ la base canonica di $\mathbb{R}^2$, abbiamo:

$$L_A(\boldsymbol{e}_1) = \boldsymbol{e}_2 \quad L_A(\boldsymbol{e}_2) = -\boldsymbol{e}_1.$$

Consideriamo ora la base $\mathcal{B}' = \{\boldsymbol{u}_1, \boldsymbol{u}_2\}$ di $\mathbb{R}^2$:

$$\boldsymbol{u}_1 = \boldsymbol{e}_2 \quad \boldsymbol{u}_2 = \boldsymbol{e}_1;$$

abbiamo:

$$\begin{aligned} L_A(\boldsymbol{u}_1) = L_A(\boldsymbol{e}_2) = -\boldsymbol{e}_1 = -\boldsymbol{u}_2, \\ L_A(\boldsymbol{u}_2) = L_A(\boldsymbol{e}_1) = \boldsymbol{e}_2 = \boldsymbol{u}_1. \end{aligned}$$

Osserviamo infine che la matrice associata all'operatore L_A rispetto alla base $\mathcal{B}'$ coincide con la matrice B:

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

possiamo quindi concludere che le matrici sono simili.

APPENDICE

6. Composizione di applicazioni lineari

Come si comportano le applicazioni lineari quando vengono composte? Riportiamo le considerazioni essenziali nella seguente

OSSERVAZIONE 4.23.

(1) Siano $L_A\colon \mathbb{R}^n \to \mathbb{R}^m$ e $L_B\colon \mathbb{R}^m \to \mathbb{R}^p$ due applicazioni lineari definite rispettivamente dalle matrici $A \in \mathcal{M}_{\mathbb{R}}(m,n)$ e $B \in \mathcal{M}_{\mathbb{R}}(p,m)$. Ricordiamo (0.15) che la **composizione delle applicazioni** L_B ed L_A è l'applicazione
$$L_B \circ L_A\colon \mathbb{R}^n \to \mathbb{R}^p,$$
definita nel modo seguente:
$$\forall X \in \mathbb{R}^n : X \ \to \ AX \ \to \ B(AX) = (BA)X.$$
La composizione $L_B \circ L_A$ è quindi l'applicazione lineare L_{BA} definita dalla matrice BA.

(2) Sia $L_A\colon \mathbb{R}^n \to \mathbb{R}^n$ l'operatore lineare definito dalla matrice $A \in \mathcal{M}_{\mathbb{R}}(n)$. Ricordiamo che L_A è un isomorfismose e solo se è iniettivo; questo avviene se e solo se $\operatorname{Ker} L_A = \{\mathbf{0}_{\mathbb{R}^n}\}$, ovvero se e solo se $\operatorname{rg}(A) = n$, ovvero se e solo se la matrice A è invertibile.

Proviamo che l'operatore lineare $L_{A^{-1}}$ definito dalla matrice A^{-1} è l'operatore inverso di L_A. Infatti, dall'osservazione precedente (1) risulta:
$$L_A \circ L_{A^{-1}} = L_{AA^{-1}} = L_{I_n},$$
poiché L_{I_n} risulta essere l'operatore identità di $\mathbb{R}^n$, segue la tesi.

I risultati che abbiamo appena verificato valgono in generale per la composizione di applicazioni lineari tra spazi vettoriali di dimensione finita. Abbiamo infatti il risultato seguente che ci limitiamo ad enunciare:

PROPOSIZIONE 4.23.

(1) *Siano $L_1\colon V \to W$ e $L_2\colon W \to U$ due applicazioni lineari tra spazi vettoriali reali di dimensioni finite. Fissate le basi $\mathcal{B}_V$, $\mathcal{B}_W$ e $\mathcal{B}_U$, siano $A \in \mathcal{M}_{\mathbb{R}}(n,m)$ e $B \in \mathcal{M}_{\mathbb{R}}(m,p)$ le matrici associate rispettivamente alle applicazioni lineari L_1 ed L_2.*
La composizione $L_2 \circ L_1\colon V \to U$ è un'applicazione lineare e la matrice ad essa associata nelle basi fissate è la matrice BA.

(2) *Sia $L\colon V \to W$ un'applicazione lineare tra due spazi vettoriali reali di dimensione n. Fissate le basi $\mathcal{B}_V$ e $\mathcal{B}_W$ sia $A \in \mathcal{M}_{\mathbb{R}}(n)$ la matrice associata ad L.*
L è un isomorfismo se e solo se la matrice A è invertibile.

L'applicazione inversa di L, $L^{-1}\colon W \to V$, è un'applicazione lineare e la matrice ad essa associata nelle basi fissate è la matrice A^{-1}.

OSSERVAZIONE 4.24.

- Se $L_1\colon V \to W$ e $L_2\colon W \to U$ sono isomorfismi, allora lo loro composizione $L_2 \circ L_1$ è un isomorfismo di V in U.
- Se $L\colon V \to W$ è un isomorfismo, allora $L^{-1}\colon W \to V$ è un isomorfismo.

7. Matrici equivalenti (o congruenti)

Come è stato discusso nella sezione 5, un'applicazione lineare $L\colon V \to W$ tra due spazi vettoriali di dimensione finita può essere rappresentata da matrici differenti a seconda delle basi usate in V e W. Detta $A \in \mathcal{M}_{\mathbb{R}}(m,n)$ la matrice associata a L rispetto alle basi $\mathcal{B}_V$ e $\mathcal{B}_W$ e $B \in \mathcal{M}_{\mathbb{R}}(m,n)$ la matrice associata a L rispetto alle basi $\mathcal{B}'_V$ e $\mathcal{B}'_W$, la relazione fra le due matrici di rappresentazione è data dall'equazione (4.9), basata sulle matrici invertibili $N \in \mathcal{GL}(n,\mathbb{R})$ ed $M \in \mathcal{GL}(m,\mathbb{R})$ che realizzano il rispettivi cambiamenti di base in V e W. Questa equazione ci consente di definire una relazione di equivalenza più "debole" rispetto alla similitudine.

Sulla scorta di quanto osservato, risulta, quindi, naturale la seguente

DEFINIZIONE 4.7. Siano $A, B \in \mathcal{M}_{\mathbb{R}}(m,n)$; le due matrici si dicono *equivalenti (o congruenti)* se esistono una matrice $M \in \mathcal{GL}(m,\mathbb{R})$ ed una matrice $N \in \mathcal{GL}(m,\mathbb{R})$ tali che

$$B = M^{-1}AN\,. \tag{4.12}$$

PROPOSIZIONE 4.24. *La relazione 4.12 della definizione 4.7 è una relazione di equivalenza (la dimostrazione è lasciata per esercizio).*

Vale la pena però di effettuare la seguente

OSSERVAZIONE 4.25. Se due matrici sono simili, allora sono equivalenti: basta prendere $N = M$ nella (4.12). Non vale però il viceversa (banalmente, per essere simili le matrici devono essere quadrate, ma questo non è necessario per la relazione di equivalenza).

Stavolta, l'uguaglianza di rango fra due matrici in $\mathcal{M}_{\mathbb{R}}(m,n)$ diventa condizione necessaria e sufficiente per l'equivalenza. Per convincersi della necessità, basta considerare la seguente

PROPOSIZIONE 4.25. *Siano $A, B \in \mathcal{M}_{\mathbb{R}}(m,n)$ due matrici equivalenti. Allora,* $\operatorname{rg}(A) = \operatorname{rg}(B)$.

DIMOSTRAZIONE. Infatti, come per il caso delle matrici simili, A e B possono essere interpretate come matrici di rappresentazione della stessa applicazione lineare $L\colon V \to W$ tra due spazi vettoriali di dimensione n ed m, rispettivamente. Si ha, quindi:

$$\mathrm{rg}(A) = \dim(\mathrm{Im}\, L) = \mathrm{rg}(B)\,. \tag{4.13}$$

□

Vediamo ora che la condizione è anche sufficiente: per arrivare a questo, fa comodo mostrare un risultato utile anche in altri contesti.

Lemma 4.26. *Sia $L\colon\ V \to W$ un'applicazione lineare fra spazi vettoriali. Sia $\mathcal{S} = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \dots \boldsymbol{v}_h\}$ un insieme di h vettori linearmente* dipendenti. *Allora i vettori immagine $\mathcal{S}' = \{\boldsymbol{w}_1 = L(\boldsymbol{v}_1), \boldsymbol{w}_2 = L(\boldsymbol{v}_2), \dots \boldsymbol{w}_h = L(\boldsymbol{v}_h)\}$ sono linearmente dipendenti (in altre parole* un'applicazione lineare trasforma vettori dipendenti in vettori dipendenti*).*

Dimostrazione. Per ipotesi, esiste una combinazione lineare non banale dei vettori di $\mathcal{S}$ che produce il vettore nullo:

$$\alpha_1 \boldsymbol{v}_1 + \alpha_2 \boldsymbol{v}_2 + \dots \alpha_h \boldsymbol{v}_h = \mathbf{0}_n\,, \tag{4.14}$$

ossia, con i coefficienti α_i non tutti nulli.

Applichiamo la L ad entrambi i membri della (4.14); otteniamo, sfruttando la linearità:

$$\alpha_1 \boldsymbol{w}_1 + \alpha_2 \boldsymbol{w}_2 + \dots \alpha_h \boldsymbol{w}_h = \mathbf{0}_m\,,$$

ossia, abbiamo trovato una c.l. non banale dei vettori di $\mathcal{S}'$ che produce il vettore nullo. □

Siamo adesso in grado di mostrare il seguente

Teorema 4.27. *Siano $A, B \in \mathcal{M}_{\mathbb{R}}(m, n)$ due matrici tali che* $\mathrm{rg}(A) = \mathrm{rg}(B) = s$. *Allora le due matrici sono equivalenti.*

Dimostrazione. In analogia con quanto detto nelle sezione (5), consideriamo l'operatore

$$L_A\colon \mathbb{R}^n \to \mathbb{R}^m$$

dato da $L_A(X) = AX$. La matrice A è la matrice associata a L_A nelle basi canoniche $\mathcal{B}_n$ di $\mathbb{R}^n$ e $\mathcal{B}_m$ di $\mathbb{R}^m$. Per ipotesi, la dimensione dell'immagine dell'applicazione associata è $\dim(\mathrm{Im}\, L_A) = \mathrm{rg}(A) = s$.

Siano ora $\mathcal{B}_{Im} = \{\boldsymbol{w}_1, \boldsymbol{w}_2, \dots \boldsymbol{w}_s\}$ una base di $\mathrm{Im}(L_A)$. Essendo vettori nell'immagine, per ogni $\boldsymbol{w}_i$ esiste almeno un vettore $\boldsymbol{v}_i$ nel dominio tale che $\boldsymbol{w}_i = L_A(\boldsymbol{v}_i)$. I vettori

$$\mathcal{S} = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \dots \boldsymbol{v}_s\} \tag{4.15}$$

sono linearmente indipendenti; infatti, se così non fosse, per il lemma 4.26 i vettori di $\mathcal{B}_{Im}$ dovrebbero essere linearmente dipendenti.

Ora, se $n = m = s$ gli insiemi $\mathcal{S}$ ed $\mathcal{B}_{Im}$ sono una base per $\mathbb{R}^n$ e $\mathbb{R}^m$, e scegliendo queste basi rispettivamente nel dominio e nel codominio l'applicazione è rappresentata dalla matrice diagonale

$$I(s) = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & & \vdots & \ddots & \vdots \\ 0 & & \cdots & 0 & 1 \end{pmatrix} \tag{4.16}$$

Altrimenti, completiamo gli insiemi di vettori linearmente indipendenti $\mathcal{S}$ e/o $\mathcal{B}_{Im}$ per avere una base di $\mathbb{R}^n$ e $\mathbb{R}^m$. Siano

$$\mathcal{B}'_n = \{\boldsymbol{v}_1, \dots \boldsymbol{v}_s, \boldsymbol{v}_{s+1}, \dots \boldsymbol{v}_n\}$$

e

$$\mathcal{B}'_m = \{\boldsymbol{w}_1, \dots \boldsymbol{w}_s, \boldsymbol{w}_{s+1}, \dots \boldsymbol{w}_m\}$$

le basi così ottenute; in particolare, possiamo scegliere i vettori

$$\{\boldsymbol{v}_{s+1}, \dots \boldsymbol{v}_n\}$$

come i vettori della base del nucleo $\operatorname{Ker} L_A$ di dimensione $k = n - s$.

La matrice E che rappresenta la L_A in questa nuova base ha una struttura molto semplice: infatti, per i primi s vettori si ha

$$L(\boldsymbol{v}_i) = \boldsymbol{w}_i\,, \qquad i = 1 \dots s$$

e per gli altri

$$L(\boldsymbol{v}_i) = \boldsymbol{0}_m\,, \qquad i = s+1 \dots n$$

Risulta, pertanto, individuata una *struttura a blocchi* per C:

$$E = \left(\begin{array}{c|c} I(s) & 0_{s,k} \\ \hline 0_{m-s,s} & 0_{m-s,k} \end{array}\right) = \left(\begin{array}{ccc|ccc} 1 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & 1 & \vdots & \vdots & & \vdots \\ \vdots & \cdots & \ddots & 0 & \cdots & 0 \\ \hline 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 \end{array}\right), \tag{4.17}$$

dove il primo blocco in alto a sinistra coincide con la matrice (4.16), e gli altri blocchi sono matrici nulle 0_{pq}, con p righe e q colonne. Nel caso particolare che $s = m$ oppure $s = n$ ($k = 0$), sopravvivono solo uno dei tre blocchi, ma la struttura rimane.

Se consideriamo le matrici N ed M che realizzano i cambi di base da $\mathcal{B}_n, \mathcal{B}_m$ a $\mathcal{B}'_n, \mathcal{B}'_m$ che abbiamo descritto, possiamo scrivere

$$E = M^{-1} A N\,,$$

ossia le matrici A ed E sono equivalenti.

Ripetiamo ora il procedimento per la matrice B; questa verrà associata ad una diversa applicazione lineare L_B; seguendo però i passi di prima, poiché per ipotesi $\operatorname{rg}(A) = \operatorname{rg}(B) = s$, e, quindi, $\dim(\operatorname{Ker}(L_B)) = n - s = k$, troveremo due nuove matrici di cambio di base M' ed N', che porterebbero a rappresentare l'applicazione con la *stessa* matrice a blocchi E trovata in (4.17). Ossia, anche le matrici B e E sono equivalenti.

Ma l'equivalenza fra matrici è una relazione di equivalenza (Proposizione 4.24); le matrici A e B sono equivalenti entrambe ad E e, quindi, equivalenti fra loro per la proprietà transitiva. □

Osservazione 4.26. Si noti che, in questo contesto, la matrice (4.16) non rappresenta, in generale l'applicazione identica perché ***le basi scelte nel dominio e nel codominio non coincidono***.

CAPITOLO 5

Sistemi lineari

1. Definizioni ed esempi.

Come discusso anche nel Capitolo 2, per molti problemi, in differenti campi, ci troviamo a trattare con un numero $n \geq 1$ di incognite, legate da relazioni che si traducono in una serie di equazioni lineari nelle stesse incognite. La risoluzione del problema si traduce quindi nel trovare (tutte) le soluzioni di un sistema lineare.

Ad esempio, abbiamo visto nei capitoli precedenti che un'equazione lineare (cioè di primo grado) nelle variabili x, y, z:

$$ax + by + cz + d = 0,$$

con $A_1 = (a, b, c) \neq (0, 0, 0)$, ammette soluzioni $\forall d \in \mathbb{R}$; inoltre, fissato un sistema di riferimento cartesiano nello spazio, l'insieme delle soluzioni costituisce un piano π.
Un sistema di due equazioni lineari nelle variabili x, y, z:

$$\begin{cases} a_1x + b_1y + c_1z = d_1 \\ a_2x + b_2y + c_2z = d_2, \end{cases}$$

con $A_i = (a_i, b_i, c_i) \neq (0, 0, 0)$, $\forall i = 1, 2$, ammette sicuramente soluzioni se le due equazioni rappresentano due piani π_1 e π_2 non paralleli, cioè i vettori riga A_1 e A_2 (ossia, i vettori normali ai piani) sono linearmente indipendenti in $\mathbb{R}^3$; in tal caso l'insieme delle soluzioni è dato dai punti della retta r intersezione dei due piani. Se invece i vettori A_1 e A_2 sono linearmente dipendenti, allora: o le due equazioni rappresentano due piani paralleli, che hanno intersezione vuota e quindi il sistema lineare non ammette soluzioni, oppure le due equazioni rappresentano lo stesso piano $\pi_1 = \pi_2$, che pertanto risulta essere l'insieme delle soluzioni del sistema.

In questo capitolo ci proponiamo di trattare in modo completo la teoria dei sistemi lineari; iniziamo a formalizzare il discorso.

Definizione 5.1. Un'equazione lineare a coefficienti reali nelle incognite $x_1 \dots x_n$ è un'equazione di primo grado a coefficienti reali nelle variabili $x_1, \dots x_n$:

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b, \tag{5.1}$$

con $a_i \in \mathbb{R}$, $i = 1, \dots n$, $b \in \mathbb{R}$.

DEFINIZIONE 5.2. Un sistema lineare di $m \geq 1$ equazioni in $n \geq 1$ incognite a coefficienti reali è un insieme di m equazioni lineari nelle stesse incognite $x_1 \dots x_n$:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m, \end{cases} \tag{5.2}$$

con $a_{ij} \in \mathbb{R}$, $i = 1, \dots m$ $j = 1, \dots n$, $b_i \in \mathbb{R}$, $i = 1, \dots m$.

Osserviamo che non esiste alcuna relazione tra il numero m delle equazioni ed il numero n delle incognite!!! I sistemi lineari in cui $m = n$, cioè in cui si hanno tante equazioni quante incognite, sono particolari sistemi che sono detti sistemi *quadrati di ordine n*. Gli studenti ad esempio hanno incontrato e studiato i sistemi lineari di due equazioni in due incognite.

NOTAZIONE 5.3. Dato il sistema lineare (5.2), introduciamo:

- *il vettore colonna delle incognite* del sistema: $X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$;
- *il vettore colonna dei termini noti* del sistema: $B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} \in \mathbb{R}^m$;
- *i vettori colonna dei coefficienti* del sistema, ($\forall j = 1, 2 \dots n$, A^j è il vettore dei coefficienti dell'incognita x_j):
$$A^j = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix} \in \mathbb{R}^m \quad \forall j = 1 \dots n;$$
- *la matrice dei coefficienti* del sistema, di colonne $A^1 \dots A^n$:
$$A = (A^1 | \dots | A^n) \in \mathcal{M}_{\mathbb{R}}(m, n).$$

Usando le notazioni introdotte sopra il sistema (5.2) dato può essere scritto in *forma vettoriale*:

$$x_1 A^1 + x_2 A^2 + \cdots + x_n A^n = B, \tag{5.3}$$

e possiamo, quindi, leggerlo come *un'equazione vettoriale lineare* a coefficienti in $\mathbb{R}^m$ nelle incognite $x_1 \dots x_n$.

Infine il sistema può essere scritto anche in modo più compatto in *forma matriciale*:

$$AX = B, \tag{5.4}$$

e possiamo, quindi, leggerlo come *un'equazione matriciale lineare* nell'incognita $X \in \mathbb{R}^n$.

D'ora in poi, rappresenteremo sempre un sistema di m equazioni in n incognite con la scrittura matriciale:

$$AX = B, \quad A \in \mathcal{M}_{\mathbb{R}}(m,n), \quad B \in \mathbb{R}^m\,, \tag{5.5}$$

tenendo sempre presente le scritture alternative, utili per certi calcoli o dimostrazioni.

Definizione 5.4. Il sistema lineare (5.5) è detto:

- *omogeneo* se B è il vettore nullo di $\mathbb{R}^m$, cioè
$$b_i = 0 \ \ \forall i = 1, \dots m;$$
- *non omogeneo* in caso contrario, cioè
$$\exists\, b_i \neq 0, \ \ i \in \{1, \dots m\}.$$

Esempio 5.1. Analizziamo alla luce di quanto detto qualche sistema lineare

(1) Il sistema lineare
$$\begin{cases} 2x + 4y - z + t = 0 \\ x + y + 2t = 0 \\ 4x + 2z - 3t = 0, \end{cases}$$
è un sistema lineare omogeneo di 3 equazioni in 4 incognite. I vettori colonna dei coefficienti si ottengono nel seguente modo:

- il vettore A^1 è dato dai coefficienti di x: $A^1 = \begin{pmatrix} 2 \\ 1 \\ 4 \end{pmatrix}$,

- il vettore A^2 è dato dai coefficienti di y: $A^2 = \begin{pmatrix} 4 \\ 1 \\ 0 \end{pmatrix}$,

- il vettore A^3 è dato dai coefficienti di z: $A^3 = \begin{pmatrix} -1 \\ 0 \\ 2 \end{pmatrix}$,

- il vettore A^4 è dato dai coefficienti di t: $A^4 = \begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix}$.

La matrice dei coefficienti è, quindi, la seguente:
$$A = \begin{pmatrix} 2 & 4 & -1 & 1 \\ 1 & 1 & 0 & 2 \\ 4 & 0 & 2 & -3 \end{pmatrix},$$

e la forma matriciale del sistema è:

$$\begin{pmatrix} 2 & 4 & -1 & 1 \\ 1 & 1 & 0 & 2 \\ 4 & 0 & 2 & -3 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \mathbf{0}_4.$$

(2) Il sistema lineare

$$\begin{cases} 2x + y - z = 1 \\ x + y = 0 \\ x - 2z = 1 \\ y - z = 1, \end{cases}$$

è un sistema lineare non omogeneo di 4 equazioni in 3 incognite. La matrice dei coefficienti è la seguente:

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

il vettore colonna dei termini noti è il seguente: $B = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \end{pmatrix}$,

e la scrittura matriciale del sistema è la seguente:

$$\begin{pmatrix} 2 & 1 & -1 \\ 1 & 1 & 0 \\ 1 & 0 & -2 \\ 0 & 1 & -1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \end{pmatrix}.$$

Una soluzione di un sistema lineare di m equazioni in n incognite è una n-upla ordinata di numeri reali che sostituiti al posto delle incognite verificano tutte le equazioni del sistema. Formalizzando, abbiamo:

Definizione 5.5. Dato un sistema lineare $AX = B$, con $A \in \mathcal{M}_{\mathbb{R}}(m,n)$, $B \in \mathbb{R}^m$:

(1) una soluzione del sistema è un vettore $X_0 = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} \in \mathbb{R}^n$ che soddisfa l'uguaglianza:

$$A \cdot X_0 = B,$$

equivalentemente

$$\alpha_1 A^1 + \alpha_2 A^2 + \cdots + \alpha_n A^n = B;$$

(2) il sistema è detto *risolubile* o *compatibile* se esiste almeno una soluzione in $\mathbb{R}^n$; in caso contrario si dice non risolubile.

Osservazione 5.2.

(1) **Un sistema lineare omogeneo è sempre risolubile!!**
Sia $A \in \mathcal{M}_{\mathbb{R}}(m,n)$, consideriamo il sistema lineare omogeneo
$$AX = \mathbf{0}_m.$$
Il vettore nullo di $\mathbb{R}^n$ è soluzione del sistema, infatti risulta:
$$A\mathbf{0}_n = \mathbf{0}_m.$$
Tale soluzione è detta *soluzione banale* del sistema.
(2) **Un sistema lineare ammette la soluzione banale se e solo se è omogeneo.**
Infatti, se il vettore nullo di $\mathbb{R}^n$ è soluzione del sistema, risulta:
$$A\mathbf{0}_n = B,$$
ma sappiamo che si ha $A\mathbf{0}_n = \mathbf{0}_m$, deduciamo quindi che $B = \mathbf{0}_m$, cioè il sistema è omogeneo.

Esempio 5.3. Esempi di sistemi lineari

(1) Il sistema lineare
$$\begin{cases} 2x + y = 1 \\ 4x + 2y = 0 \end{cases}$$
è un sistema lineare non omogeneo di 2 equazioni in 2 incognite. Il sistema non è risolubile, infatti, fissato un sistema di riferimento cartesiano nel piano, le due equazioni lineari rappresentano due rette parallele, la loro intersezione è quindi l'insieme vuoto.
(2) Il sistema lineare
$$\begin{cases} x + y = 2 \\ 2x - 2y = 0 \\ x + 2y = 3 \end{cases}$$
è un sistema lineare non omogeneo di 3 equazioni in 2 incognite. Scriviamo il sistema in forma matriciale: $AX = B$, dove
$$A = \begin{pmatrix} 1 & 1 \\ 2 & -2 \\ 1 & 2 \end{pmatrix} \quad B = \begin{pmatrix} 2 \\ 0 \\ 3 \end{pmatrix}, \quad X = \begin{pmatrix} x \\ y \end{pmatrix}.$$
Osserviamo che il vettore $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ è soluzione del sistema, infatti si ha:
$$A\begin{pmatrix} 1 \\ 1 \end{pmatrix} = A^1 + A^2 = B.$$
Il sistema dato è quindi risolubile.

Problemi: I principali problemi riguardanti i sistemi lineari che ci poniamo sono i seguenti:

(1) determinare quando un sistema è risolubile;
(2) descrivere l'insieme delle soluzioni;
(3) trovare tecniche per la risoluzione.

2. Sistemi omogenei.

Iniziamo a studiare i sistemi lineari omogenei:

$$AX = \mathbf{0}_m, \quad A \in \mathcal{M}_{\mathbb{R}}(m,n).$$

Come abbiamo già osservato, un sistema lineare omogeneo è sempre risolubile: infatti ammette sempre la soluzione banale $X = \mathbf{0}_n$.
Spesso tuttavia, in alcuni problemi, siamo interessati a soluzioni diverse dal vettore nullo, cioè a soluzioni *non banali*. Ad esempio, per verificare che n vettori $\{\boldsymbol{v}_1, \dots \boldsymbol{v}_n\} \subset V$ sono linearmente indipendenti, dobbiamo verificare che *l'unica* n-upla di numeri reali $x_1 \dots x_n$ che verifica la relazione

$$x_1\boldsymbol{v}_1 + \dots + x_n\boldsymbol{v}_n = \mathbf{0}_V$$

è quella banale $x_1 = x_2 = \dots = x_n = 0$.

Ci proponiamo quindi innanzittutto di stabilire quando un sistema lineare omogeneo ammette solo la soluzione banale.

Proposizione 5.1. *Un sistema lineare omogeneo*

$$AX = \mathbf{0}_m, \quad A \in \mathcal{M}_{\mathbb{R}}(m,n),$$

ammette ***solo*** *la soluzione banale* $\iff$ *le colonne di* A *sono linearmente indipendenti in* $\mathbb{R}^m$ $\iff$ $\mathrm{rg}(A) = n$.

Dimostrazione. Ricordiamo che un vettore $X_0 \in \mathbb{R}^n$ è una soluzione del sistema se

$$AX_0 = \mathbf{0}_m, \quad X_0 = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix},$$

cioè risulta

$$\alpha_1 A^1 + \dots + \alpha_n A^n = \mathbf{0}_m$$

Possiamo allora concludere che:

– esiste una soluzione non banale $X_0 \neq \mathbf{0}_m$ se e solo se i vettori A^1, $\dots A^n$ sono linearmente dipendenti, ossia se e solo se $\mathrm{rg}(A) < n$;
– l'unica soluzione è quella banale $X_0 = \mathbf{0}_m$ se e solo se i vettori A^1, $\dots A^n$ sono linearmente indipendenti, ossia se e solo se $\mathrm{rg}(A) = n$.

□

Prima di proseguire, vale la pena di effettuare la seguente

Osservazione 5.4. Se in un sistema omogeneo si ha $n > m$ (cioè il numero di incognite è strettamente maggiore del numero di equazioni), allora necessariamente le colonne $A^1 \dots A^n \in \mathbb{R}^m$ della matrice A sono vettori linearmente dipendenti in $\mathbb{R}^m$, quindi $\operatorname{rg}(A) \leq m < n$ ed il sistema ammette soluzioni non banali.

Vogliamo ora studiare l'insieme delle soluzioni di un sistema lineare omogeneo.

Proposizione 5.2. *Dato un sistema lineare omogeneo:*

$$AX = \mathbf{0}_m, \quad A \in \mathcal{M}_{\mathbb{R}}(m, n),$$

l'insieme V delle soluzioni del sistema è un sottospazio di $\mathbb{R}^n$ e risulta:

$$\dim V = n - \operatorname{rg}(A).$$

Dimostrazione. Consideriamo l'applicazione lineare L_A definita dalla matrice A:

$$L_A \colon \mathbb{R}^n \to \mathbb{R}^m, \qquad X \mapsto AX,$$

basta osservare che l'insieme V delle soluzioni del sistema lineare coincide con il nucleo di L_A:

$$\operatorname{Ker} L_A = \{X \in \mathbb{R}^n | \ AX = \mathbf{0}_m \ \} = V.$$

Possiamo quindi concludere che V è un sottospazio di $\mathbb{R}^n$, dal Teorema delle dimensioni (4.8) risulta inoltre che $\dim V = n - \operatorname{rg}(A)$. $\square$

Abbiamo quindi dimostrato che l'insieme delle soluzioni di un sistema lineare omogeneo è un sottospazio di $\mathbb{R}^n$; proviamo ora che, viceversa, ogni sottospazio di $\mathbb{R}^n$ è definito da un sistema lineare omogeneo.

Proposizione 5.3. *Un sottospazio $V \subset \mathbb{R}^n$ di dimensione r, $0 < r < n$, è definito da un sistema lineare omogeneo di $n - r$ equazioni indipendenti nelle incognite $x_1 \dots x_n$.*

Dimostrazione. Poiché $\dim V = r$, esistono r vettori $X_1, \dots X_r$ di $\mathbb{R}^n$ linearmente indipendenti che generano V:

$$V = \operatorname{Span}(X_1, \dots X_r),$$

Identifichiamo questi vettori con le colonne $A^1 \dots A^j$ della matrice

$$A = (A^1 | A^2 | \dots | A^r) \in \mathcal{M}_{\mathbb{R}}(n, r),$$

ossia $X_i = A^i = \begin{pmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{ni} \end{pmatrix} \in \mathbb{R}^n, \quad i = 1, \dots r.$

Sia $X \in \mathbb{R}^n$, osserviamo che $X \in V$ se e solo se $X \in \operatorname{Span}(X_1, \dots X_r)$, e questo avviene se e solo se i seguenti sottospazi coincidono

$$\operatorname{Span}(X_1, \dots X_r, X) = \operatorname{Span}(X_1, \dots X_r).$$

Poiché $\mathrm{Span}(X_1, \dots X_r) \subseteq \mathrm{Span}(X_1, \dots X_r, X)$, i due sottospazi coincidono se e solo se hanno la stessa dimensione:

$$\dim(\mathrm{Span}(X_1, \dots X_r, X)) = \dim(\mathrm{Span}(X_1, \dots X_r)).$$

Poiché le colonne di A sono linearmente indipendenti, per come le abbiamo costruite, risulta $\mathrm{rg}(A) = r$.
Per ogni vettore $X \in \mathbb{R}^n$ consideriamo ora la matrice $(A|X) \in \mathcal{M}_{\mathbb{R}}(n, r+1)$: la condizione precedente si traduce nella seguente:

$$X \in V \iff \mathrm{rg}(A|X) = r.$$

Ma $\mathrm{rg}(A) = r$, quindi esiste un minore Δ non nullo di ordine r in A, supponiamo, per semplicità, che tale minore si ottenga considerando le prime r righe di A (è sempre possibile riordinare le righe del sistema per avere questa condizione, senza cambiarne l'insieme delle soluzioni):

$$\Delta = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1r} \\ a_{21} & a_{22} & \cdots & a_{2r} \\ \vdots & & & \\ a_{r1} & a_{r2} & \cdots & a_{rr} \end{vmatrix}.$$

Allora, $\mathrm{rg}(A|X) = r$ se e solo se sono nulli **tutti** i minori di ordine $r+1$ della matrice $(A|X)$. Tuttavia, per la regola di Kronecker (3.27), basta provare che sono nulli tutti i minori di ordine $r+1$ ottenuti "orlando" quelli contenenti le prime r righe. Poiché la matrice $(A|X)$ è di ordine $n \times (r+1)$, questi minori sono tanti quante le righe restanti della matrice, cioè: $n - r$.
Indichiamo con $\Delta_i(X)$ il minore di ordine $r+1$ di $(A|X)$ contenente le prime r righe e la i-esima riga, in modo che cada dopo di esse $(r < i \leq n)$:

$$\Delta_i(X) = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1r} & x_1 \\ a_{21} & a_{22} & \cdots & a_{2r} & x_2 \\ \vdots & & & & \\ a_{r1} & a_{r2} & \cdots & a_{rr} & x_r \\ a_{i1} & a_{i2} & \cdots & a_{ir} & x_i \end{vmatrix}.$$

Ciascuno di questi minori è un'equazione lineare omogenea in $x_1 \dots x_n$; infatti, sviluppando il determinante sull'ultima ($(r+1)$-esima) colonna, si ottiene che $\Delta_i(X)$ è una funzione lineare e omogenea nelle coordinate $x_1 \dots x_r$, x_i a coefficienti reali non tutti nulli: infatti il coefficiente di x_i è proprio $\Delta \neq 0$.

La condizione che tutti i minori orlati siano nulli, quindi, equivale ad imporre che $\forall i = r+1, \dots n$, l'equazione $\Delta_i(X) = 0$ sia soddisfatta.
Riassumendo, possiamo concludere che V è il sottospazio delle soluzioni del seguente sistema lineare omogeneo di $n - r$ equazioni nelle incognite $x_1 \dots x_n$:

$$\begin{cases} \Delta_{r+1}(X) = 0 \\ \vdots \\ \Delta_n(X) = 0, \end{cases}$$

ossia, in forma matriciale $DX = \mathbf{0}_n$, $D \in \mathcal{M}_{\mathbb{R}}(n - r, n)$. Infine, le $n - r$ equazioni risultano indipendenti e quindi necessarie. Infatti, poiché V è lo spazio delle soluzioni del sistema, per il teorema precedente si ha

$$r = \dim V = n - \mathrm{rg}(D),$$

da cui ricaviamo che $\mathrm{rg}(D) = n - r$. □

ESEMPIO 5.5. Vediamo qualche esempio.

(1) Sia $V \subset \mathbb{R}^4$ il sottospazio generato dai seguenti vettori:

$$\boldsymbol{u} = \begin{pmatrix} 1 \\ 2 \\ 1 \\ 1 \end{pmatrix} \quad \boldsymbol{v} = \begin{pmatrix} 1 \\ 0 \\ -1 \\ 1 \end{pmatrix}.$$

Vogliamo trovare un sistema di equazioni lineari che definiscono V.
Osserviamo che i vettori $\boldsymbol{u}$ e $\boldsymbol{v}$ sono linearmente indipendenti, quindi sono una base di V e il sottospazio ha dimensione 2. Pertanto sono necessarie $4 - 2 = 2$ equazioni lineari *indipendenti* per definire il sottospazio.
Sia $X \in \mathbb{R}^4$, $X \in V$ se e solo se $X \in \operatorname{Span}(\boldsymbol{u}, \boldsymbol{v})$, seguendo il procedimento del teorema precedente, $X \in V$ se e solo se

$$\operatorname{rg} \begin{pmatrix} 1 & 1 & x_1 \\ 2 & 0 & x_2 \\ 1 & -1 & x_3 \\ 1 & 1 & x_4 \end{pmatrix} = 2.$$

Osserviamo che il minore di ordine 2 estratto dalle prime due righe e dalle prime due colonne è non nullo:

$$\Delta = \begin{vmatrix} 1 & 1 \\ 2 & 0 \end{vmatrix} = -2 \neq 0.$$

Calcoliamo i minori di ordine 3 che si ottengono orlando la sottomatrice analizzata, ossia che contengono le prime due righe:

$$\Delta_3(X) = \begin{vmatrix} 1 & 1 & x_1 \\ 2 & 0 & x_2 \\ 1 & -1 & x_3 \end{vmatrix} = -2x_1 + 2x_2 - 2x_3,$$

$$\Delta_4(X) = \begin{vmatrix} 1 & 1 & x_1 \\ 2 & 0 & x_2 \\ 1 & 1 & x_4 \end{vmatrix} = 2x_1 - 2x_4.$$

Il sottospazio V è lo spazio delle soluzioni del seguente sistema lineare omogeneo di 2 equazioni nelle incognite x_1, x_2, x_3, x_4:

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

(2) Sia $V \subset \mathbb{R}^n$ un sottospazio generato da $n-1$ vettori linearmente indipendenti. Osserviamo che se $n = 2$, $V \subset \mathbb{R}^2$ è una retta per l'origine; se $n = 3$, $V \subset \mathbb{R}^3$ è un piano per l'origine, infine, se $n \geq 4$, V ha dimensione $n - 1$ e viene detto *iperpiano* di $\mathbb{R}^n$.
Un iperpiano $V \subset \mathbb{R}^n$ è definito da *un'unica equazione lineare omogenea* in $x_1 \dots x_n$.
Infatti, dal teorema precedente, poiché $n - (n-1) = 1$, il sistema lineare che definisce V ha un'unica equazione lineare. Per determinare l'equazione di V procediamo come nel teorema precedente. Siano

$\{A^1, \dots A^{n-1}\}$ una base di V e $X \in \mathbb{R}^n$. Allora, $X \in V$ se e solo se $X \in \mathrm{Span}(A^1, \dots A^{n-1})$, cioè $X \in V$ se e solo se si ha

$$\mathrm{rg}(A^1|\dots|A^{n-1}|X) = n-1.$$

Poiché la matrice $(A^1|\dots|A^{n-1}|X)$ è quadrata di ordine n e le colonne $A^1, \dots A^{n-1}$ sono linearmente indipendenti allora $\mathrm{rg}(A^1|\dots|A^{n-1}|X) = n-1$ se e solo se il determinante è nullo:

$$\left|(A^1|\dots|A^{n-1}|X)\right| = 0.$$

Otteniamo in questo modo l'equazione cartesiana dell'iperpiano V.

3. Sistemi non omogenei.

Dato un sistema lineare non omogeneo

$$AX = B, \quad A \in \mathcal{M}_{\mathbb{R}}(m,n),\ B \in \mathbb{R}^m,$$

per prima cosa ci proponiamo di stabilire se il sistema è risolubile.

A tale scopo introduciamo la *matrice completa* del sistema, ottenuta da A aggiungendo una colonna formata dal vettore dei termini noti:

$$\tilde{A} = (A|B) = (A^1|A^2|\dots|A^n|B) \in \mathcal{M}_{\mathbb{R}}(m, n+1)\,. \tag{5.6}$$

Un criterio semplice ed efficace per stabilire se un sistema è risolubile è dato dal fondamentale teorema sui sistemi lineari di Rouché e Capelli:

Teorema 5.4 (di Rouché-Capelli). *Un sistema lineare*

$$AX = B, \quad A \in \mathcal{M}_{\mathbb{R}}(m,n),\ B \in \mathbb{R}^m,$$

è risolubile se e solo se $\mathrm{rg}(A) = \mathrm{rg}(\tilde{A})$.

Dimostrazione. Ricordiamo che un vettore $X_0 \in \mathbb{R}^n$ è una soluzione del sistema se

$$AX_0 = B, \quad X_0 = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix},$$

cioè risulta

$$\alpha_1 A^1 + \dots + \alpha_n A^n = B.$$

Quindi il sistema è risolubile se e solo se B è combinazione lineare dei vettori colonna di A, cioè se e solo se $B \in \mathrm{Span}(A^1, A^2, \dots A^n)$.

Osserviamo che:

– se $B \in \mathrm{Span}(A^1, A^2, \dots A^n)$, allora B è un generatore superfluo per il sottospazio $\mathrm{Span}(A^1, \dots A^n, B)$ e quindi risulta

$$\mathrm{Span}(A^1, \dots A^n, B) = \mathrm{Span}(A^1, \dots A^n),$$

che implica $\mathrm{rg}(A) = \mathrm{rg}(\tilde{A})$;

– se $\mathrm{rg}(A) = \mathrm{rg}(\tilde{A})$, allora i sottospazi $\mathrm{Span}(A^1, \dots A^n, B)$ e $\mathrm{Span}(A^1, \dots A^n)$ hanno la stessa dimensione, ma poiché si ha

$$\mathrm{Span}(A^1, \dots A^n) \subseteq \mathrm{Span}(A^1, \dots A^n, B),$$

i due sottospazi coincidono ed in particolare $B \in \mathrm{Span}(A^1, A^2, \dots A^n)$.

□

Osservazione 5.6. Se $\mathrm{rg}(A) = \mathrm{rg}(\tilde{A}) = n$, allora il sistema ammette un'unica soluzione. In altre parole, se esiste, la soluzione è *unica quando il rango della matrice dei coefficienti coincide con il numero delle incognite.*

Dimostrazione. Infatti, osserviamo che se $\mathrm{rg}(A) = \mathrm{rg}(\tilde{A}) = n$, il sistema è risolubile per il Teorema di Rouché-Capelli, inoltre i vettori colonna A^1, $A^2 \dots A^n$ della matrice A sono una base per il sottospazio $\mathrm{Span}(A^1, \dots A^n)$. Allora la scrittura di B come combinazione lineare di $A^1, \dots A^n$ è unica:

$$B = \alpha_1 A^1 + \cdots + \alpha_n A^n,$$

cioè il sistema ammette un'unica soluzione. □

Esempio 5.7.

(1) Determiniamo se il sistema lineare seguente è risolubile:

$$\begin{cases} x + 2y - z = 1 \\ x - y + 2z = 0 \\ 3y - 3z = 2. \end{cases}$$

È un sistema lineare non omogeneo di 3 equazioni in 3 incognite. La matrice dei coefficienti e la colonna dei termini noti sono:

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

Calcoliamo il rango di A. Poiché A è una matrice quadrata, innanzitutto calcoliamo il suo determinante:

$$|A| = \begin{vmatrix} 1 & 2 & -1 \\ 1 & -1 & 2 \\ 0 & 3 & -3 \end{vmatrix} = \begin{vmatrix} 0 & 3 & -3 \\ 1 & -1 & 2 \\ 0 & 3 & -3 \end{vmatrix} = 0.$$

Poiché le prime due colonne di A sono linearmente indipendenti, possiamo concludere che $\mathrm{rg}(A) = 2$.

Scriviamo ora la matrice completa:

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

e calcoliamo il suo rango. Consideriamo la seguente sottomatrice quadrata di $\tilde{A}$:

$$A' = (A^1 | A^2 | B), \qquad A' \in \mathcal{M}_{\mathbb{R}}(3)\,;$$

osserviamo che

$$|A'| = \begin{vmatrix} 1 & 2 & 1 \\ 1 & -1 & 0 \\ 0 & 3 & 2 \end{vmatrix} = \begin{vmatrix} 0 & 3 & 1 \\ 1 & -1 & 0 \\ 0 & 3 & 2 \end{vmatrix} = -1(6-3) = -3 \neq 0,$$

per cui risulta $\mathrm{rg}(\tilde{A}) = 3$. Poiché $\mathrm{rg}(\tilde{A}) \neq \mathrm{rg}(A)$, per il Teorema di Rouché-Capelli, possiamo concludere che il sistema non è risolubile.

(2) Determiniamo se il sistema lineare seguente è risolubile:

$$\begin{cases} x + y = 1 \\ 2x + 3y = 2 \\ 7x + 6y = 7 \end{cases}$$

È un sistema lineare non omogeneo di 3 equazioni in 2 incognite. La matrice dei coefficienti e la colonna dei termini noti sono:

$$A = \begin{pmatrix} 1 & 1 \\ 2 & 3 \\ 7 & 6 \end{pmatrix}, \quad B = \begin{pmatrix} 1 \\ 2 \\ 7 \end{pmatrix}.$$

Osserviamo che le colonne di A sono linearmente indipendenti, quindi $\mathrm{rg}(A) = 2$. Scriviamo ora la matrice completa:

$$\tilde{A} = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 3 & 2 \\ 7 & 6 & 7 \end{pmatrix},$$

e calcoliamo il suo rango. Poiché $\tilde{A}$ è una matrice quadrata, calcoliamo il determinante:

$$|\tilde{A}| = 0,$$

poiché la matrice ha due colonne uguali. Possiamo concludere che $\mathrm{rg}(\tilde{A}) = \mathrm{rg}(A) = 2$, quindi per il Teorema di Rouché-Capelli il sistema è risolubile.

Ci proponiamo ora di descrivere l'insieme delle soluzioni di un sistema lineare **non omogeneo** risolubile in n incognite. Osserviamo che tale insieme non è un sottospazio di $\mathbb{R}^n$, poiché il vettore nullo non è soluzione del sistema. Facciamo un esempio. Consideriamo un sistema lineare di due equazioni nelle variabili x, y, z, che possono essere viste come le coordinate di un generico vettore di $\mathbb{E}^3_O$ nella base canonica:

$$\begin{cases} a_1x + b_1y + c_1z = d_1 \\ a_2x + b_2y + c_2z = d_2, \end{cases}$$

con $d_i \neq 0$ per almeno un valore dell'indice i, e $A^i = \begin{pmatrix} a_i \\ b_i \\ c_i \end{pmatrix} \neq \mathbf{0}$, $\forall i = 1, 2$. Supponiamo ora che i vettori colonna A^1 e A^2 siano linearmente indipendenti in $\mathbb{R}^3$. Fissato nello spazio $\mathbb{E}^3_O$ un sistema di riferimento cartesiano ortogonale $\mathcal{R}(O, \hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}, \hat{\boldsymbol{k}})$, l'insieme delle soluzioni è la retta r intersezione dei piani definiti dalle due equazioni lineari, ed è quindi una retta che non passa per l'origine. Osserviamo che l'insieme delle soluzioni non è un sottospazio di $\mathbb{R}^3$, tuttavia si ottiene traslando un sottospazio di $\mathbb{R}^3$.

Formalizziamo ora questo concetto.

DEFINIZIONE 5.6. Un sottoinsieme $\mathcal{V} \subset \mathbb{R}^n$ è detto **varietà lineare** o **sottospazio affine** di dimensione $h \geq 0$ se esistono un sottospazio $V \subset \mathbb{R}^n$ con $\dim V = h$ ed un vettore $X_0 \in \mathbb{R}^n$ tali che ogni vettore $X \in \mathcal{V}$ si scrive nel seguente modo:

$$X = X_0 + Y, \quad Y \in V.$$

Scriveremo questo sinteticamente come

$$\mathcal{V} = X_0 + V.$$

OSSERVAZIONE 5.8. Osserviamo che:

- Le rette e i piani sono varietà lineari di $\mathbb{R}^3$ di dimensione rispettivamente 1 e 2;
- $\mathcal{V} \subset \mathbb{R}^n$ è una varietà lineare di dimensione 0 se e solo se $\mathcal{V} = \{X_0\}$.

TEOREMA 5.5 (di struttura). *Dato un sistema lineare non omogeneo risolubile:*

$$AX = B, \quad A \in \mathcal{M}_{\mathbb{R}}(m, n), \ B \in \mathbb{R}^m,$$

l'insieme delle soluzioni $\mathcal{V}$ è una varietà lineare di dimensione $n - \operatorname{rg}(A)$, in particolare risulta:

$$\mathcal{V} = X_0 + \operatorname{Ker} L_A,$$

dove $X_0 \in \mathbb{R}^n$ è una soluzione del sistema e $L_A \colon \mathbb{R}^n \to \mathbb{R}^m$ è l'applicazione lineare definita dalla matrice A.

DIMOSTRAZIONE. Poiché il sistema è risolubile, l'insieme delle soluzioni del sistema $\mathcal{V} \subset \mathbb{R}^n$ è non vuoto: sia $X_0 \in \mathcal{V}$ una soluzione.
Sia $L_A \colon \mathbb{R}^n \to \mathbb{R}^m$ l'applicazione lineare definita dalla matrice A. Sappiamo che il nucleo $\operatorname{Ker} L_A$ è un sottospazio vettoriale di $\mathbb{R}^n$ di dimensione $n - \operatorname{rg}(A)$. Verfichiamo che $\mathcal{V}$ è un traslato del sottospazio $\operatorname{Ker} L_A$.

- Prima verifichiamo che $\forall\, Y \in \operatorname{Ker} L_A$, $X_0 + Y$ è una soluzione del sistema. Basta osservare che
$$A(X_0 + Y) = AX_0 + AY = B + \mathbf{0}_m = B,$$
da cui deduciamo che $X_0 + Y \in \mathcal{V}$;
- ora verifichiamo che $\forall X \in \mathbb{R}^n$ soluzione del sistema, si ha: $X = X_0 + Y$, con $Y \in \operatorname{Ker} L_A$. Basta verficare che il vettore $X - X_0 \in \mathbb{R}^n$ è un vettore del nucleo $\operatorname{Ker} L_A$; infatti:
$$A(X - X_0) = AX - AX_0 = B - B = \mathbf{0}_m,$$
da cui deduciamo che esiste un vettore $Y \in \operatorname{Ker} L_A$ tale che $X - X_0 = Y$.

□

OSSERVAZIONE 5.9. La varietà lineare $\mathcal{V}$ si riduce ad essere un unico vettore $\{X_0\}$ se e solo se risulta $\operatorname{Ker} L_A = \{\mathbf{0}_n\}$ se e solo se $\dim \operatorname{Ker} L_A = 0$ se e solo se $n = \operatorname{rg}(A)$.

Come immediata conseguenza otteniamo il seguente criterio:

Corollario 5.6. *Un sistema lineare non omogeneo risolubile:*

$$AX = B, \quad A \in \mathcal{M}_{\mathbb{R}}(m, n),\ B \in \mathbb{R}^m,$$

ammette un'unica soluzione se e solo se $\mathrm{rg}(A) = n$.

4. Sistemi quadrati non singolari.

Vediamo ora una classe di sistemi lineari che ammettono un'unica soluzione. Consideriamo sistemi lineari *quadrati*, cioè con lo stesso numero di equazioni ed incognite; vogliamo caratterizzare i sistemi quadrati che ammettono *una sola soluzione*; per fare questo, occorre premettere una

Definizione 5.7. sia

$$AX = B, \quad A \in \mathcal{M}_{\mathbb{R}}(n),\ B \in \mathbb{R}^n,$$

un sistema lineare di n equazioni in n incognite, supponiamo inoltre che la matrice A abbia rango massimo:

$$\mathrm{rg}(A) = n.$$

Un tale sistema è detto *sistema lineare quadrato non singolare* di ordine n, o anche sistema di Cramer di ordine n.

Sulla scorta di questo possiamo mostrare che la caratterizzazione cercata è molto semplice.

Proposizione 5.7. *Un sistema lineare quadrato non singolare*

$$AX = B, \quad A \in \mathcal{M}_{\mathbb{R}}(n),\ B \in \mathbb{R}^n,$$

ammette un'unica soluzione:

$$X_0 = A^{-1}B\,.$$

Dimostrazione. Notiamo che il sistema è sempre risolubile $\forall B \in \mathbb{R}^m$; infatti, poiché la matrice completa $\tilde{A} = (A|B) = (A|B) \in \mathcal{M}_{\mathbb{R}}(n, n+1)$, risulta necessariamente $\mathrm{rg}(\tilde{A}) = \mathrm{rg}(A)$ (il rango non può diminuire aggiungendo una colonna, e non può superare il minimo fra n e $n+1$, cioè n). Inoltre poiché $\mathrm{rg}(A) = n$, il sistema ammette un'unica soluzione.

Essendo $\mathrm{rg}(A) = n$, matrice A è invertibile, sia A^{-1} la matrice inversa di A. Osserviamo che dall'uguaglianza

$$AX = B,$$

moltiplicando a sinistra per la matrice inversa di A entrambi i membri si ha:

$$A^{-1}(AX) = A^{-1}B,$$

applicando le proprietà del prodotto di matrici otteniamo:

$$X = A^{-1}B\,.$$

□

Vogliamo ora determinare in modo esplicito la soluzione e ricavare una *regola di calcolo* per i sistemi quadrati non singolari. Consideriamo quindi un sistema lineare:

$$AX = B, \quad A \in \mathcal{M}_{\mathbb{R}}(n), \; B \in \mathbb{R}^n.$$

Poiché la soluzione è $X_0 = A^{-1}B$, per ogni $i = 1, \dots n$ la $i-$esima componente del vettore X_0 risulta essere il prodotto della i-esima riga di A^{-1} per il vettore B:

$$(X_0)_i = (A^{-1})_i B.$$

Ricordiamo che (formula di Cramer per l'inversa, Proposizione 3.19)

$$(A^{-1})_{ij} = (-1)^{i+j} \frac{1}{|A|} |A_{[j,i]}|,$$

dove $A_{[j,i]}$ è la sottomatrice quadrata di A che si ottiene eliminando la j-esima riga di A e la i-esima colona di A. Si ha allora:

$$(X_0)_i = (A^{-1})_i B = \frac{(-1)^{i+1} b_1 |A_{[1,i]}| + \cdots + (-1)^{i+n} b_n |A_{[n,i]}|}{|A|}.$$

Consideriamo ora la matrice che si ottiene da A sostituendo la i-esima colonna A^i con la colonna dei termini noti B:

$$(A^1 | \dots | A^{i-1} | B | A^{i+1} | \dots | A^n),$$

osserviamo che sviluppando il determinante di tale matrice con la regola di Laplace sulla i-esima colonna otteniamo:

$$\left|(A^1 | \dots A^{i-1} | B | A^{i+1} \dots A^n)\right| = (-1)^{i+1} b_1 |A_{[1,i]}| + \cdots + (-1)^{i+n} b_n |A_{[n,i]}|.$$

Possiamo quindi concludere che:

$$(X_0)_i = \frac{\left|(A^1 | \dots | A^{i-1} | B | A^{i+1} | \dots | A^n)\right|}{|A|}, \qquad i = 1, \dots n\,.$$

In altre parole, abbiamo appena mostrato la seguente regola

Proposizione 5.8 (Regola di Cramer).
Sia

$$AX = B$$

un sistema quadrato non singolare di ordine n; sia X_0 la sua unica soluzione. La componente i-esima $(X_0)_i$ del vettore soluzione si ottiene dividendo per il determinante di A il determinante della matrice $A_B(i) = (A^1 | \dots | A^{i-1} | B | A^{i+1} | \dots | A^n)$, ottenuta sostituendo la $i-$esima colonna della matrice A con la colonna B dei termini noti.

Esempio 5.10. Studiamo il seguente sistema lineare quadrato in 3 incognite:

$$\begin{cases} 2x + 2z = 0 \\ x - y + z = 1 \\ x + 2y = 2. \end{cases}$$

La matrice dei coefficienti del sistema e il vettore colonna dei termini noti sono i seguenti:

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

Osserviamo che risulta:

$$|A| = \begin{vmatrix} 2 & 0 & 2 \\ 1 & -1 & 1 \\ 1 & 2 & 0 \end{vmatrix} = \begin{vmatrix} 0 & 0 & 2 \\ 0 & -1 & 1 \\ 1 & 2 & 0 \end{vmatrix} = 2,$$

da cui deduciamo che $\operatorname{rg}(A) = 3$, il sistema è non singolare e quindi ammette un'unica soluzione. Calcoliamo la soluzione con il metodo che abbiamo ricavato sopra. Abbiamo:

$$x = \frac{\left|(B|A^2|A^3)\right|}{|A|} = \frac{1}{2} \begin{vmatrix} 0 & 0 & 2 \\ 1 & -1 & 1 \\ 2 & 2 & 0 \end{vmatrix} = 4,$$

$$y = \frac{\left|(A^1|B|A^2)\right|}{|A|} = \frac{1}{2} \begin{vmatrix} 2 & 0 & 2 \\ 1 & 1 & 1 \\ 1 & 2 & 0 \end{vmatrix} = -1,$$

$$z = \frac{\left|(A^1|A^2|B)\right|}{|A|} = \frac{1}{2} \begin{vmatrix} 2 & 0 & 0 \\ 1 & -1 & 1 \\ 1 & 2 & 2 \end{vmatrix} = -4,$$

la soluzione del sistema è la seguente:

$$X = \begin{pmatrix} 4 \\ -1 \\ -4 \end{pmatrix}.$$

5. Risoluzione di sistemi lineari.

Dato un sistema lineare possiamo effettuare delle operazioni *elementari* sulle equazioni senza modificare le soluzioni del sistema, ad esempio cambiare l'ordine delle equazioni.

Definizione 5.8. Due sistemi lineari nelle stesse incognite sono equivalenti se tutte le soluzioni di uno sono soluzioni dell'altro e viceversa.

Osservazione 5.11. Vi sono operazioni che possiamo compiere che permettono di ottenere sistemi sicuramente equivalenti ad uno dato. In particolare, osserviamo che:

- cambiando l'ordine delle equazioni, si ottiene un sistema equivalente a quello dato;
- aggiungendo ad un sistema un'equazione che è combinazione lineare delle stesse equazioni del sistema si ottiene un sistema equivalente a quello dato;

- sommando ad una equazione del sistema una combinazione lineare delle altre equazioni del sistema si ottiene un sistema equivalente a quello dato;
- un sistema è equivalente ad ogni altro sistema ottenuto eliminando tutte le equazioni che sono combinazione lineare delle altre (*equazioni superflue*), cioè ad un sistema in cui tutte le equazioni sono linearmente indipendenti.

Dato un sistema lineare risolubile

$$AX = B, \quad A \in \mathcal{M}_{\mathbb{R}}(m,n), \quad B \in \mathbb{R}^m,$$

è naturale chiedersi in che modo si possa stabilire se la i-esima equazione del sistema è superflua. Osserviamo che i coefficienti della i-esima equazione del sistema sono dati dalla i-esima riga della matrice completa $\tilde{A} = (A|B)$. Abbiamo quindi le seguenti

Proprietà 5.9.

(1) la i-esima equazione del sistema è combinazione lineare delle altre equazioni se e solo se la i-esima riga della matrice completa $\tilde{A}$ è combinazione lineare delle altre righe;
(2) se $\mathrm{rg}(A) = \mathrm{rg}(\tilde{A}) = r$, allora ci sono r equazioni linearmente indipendenti nel sistema e le altre sono superflue.

Illustriamo ora il **procedimento** che possiamo seguire per risolvere un sistema lineare:

$$AX = B, \quad A \in \mathcal{M}_{\mathbb{R}}(m,n), \quad B \in \mathbb{R}^m,$$

con $\mathrm{rg}(A) = \mathrm{rg}(\tilde{A}) = r$.

(1) $\mathrm{rg}(\tilde{A}) = r$: ci sono al massimo r righe in $\tilde{A}$ linearmente indipendenti, possiamo supporre, a meno di cambiare l'ordine delle equazioni del sistema, che siano le righe $\tilde{A}_1 \dots \tilde{A}_r$. Le prime r equazioni del sistema sono quindi indipendenti e le altre sono superflue: possiamo eliminare queste ultime equazioni.
(2) A questo punto, indichiamo con $\bar{A}$ la matrice ottenuta da A eliminando le righe superflue, e con $\bar{B}$ il vettore colonna ottenuto da B eliminando le componenti *corrispondenti*. Il sistema originario è, adesso, equivalente al sistema

$$\bar{A}X = \bar{B}, \quad \bar{A} \in \mathcal{M}_{\mathbb{R}}(r,n)\,, \quad \bar{B} \in \mathbb{R}^r,$$

con $\mathrm{rg}(\bar{A}) = (\tilde{\bar{A}}) = r$ (il sistema rimane risolubile, quindi i due ranghi devono coincidere).
(3) Se $r = n$, allora abbiamo un sistema quadrato non singolare, che ammette quindi un'unica soluzione, che possiamo calcolare con il metodo esposto nel paragrafo precedente 4.
(4) Se $r < n$, allora poiché $\mathrm{rg}(\bar{A}) = r$ esistono r colonne linearmente indipendenti in $\bar{A}$ e le altre sono combinazione lineare di queste. A meno di cambiare l'ordine delle incognite del sistema, possiamo supporre che

queste siano le colonne $\bar{A}^1, \dots \bar{A}^r$. Indichiamo con A' la matrice (*invertibile* per il Teorema 3.10!) formata da queste prime r colonne, e con A'' quella formata dalle rimanenti colonne. In altre parole, poniamo:

$$\bar{A} = (A'|A'') \quad A' \in \mathcal{GL}(r, \mathbb{R})\,, \quad A'' \in \mathcal{M}_{\mathbb{R}}(r, n-r)\,;$$

in corrispondenza, scriveremo per il vettore delle incognite

$$X = \begin{pmatrix} X' \\ X'' \end{pmatrix}, \quad X' \in \mathbb{R}^r, \quad X'' \in \mathbb{R}^{n-r}.$$

Possiamo allora scrivere il sistema nella forma seguente:

$$A'X' + A''X'' = \bar{B},$$

o, equivalentemente, portando a secondo membro:

$$A'X' = \bar{B} - A''X'' = B'(X''),$$

dove abbiamo evidenziato il fatto che il nuovo vettore dei termini noti B' dipende dai parametri X''. Osserviamo che per ogni vettore $X'' \in \mathbb{R}^{n-r}$, questo sistema lineare è un sistema lineare quadrato non singolare, quindi ammette un'unica soluzione: $X' = A'^{-1}B'(X'')$, che possiamo calcolare esplicitamente con il metodo esposto nel paragrafo precedente. La soluzione trovata dipende dal vettore $X'' \in \mathbb{R}^{n-r}$, ossia, le $n-r$ incognite, $x_{r+1}, \dots x_n$, sono *libere*, e le prime r, $x_1, \dots x_r$, vengono espresse in funzione (*lineare*) di questi parametri. Otteniamo in questo modo una rappresentazione *parametrica* della varietà lineare delle soluzioni.

Esempio 5.12. Studiamo il seguente sistema lineare non omogneo di 3 equazioni in 4 incognite:

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

La matrice dei coefficienti del sistema ed il vettore colonna dei termini noti sono:

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

Calcoliamo il rango di A: osserviamo che le colonne A^1 e A^2 sono linearmente indipendenti, quindi $\mathrm{rg}(A) \geq 2$. Calcoliamo i due minori di ordine 3 che contengono le due colonne A^1 e A^2:

$$\Delta_1 = \begin{vmatrix} 1 & -2 & 3 \\ 2 & 5 & 0 \\ 3 & 0 & 5 \end{vmatrix} = -45 + 45 = 0,$$

$$\Delta_2 = \begin{vmatrix} 1 & -2 & -1 \\ 2 & 5 & 1 \\ 3 & 0 & -1 \end{vmatrix} = 9 - 9 = 0,$$

per la regola di Kronecker degli orlati (3.27 nel Capitolo 3) possiamo concludere che $\text{rg}(A) = 2$.

Scriviamo ora la matrice completa:

$$\tilde{A} = (A|B) = \begin{pmatrix} 1 & -2 & 3 & -1 & 2 \\ 2 & 5 & 0 & 1 & -2 \\ 3 & 0 & 5 & -1 & 2 \end{pmatrix},$$

osserviamo che $B \in \text{Span}(A^4)$, quindi $B \in \text{Span}(A^1, A^2, A^3, A^4)$ e possiamo concludere che

$$\text{rg}(\tilde{A}) = \text{rg}(A) = 2.$$

Il sistema è quindi risolubile e l'insieme delle soluzioni è una varietà lineare $\mathcal{V} \subset \mathbb{R}^4$ di dimensione $n - \text{rg}(A) = 4 - 2 = 2$.
Per risolvere il sistema procediamo come sopra. Poiché $\text{rg}(\tilde{A}) = 2$ e le prime due righe di $\tilde{A}$ sono linearmente indipendenti, possiamo concludere che le prime due equazioni del sistema sono indipendenti e la terza è superflua: la possiamo eliminare. Abbiamo quindi:

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

Osserviamo ora che le prime due colonne sono linearmente indipendenti, possiamo risolvere quindi il sistema nelle incognite x e y e portare a secondo membro le altre incognite:

$$\begin{cases} x - 2y = t - 3z + 2 \\ 2x + 5y = -2 - t, \end{cases} \qquad A' = \begin{pmatrix} 1 & -2 \\ 2 & 5 \end{pmatrix}, \quad B' = \begin{pmatrix} t - 3z + 2 \\ -2 - t \end{pmatrix}.$$

Per ogni $(z, t) \in \mathbb{R}^2$, abbiamo un sistema lineare quadrato non singolare in 2 incognite, che ammette un'unica soluzione. Calcoliamo ora la soluzione, in funzione di z e t:

$$x = \frac{\begin{vmatrix} t - 3z + 2 & -2 \\ -2 - t & 5 \end{vmatrix}}{|A'|} = \frac{2 + t - 5z}{3},$$

$$y = \frac{\begin{vmatrix} 1 & t - 3z + 2 \\ 2 & -2 - t \end{vmatrix}}{|A'|} = \frac{2z - t - 2}{3}, \quad z, t \in \mathbb{R}.$$

Possiamo concludere che la varietà lineare delle soluzioni ha le seguenti equazioni parametriche:

$$\begin{cases} x = \frac{2}{3} + \frac{1}{3}\beta - \frac{5}{3}\alpha, \\ y = \frac{2}{3}\alpha - \frac{1}{3}\beta - \frac{2}{3}, \\ z = \alpha, \\ t = \beta, \end{cases} \qquad \alpha, \beta \in \mathbb{R}.$$

6. Il metodo di riduzione di Gauss

Il metodo di risoluzione, che abbiamo esposto nel paragrafo precedente, non è ottimale sul piano pratico, quando il sistema ha un numero elevato di equazioni. Infatti, come abbiamo visto, ci si riduce a risolvere un sistema quadrato non singolare di r equazioni:

$$A'X' = B', \qquad A' \in \mathcal{GL}(r, \mathbb{R}), \quad B' \in \mathbb{R}^r.$$

Ciò comporta il calcolo di un numero elevato di determinanti, osserviamo che per calcolare ogni determinante di ordine r si devono calcolare $\frac{r!}{2}$ determinanti di ordine 2.

Esistono metodi alternativi che consentono di ridurre notevolmente i calcoli, uno di questi è il *metodo di riduzione di Gauss*. Tale metodo consiste nel trasformare, tramite operazioni elementari sulle equazioni, un sistema lineare quadrato in un sistema lineare "triangolare superiore" equivalente ad esso.

Definizione 5.9. Un sistema lineare quadrato $AX = B$, con $A \in \mathcal{M}_{\mathbb{R}}(n)$ e $B \in \mathbb{R}^n$, è detto triangolare superiore se la matrice A dei coefficienti è una matrice triangolare superiore, cioè se tutti gli elementi al di sotto della diagonale principale di A sono nulli:

$$\forall i > j \quad A_{ij} = 0, \quad i, j = 1, \dots n.$$

Vale la pena di effettuare una semplice

Osservazione 5.13. Un sistema triangolare superiore, $AX = B$ con $A \in \mathcal{M}_{\mathbb{R}}(n)$ e $B \in \mathbb{R}^n$, è non singolare se e solo se tutti gli elementi sulla diagonale principale di A sono non nulli:

$$A_{ii} \neq 0, \quad \forall i = 1, \dots n.$$

Dimostrazione. Basta ricordare che il determinante di una matrice triangolare è il prodotto dei termini sulla diagonale (Proposizione 3.13). □

Un sistema triangolare superiore non singolare può essere risolto facilmente con un procedimento detto *algoritmo di riduzione all'indietro*. Illustriamo il procedimento direttamente su di un esempio.

Esempio 5.14. Consideriamo il sistema lineare di 3 equazioni in 3 incognite:

$$\begin{cases} x - 2y + 3z = 3 \\ 2y + 2z = 0 \\ 4z = -4. \end{cases}$$

La matrice dei coefficienti ed il vettore colonna dei termini noti sono:

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

Osserviamo che la matrice A è triangolare superiore, inoltre tutti gli elementi sulla diagonale principale di A sono non nulli. Il sistema è quidi un sistema triangolare superiore non singolare e quindi ammette un'unica soluzione. Procediamo ora al calcolo della soluzione. Partendo dall'ultima equazione abbiamo:

$$4z = -4 \implies z = -1,$$

sostituendo $z = -1$ nella seconda equazione otteniamo:

$$2y + 2(-1) = 0 \implies y = 1,$$

sostituendo infine nella prima equazione $z = -1$ e $y = 1$ otteniamo:

$$x - 2(1) + 3(-1) = 3 \implies x = 8.$$

La soluzione del sistema è quindi la seguente:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 8 \\ 1 \\ -1 \end{pmatrix}.$$

Il metodo di riduzione di Gauss consiste nel trasformare un sistema lineare quadrato in un sistema lineare triangolare superiore equivalente ad esso, attraverso operazioni elementari sulle equazioni. Poiché le operazioni elementari sulle equazioni si traducono in analoghe operazioni elementari sulle righe della matrice completa del sistema, possiamo procedere direttamente sulle righe della matrice completa. Lo scopo è quello di trasformare la matrice A in una matrice triangolare superiore: si procede per passi a partire dalla prima colonna di A. Il passo i-esimo consiste nell'annullare tutti gli elementi della colonna A^i che sono sotto la diagonale principale di A; ad ogni passo, si scambiano eventualmente fra loro delle righe, se vi sono degli elementi sulla colonna che sono già nulli, per portarli al di sotto. Se il sistema è, come ipotizzato, non singolare, ad ogni passo esiste almeno un elemento non nullo su una colonna che si posiziona sulla diagonale principale della matrice dei coefficienti del sistema equivalente ottenuto.
Illustriamo il procedimento direttamente su di un esempio.

Esempio 5.15. Consideriamo il sistema lineare di 4 equazioni in 4 incognite:

$$\begin{cases} x + 3y + z - t = 3 \\ 3x + 9y + 4z + t = 0 \\ 2x + y + 5z + 2t = 1 \\ y - z - t = 1. \end{cases}$$

La matrice dei coefficienti ed il vettore colonna dei termini noti sono:

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

Partiamo dalla prima colonna di A, infatti la prima coordinata del vettore A^1 è non nulla. In caso contrario, basta scambiare le equazioni del sistema per trovarsi in questa situazione.

Passo 1: Vogliamo annullare tutti gli elementi della prima colonna che sono sotto a_{11} attraverso operazioni elementari (conviene rappresentare direttamente la matrice completa $(A|B)$, con una barra verticale per separare la colonna dei termini noti). Quindi operiamo nel seguente modo: al posto della riga R_2 sostituiamo $R_2 - 3R_1$, al posto della riga R_3 sostituiamo $R_3 - 2R_1$, lasciamo infine invariata R_4.

$$\left(\begin{array}{cccc|c} 1 & 3 & 1 & -1 & 3 \\ 3 & 9 & 4 & 1 & 0 \\ 2 & 1 & 5 & 2 & 1 \\ 0 & 1 & -1 & -1 & 1 \end{array}\right) \Longrightarrow \left(\begin{array}{cccc|c} 1 & 3 & 1 & -1 & 3 \\ 0 & 0 & 1 & 4 & -9 \\ 0 & -5 & 3 & 4 & -5 \\ 0 & 1 & -1 & -1 & 1 \end{array}\right).$$

Osserviamo che l'elemento della seconda colonna sulla diagonale è nullo. Possiamo scambiare tra loro la seconda riga e la terza riga:

$$\left(\begin{array}{cccc|c} 1 & 3 & 1 & -1 & 3 \\ 0 & 0 & 1 & 4 & -9 \\ 0 & -5 & 3 & 4 & -5 \\ 0 & 1 & -1 & -1 & 1 \end{array}\right) \Longrightarrow \left(\begin{array}{cccc|c} 1 & 3 & 1 & -1 & 3 \\ 0 & -5 & 3 & 4 & -5 \\ 0 & 0 & 1 & 4 & -9 \\ 0 & 1 & -1 & -1 & 1 \end{array}\right).$$

Passo 2: Vogliamo annullare tutti gli elementi della seconda colonna che sono sotto a_{22} attraverso operazioni elementari. Quindi operiamo nel seguente modo: al posto della riga R_4 sostituiamo $R_4 + \frac{1}{5}R_2$, lasciamo invariate le altre.

$$\left(\begin{array}{cccc|c} 1 & 3 & 1 & -1 & 3 \\ 0 & -5 & 3 & 4 & -5 \\ 0 & 0 & 1 & 4 & -9 \\ 0 & 1 & -1 & -1 & 1 \end{array}\right) \Longrightarrow \left(\begin{array}{cccc|c} 1 & 3 & 1 & -1 & 3 \\ 0 & -5 & 3 & 4 & -5 \\ 0 & 0 & 1 & 4 & -9 \\ 0 & 0 & -\frac{2}{5} & -\frac{1}{5} & 0 \end{array}\right).$$

Passo 3: Vogliamo annullare tutti gli elementi della terza colonna che sono sotto a_{33} attraverso operazioni elementari. Quindi operiamo nel seguente modo: al posto della riga R_4 sostituiamo $R_4 + \frac{2}{5}R_3$ e lasciamo invariate le altre.

$$\left(\begin{array}{cccc|c} 1 & 3 & 1 & -1 & 3 \\ 0 & -5 & 3 & 4 & -5 \\ 0 & 0 & 1 & 4 & -9 \\ 0 & 0 & -\frac{2}{5} & -\frac{1}{5} & 0 \end{array}\right) \Longrightarrow \left(\begin{array}{cccc|c} 1 & 3 & 1 & -1 & 3 \\ 0 & -5 & 3 & 4 & -5 \\ 0 & 0 & 1 & 4 & -9 \\ 0 & 0 & 0 & \frac{7}{5} & -\frac{18}{5} \end{array}\right).$$

Il sistema dato è equivalente al seguente sistema lineare triangolare superiore:

$$\begin{cases} x + 3y + z + t = 3 \\ -5y + 3z + 4t = -5 \\ z + 4t = -9 \\ \frac{7}{5}t = -\frac{18}{5}. \end{cases}$$

Possiamo risolvere il sistema con il procedimento di risoluzione all'indietro. Partiamo dall'ultima equazione, abbiamo:

$$\frac{7}{5}t = -\frac{18}{5} \implies t = -\frac{18}{7}.$$

Sostituiamo $t = -\frac{18}{7}$ nella terza equazione, abbiamo:

$$z + 4(-\frac{18}{7}) = -9 \implies z = \frac{9}{7}.$$

Sostituiamo ora $z = \frac{9}{7}$ e $t = -\frac{18}{7}$ nella seconda equazione, abbiamo:

$$-5y + 3.\frac{9}{7} + 4(-\frac{18}{7}) = -5 \implies y = -\frac{2}{7}.$$

Infine, sostituiamo $y = -\frac{2}{7}$, $z = \frac{9}{7}$ e $t = -\frac{18}{7}$ nella prima equazione, abbiamo:

$$x + 3(-\frac{2}{7}) + \frac{9}{7} + (-\frac{18}{7}) = 3 \implies x = 0.$$

La soluzione del sistema è quindi la seguente:

$$\begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} 0 \\ -\frac{2}{7} \\ \frac{9}{7} \\ -\frac{18}{7} \end{pmatrix}.$$

Il metodo può anche essere generalizzato, per risolvere un sistema qualunque (non necessariamente quadrato). La riduzione di Gauss porta ad un cosiddetto *sistema a scala*, la cui analisi consente di ricavare rango della matrice dei coefficienti, e, corrispondentemente, dimensione della soluzione, nonché di individuare le variabili libere (parametri) e quelle dipendenti. Per una trattazione sistematica, si rimanda alla sezione dedicata posta nell'appendice di questo capitolo.

7. Discussione di sistemi lineari parametrici

Nella risoluzione di molti problemi, in geometria, meccanica o fisica, in cui ci troviamo a trattare con un numero $n \geq 1$ di incognite, i dati possono anche dipendere da uno o più parametri variabili in $\mathbb{R}$. Se le relazioni tra le incognite sono relazioni lineari, il problema si traduce quindi in un sistema lineare *parametrico* , cioè in un sistema in cui i coefficienti, ed eventualmente anche i termini noti, non sono numeri reali assegnati, ma sono funzioni di uno o più parametri reali.

Un sistema lineare parametrico, in generale, può essere o non essere risolubile a seconda dei valori che assumono i parametri. È quindi naturale chiedersi:

(1) per quali valori dei parametri il sistema è risolubile;
(2) per quali valori dei parametri il sistema ammette un'unica soluzione;
(3) per quali valori dei parametri l'insieme delle soluzioni è un sottospazio vettoriale o una varietà lineare di dimensione $h \geq 1$.

La *discussione* di un sistema lineare parametrico consiste nel dare una risposta a tutte le domande precedenti e costituisce quindi lo studio completo del sistema al variare dei parametri.

Analizziamo il caso di un sistema lineare che dipende da un solo parametro reale $h \in \mathbb{R}$:

$$AX = B, \quad A = A(h), \quad B = B(h),$$

dove A e B sono rispettivamente matrici di ordine $m \times n$ e di ordine $m \times 1$ le cui entrate sono funzioni a valori reali di h. Spesso sono funzioni algebriche, cioè polinomi $p(h) \in \mathbb{R}[h]$.
Per stabilire per quali valori del parametro il sistema è risolubile, studiamo in funzione del parametro h il rango della matrice A e della matrice completa $(A|B)$. Tale studio richiede particolare attenzione, esistono diversi procedimenti per il calcolo del rango quando la matrice dipende da uno o più parametri, si tratta di volta in volta di capire quale è il più conveniente; in generale, ove possibile, se una delle due matrici (A e $(A|B)$) è quadrata, può convenire iniziare con il calcolo del determinante di quella matrice, trovare i valori di h che annullano il determinante, ed iniziare la discussione a partire da questo discrimine. Confrontando poi il rango delle due matrici, possiamo stabilire per quali valori di h risulta $\mathrm{rg}(A) = \mathrm{rg}(A|B)$.
Una volta stabilito per quali valori del parametro il sistema ammette soluzioni, completiamo lo studio del sistema per tali valori, descrivendo l'insieme delle soluzioni $\mathcal{V}_h$: possiamo stabilire per quali h, $\mathcal{V}_h$ è un sottospazio, per quali invece è una varietà lineare, infine ne calcoliamo la dimensione i funzione del parametro h.

A questo punto possiamo risolvere il sistema scegliendo, tra i metodi che abbiamo esposto nei paragrafi precedenti, quello che riteniamo più conveniente, sempre controllando che il percorso scelto sia percorribile per tutti i valori del parametro che stiamo considerando. In caso contrario, ci ricordiamo di esaminare a parte alcuni parametri con un procedimento alternativo.

Per chiarire quanto detto sopra, esaminiamo alcuni esempi.

Esempio 5.16.

(1) Discutere e risolvere, al variare del parametro reale h, il seguente sistema lineare:

$$\begin{cases} hx + y + z = 1 \\ x + hy + z = 1 \\ x + y + hz = 1. \end{cases}$$

Osserviamo che il sistema dato è un sistema quadrato non omogeneo, con 3 equazioni e 3 incognite. La matrice dei coeffcienti e la colonna dei termini noti sono le seguenti:

$$A = \begin{pmatrix} h & 1 & 1 \\ 1 & h & 1 \\ 1 & 1 & h \end{pmatrix} \quad B = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Prima di tutto calcoliamo il rango di A al variare di h. Poiché la matrice A è quadrata di ordine 3, procediamo calcolando il suo determinante. Nel fare questo, **cerchiamo di evidenziare un fattore comune nel determinante che contenga il parametro** h: infatti, poiché la matrice A contiene il parametro in tutte le righe e le colonne, ci aspettiamo che il determinante sia un polinomio di terzo grado, per il quale la ricerca delle radici non è semplice. Abbiamo:

$$|A| = \begin{vmatrix} h & 1 & 1 \\ 1 & h & 1 \\ 1 & 1 & h \end{vmatrix} = \begin{vmatrix} h-1 & 1 & 1 \\ 1-h & h & 1 \\ 0 & 1 & h \end{vmatrix} = (h-1)\begin{vmatrix} 1 & 1 & 1 \\ -1 & h & 1 \\ 0 & 1 & h \end{vmatrix};$$

dove abbiamo sostituito al posto della colonna A^1 la colonna $A^1 - A^2$ e poi abbiamo applicato la proprietà di linearità del determinante sulla prima colonna. Ora, al posto della riga A_1 sostituiamo la riga $A_1 + A_2$ e calcoliamo infine il determinante sviluppandolo sulla prima colonna:

$$\begin{aligned} |A| &= (h-1)\begin{vmatrix} 0 & h+1 & 2 \\ -1 & h & 1 \\ 0 & 1 & h \end{vmatrix} \\ &= (h-1)\begin{vmatrix} h+1 & 2 \\ 1 & h \end{vmatrix} = (h-1)(h^2+h-2). \end{aligned}$$

Osserviamo che risulta:

$$|A| = 0 \iff (h-1)(h^2+h-2) = 0, \iff h = 1 \text{ o } h = -2.$$

Possiamo quindi affermare che:

$$\forall\, h \in \mathbb{R}, \text{ con } h \neq 1 \text{ e } h \neq -2, \text{ risulta } \operatorname{rg}(A) = 3.$$

Per tali valori del parametro, il sistema è un sistema quadrato non singolare e quindi ammette un'unica soluzione $\forall B \in \mathbb{R}^3$.

Analizziamo ora il sistema per i valori del parametro che rimangono, per i quali risulta $|A| = 0$. Consideriamo caso per caso.
Sia $h = 1$: sostituiamo tale valore nel sistema, abbiamo

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

Osserviamo che $\operatorname{Span}(A^1, A^2, A^3) = \operatorname{Span}(A^1)$, quindi si ha $\operatorname{rg}(A) = 1$, inoltre $B \in \operatorname{Span}(A^1)$. Indichiamo con $\tilde{A} = (A|B)$ la matrice completa del sistema, risulta allora:

$$\operatorname{rg}(A) = \operatorname{rg}(\tilde{A}) = 1.$$

Allora, per il teorema di Rouchè-Capelli, il sistema ammette soluzioni e l'insieme delle soluzioni è una varietà lineare di dimensione $3 - 1 = 2$ in $\mathbb{R}^3$, quindi è un piano in $\mathbb{R}^3$, che non passa per l'origine (il sistema non è omogeneo).
Sia ora $h = -2$: sostituiamo tale valore nel sistema, abbiamo

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

Osserviamo che, le colonne A^1 e A^2 sono linearmente indipendenti, quindi risulta $\operatorname{rg}(A) = 2$. Calcoliamo ora il rango della matrice completa $\tilde{A} = (A|B)$. Consideriamo il minore di ordine 3 di $\tilde{A}$ determinato dalle colonne A^1, A^2 e B:

$$\Delta = \begin{vmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & 1 \end{vmatrix} = \begin{vmatrix} -3 & 0 & 0 \\ 1 & -2 & 1 \\ 1 & 1 & 1 \end{vmatrix} = -3 \begin{vmatrix} -2 & 1 \\ 1 & 1 \end{vmatrix} = 9 \neq 0.$$

Poiché $\Delta \neq 0$, possiamo concludere che $\operatorname{rg}(\tilde{A}) = 3$.
Osserviamo che risulta $\operatorname{rg}(A) \neq \operatorname{rg}(\tilde{A})$, per il Teorema di Rouché-Capelli, possiamo concludere che il sistema **non** ammette soluzioni per $h = -2$.

Procediamo ora alla risoluzione del sistema. Distinguiamo due casi. Sia $h \in \mathbb{R}$, con $h \neq 1$ e $h \neq -2$. Abbiamo verificato sopra che il sistema ammette un'unica soluzione: $X_h = A^{-1}B$ che dipende dal parametro reale h. Possiamo determinare la soluzione con il metodo di Cramer. Abbiamo:

$$\begin{aligned} x &= \frac{\left|(B|A^2|A^3)\right|}{|A|} = \frac{\begin{vmatrix} 1 & 1 & 1 \\ 1 & h & 1 \\ 1 & 1 & h \end{vmatrix}}{(h-1)^2(h+2)} = \frac{(h-1)^2}{(h-1)^2(h+2)} \\ &= \frac{1}{(h+2)}, \\ y &= \frac{\left|(A^1|B|A^3)\right|}{|A|} = \frac{\begin{vmatrix} h & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & h \end{vmatrix}}{(h-1)^2(h+2)} = \frac{(h-1)^2}{(h-1)^2(h+2)} \\ &= \frac{1}{(h+2)}, \\ z &= \frac{\left|(A^1|A^2|B)\right|}{|A|} = \frac{\begin{vmatrix} h & 1 & 1 \\ 1 & h & 1 \\ 1 & 1 & 1 \end{vmatrix}}{(h-1)^2(h+2)} = \frac{(h-1)^2}{(h-1)^2(h+2)} \\ &= \frac{1}{(h+2)}. \end{aligned}$$

La soluzione del sistema per ogni $h \in \mathbb{R}$ con $h \neq 1$ e $h \neq -2$ è la seguente:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \frac{1}{(h+2)} \\ \frac{1}{(h+2)} \\ \frac{1}{(h+2)} \end{pmatrix}.$$

Sia $h = 1$: abbiamo verificato prima che l'insieme delle soluzioni è una varietà lineare $\mathcal{V}_1$ di dimensione 2 in $\mathbb{R}^3$, quindi un piano non passante per l'origine. Poiché $\operatorname{rg}(A) = \operatorname{rg}(\tilde{A}) = 1$, e la prima riga di $\tilde{A}$ non è nulla, possiamo allora eliminare le seconda e la terza equazione del sistema. Il sistema è equivalente alla seguente equazione:

$$x + y + z = 1,$$

che rappresenta il piano $\mathcal{V}_1$ in $\mathbb{R}^3$. Le soluzioni si possono rappresentare in funzione di due parametri reali:

$$\mathcal{V}_1 \colon \begin{cases} x = t \\ y = s \\ z = 1 - t - s, \quad t, s \in \mathbb{R}. \end{cases}$$

(2) Discutere, al variare del parametro reale h, il seguente sistema lineare:

$$\begin{cases} 3x + 3hy + (2-h)z = 0 \\ (2h+1)x + (h+2)y + hz = 1 + h. \end{cases}$$

Osserviamo che il sistema dato è un sistema di 2 equazioni e 3 incognite. La matrice dei coefficienti e la colonna dei termini noti sono le seguenti:

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

Prima di tutto calcoliamo il rango di A al variare di $h \in \mathbb{R}$. Osserviamo che $\forall h \in \mathbb{R}$ risulta $1 \leq \mathrm{rg}(A) \leq 2$. Scegliamo un minore di ordine 2 in A, ad esempio quello determinato dalle colonne A^2 e A^3:

$$\delta = \begin{vmatrix} 3h & 2-h \\ h+2 & h \end{vmatrix} = 3h^2 - (2-h)(2+h) = 4(h^2 - 1).$$

Osserviamo che risulta:

$$\delta = 0 \iff h = -1 \text{ o } h = +1.$$

Possiamo quindi affermare che:
$\forall\, h \in \mathbb{R}$, con $h \neq -1$ e $h \neq +1$, $\mathrm{rg}(A) = 2$. Consideriamo ora la matrice completa del sistema $\tilde{A} = (A|B)$; essendo una matrice di ordine 2×4, si ha $\mathrm{rg}(\tilde{A}) \leq 2$, per ogni $h \in \mathbb{R}$. Possiamo concludere che, $\forall h \in \mathbb{R}$, con $h \neq -1$ e $h \neq +1$ si ha:

$$\mathrm{rg}(\tilde{A}) = \mathrm{rg}(A) = 2.$$

Per il Teorema di Rouché-Capelli, il sistema ammette soluzioni e l'insieme delle soluzioni $\mathcal{V}_h$ è una varietà lineare di dimensione $3 - 2 = 1$ in $\mathbb{R}^3$, quindi è una retta in $\mathbb{R}^3$, che non passa per l'origine.

Esaminiamo ora, caso per caso, i valori del parametro rimasti.
Sia $h = -1$: sostituiamo tale valore nel sistema, abbiamo

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

Osserviamo che il sistema è omogeneo, quindi ammette sempre soluzioni. Poiché risulta $A_1 = -3A_2$, possiamo concludere che $\mathrm{rg}(A) = 1$. L'insieme delle soluzioni è quindi un sottospazio V_0 di $\mathbb{R}^3$ di dimensione $3 - 1 = 2$, ed è quindi un piano di $\mathbb{R}^3$ passante per l'origine.

Sia ora $h = +1$: sostituiamo tale valore nel sistema, abbiamo

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

Osserviamo che $\mathrm{Span}(A^1, A^2, A^3) = \mathrm{Span}(A^3)$, per cui $\mathrm{rg}(A) = 1$. Indichiamo con $\tilde{A} = (A|B)$ la matrice completa del sistema, poiché $B \notin \mathrm{Span}(A^3)$, risulta $\mathrm{rg}(\tilde{A}) = 2$. Quindi abbiamo:

$$\mathrm{rg}(A) \neq \mathrm{rg}(\tilde{A}).$$

Pertanto, per il teorema di Rouché-Capelli, il sistema non ammette soluzioni per $h = +1$.

(3) Discutere, al variare del parametro reale h, il seguente sistema lineare:

$$\begin{cases} x + 2y - 5z = h \\ 3x - y + hz = 0 \\ 2x + 3y - 3z = 0 \\ x - y - 2z = 0. \end{cases}$$

Osserviamo che il sistema dato è un sistema di 4 equazioni e 3 incognite. La matrice dei coefficienti e la colonna dei termini noti sono le seguenti:

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

Notiamo che la matrice completa del sistema $\tilde{A} = (A|B)$ è quadrata di ordine 4. Osserviamo che se $|\tilde{A}| \neq 0$, allora risulta $\mathrm{rg}(\tilde{A}) = 4 \neq \mathrm{rg}(A)$, poiché $\mathrm{rg}(A) \leq 3$, essendo $A \in \mathcal{M}_{\mathbb{R}}(4,3)$. Quindi una condizione *necessaria* affinché il sistema sia risolubile è che il parametro h verifichi l'equazione

$$|\tilde{A}| = 0.$$

Cominciamo allora calcolando il determinante della matrice completa, sviluppando sulla quarta colonna:

$$|\tilde{A}| = \begin{vmatrix} 1 & 2 & -1 & h \\ 3 & -1 & h & 0 \\ 2 & h & -3 & 0 \\ 3 & -1 & -2 & 0 \end{vmatrix} = -h \begin{vmatrix} 3 & -1 & h \\ 2 & h & -3 \\ 3 & -1 & -2 \end{vmatrix},$$

ora, nel determinante della matrice 3×3 precedente, sostituiamo al posto della prima riga la differenza tra la prima e la terza riga:

$$\begin{aligned} |\tilde{A}| &= -h \begin{vmatrix} 0 & 0 & h+2 \\ 2 & h & -3 \\ 3 & -1 & -2 \end{vmatrix} \\ &= -h(h+2)(-2-3h) = h(h+2)(2+3h). \end{aligned}$$

Osserviamo che risulta:

$$|\tilde{A}| = 0 \iff h = 0 \ \vee \ h = -2 \ \vee h = -\frac{2}{3}.$$

Possiamo quindi affermare che:

$$\forall h \in \mathbb{R}, \text{ con } h \neq 0, h \neq -2 \text{ e } h \neq -\frac{2}{3}, \text{ risulta } \operatorname{rg}(\tilde{A}) = 4 \neq \operatorname{rg}(A).$$

Per il Teorema di Rouché-Capelli, il sistema non ammette soluzioni.

Esaminiamo ora il sistema, caso per caso, per i valori del parametro rimasti.

Sia $h = 0$: sostituiamo tale valore nel sistema, abbiamo:

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

Il sistema è omogeneo, quindi ammette sempre soluzioni. Calcoliamo il rango di A. Osserviamo che $\operatorname{rg}(A) \geq 2$, infatti le colonne A^1 e A^2 sono linearmente indipendenti. Consideriamo ora il minore di ordine 3 dato dalle righe A_2, A_3 e A_4:

$$\Delta = \begin{vmatrix} 3 & -1 & 0 \\ 2 & 0 & -3 \\ 3 & -1 & -2 \end{vmatrix} = \begin{vmatrix} 0 & 0 & 2 \\ 2 & 0 & -3 \\ 3 & -1 & -2 \end{vmatrix} = 2 \begin{vmatrix} 2 & 0 \\ 3 & -1 \end{vmatrix} = -4,$$

dove abbiamo sostituito la prima riga con la differenza tra la prima e la terza riga, poi abbiamo sviluppato il determinante sulla prima riga. Poiché $\Delta \neq 0$, possiamo concludere che $\operatorname{rg}(A) = 3$; pertanto, tutte le colonne A^1, A^2 e A^3 sono linearmente indipendenti; in conclusione, *per* $h = 0$*, il sistema ammette solo la soluzione banale:* $x = y = z = 0$.

Sia $h = -2$: sostituiamo questo valore nel sistema, abbiamo:

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

Osserviamo che la seconda e la quarta equazione del sistema coincidono. Quindi possiamo eliminare quest'ultima e ottenere un sistema lineare equivalente a quello dato con 3 equazioni e 3 incognite. Studiamo allora quest'ultimo sistema. La matrice dei coefficienti e il vettore colonna dei termini noti sono i seguenti:

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

Poiché la matrice A' è quadrata, calcoliamo il suo determinante. Abbiamo:

$$|A'| = \begin{vmatrix} 1 & 2 & -1 \\ 3 & -1 & -2 \\ 2 & -2 & -3 \end{vmatrix} = \begin{vmatrix} 0 & 0 & -1 \\ 1 & -5 & -2 \\ -1 & -8 & -3 \end{vmatrix} = -1 \begin{vmatrix} 1 & -5 \\ -1 & -8 \end{vmatrix} = 13,$$

dove abbiamo sostituito al posto della colonna A'^1 la somma $A'^1 + A'^3$ e al posto della colonna A'^2 la combinazione lineare $A'^2 + 2A'^3$,e, infine,

abbiamo sviluppato il determinante sulla prima riga. Poiché risulta $|A'| \neq 0$, il sistema è un sistema quadrato non singolare e quindi, concludendo, *per* $h = -2$, *ammette un'unica soluzione.*
Sia infine $h = -\frac{2}{3}$, sostituiamo questo valore nel sistema, abbiamo:

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

Calcoliamo il rango di A'. Consideriamo la sottomatrice quadrata di ordine 3 formata dalle righe A_1, A_2, e A_3:

$$\Delta = \begin{vmatrix} 1 & 2 & -1 \\ 3 & -1 & -\frac{2}{3} \\ 2 & -\frac{2}{3} & -3 \end{vmatrix} = \begin{vmatrix} 1 & 2 & -1 \\ 0 & -7 & \frac{7}{3} \\ 0 & -\frac{14}{3} & -1 \end{vmatrix} = \begin{vmatrix} -7 & \frac{7}{3} \\ -\frac{14}{3} & -1 \end{vmatrix} = \frac{161}{9},$$

dove abbiamo sostituito al posto della riga A_2 la combinazione lineare $A_2 - 3A_1$ e al posto della riga A_3 la combinazione lineare $A_3 - 2A_1$, infine abbiamo sviluppato il determinante sulla prima colonna. Poiché risulta $\Delta \neq 0$, risulta $\mathrm{rg}(A) = 3$. Essendo, per ipotesi, $\mathrm{rg}(\tilde{A}) \leq 3$ $(\det(\tilde{A}) = 0)$ *per* $h = -\frac{2}{3}$, *possiamo concludere che si ha*:

$$\mathrm{rg}(A) = \mathrm{rg}(\tilde{A}) = 3.$$

Per il Teorema di Rouché-Capelli, il sistema ammette soluzioni e l'insieme delle soluzioni è una varietà lineare di $\mathbb{R}^3$ di dimensione $3-3 = 0$, cioè il sistema amette *un'unica soluzione.*

(4) Discutere, al variare dei parametro reale h, k, il seguente sistema lineare:

$$\begin{cases} (h+1)x + hy - z = k \\ x - y + (h+2)z = k+1 \\ 2x - 2y + (h+1)z = -2k \end{cases}$$

Si tratta di un sistema di 3 equazioni in 3 incognite, le cui matrici dei coefficienti e dei termini noti sono, rispettivamente:

(5.7)
$$A = \begin{pmatrix} (h+1) & h & -1 \\ 1 & -1 & h+2 \\ 2 & -2 & h+1 \end{pmatrix} \qquad B = \begin{pmatrix} k \\ k+1 \\ -2k\,. \end{pmatrix}$$

Cominciamo a calcolare il determinante di A:

$$|A| = \begin{vmatrix} (h+1) & h & -1 \\ 1 & -1 & h+2 \\ 2 & -2 & h+1 \end{vmatrix} = 2h^2 + 7h + 3\,,$$

che si annulla per $h = -\frac{1}{2}$ ed $h = -3$.

Quindi, per $h \neq -\frac{1}{2}, -3$, il rango di A è massimo, ossia $\mathrm{rg}(A) = 3$; siccome la matrice completa $(A|B)$ è di ordine 3×4, il suo rango non può superare 3, qualunque sia il valore di k; per il teorema di Rouché-Capelli, il sistema ammette soluzione, e questa è unica, poiché il rango di A coincide con il numero di incognite. Riassumendo,

$$\forall k \in \mathbb{R} \quad h \neq -\frac{1}{2}, -3 \implies \text{ il sistema ammette una sola soluzione.}$$

Vediamo ora cosa succede per i valori di h particolari; per $h = -\frac{1}{2}$ la matrice dei coefficienti diventa

$$A(h = -\frac{1}{2}) = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} & -1 \\ 1 & -1 & \frac{3}{2} \\ 2 & -2 & \frac{1}{2}\,. \end{pmatrix}$$

Essendo $A^2 = -A^1$, e poiché il minore formato scegliendo le prime due righe e le ultime due colonne ha determinante

$$\begin{vmatrix} -\frac{1}{2} & -1 \\ -1 & \frac{3}{2} \end{vmatrix} = -\frac{3}{4} - 1 = -\frac{7}{4} \neq 0\,,$$

si ha sicuramente $\operatorname{rg}(A) = 2$. Consideriamo la matrice completa $(A|B)$, e calcoliamo il determinante con le ultime tre colonne:

$$\det(A^2|A^3|B) = \begin{vmatrix} -\frac{1}{2} & -1 & k \\ -1 & \frac{3}{2} & k+1 \\ -2 & \frac{1}{2} & -2k \end{vmatrix} = \frac{33}{4}k + \frac{9}{4}\,.$$

Ora, se $k \neq -\frac{3}{11}$, il minore corrispondente non è nullo, e questo basta per concludere che $\operatorname{rg}(A|B) = 3 > \operatorname{rg}(A)$. In altre parole, per $h = -\frac{1}{2}$ e $k \neq -\frac{3}{11}$ il sistema non ammette soluzioni. Cosa succede se $k = -\frac{3}{11}$? Anzitutto, osserviamo che $\det(A^1|A^2|B) = 0$ per tutti i valori di k (due colonne uguali); che $\det(A^1|A^3|B) = -\det(A^2|A^3|B)$, quindi se il secondo determinante è nullo, lo è anche il primo. Cioè, se $k = -\frac{3}{11}$ tutti i minori di ordine 3 sono nulli, quindi $\operatorname{rg}(A|B) = 2$, ed il sistema è risolubile; inoltre, la soluzione ha dimensione $3 - \operatorname{rg}(A) = 3 - 2 = 1$, ossia è una retta (che non passa mai per l'origine: per nessun valore di k si ha $B = \mathbf{0}_3$).

Resta da esaminare il caso $h = -3$; in corrispondenza, la matrice dei coefficienti diventa

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

Stavolta, $A_3 = 2A_2$, cioè la terza riga è il doppio della seconda: il rango non è massimo; prendendo la sottomatrice estraendo le prime due righe e le prime due colonne, si ha subito un minore non nullo:

$$\begin{vmatrix} -2 & -3 \\ 1 & -1 \end{vmatrix} = 2 + 3 = 5\,,$$

perciò $\operatorname{rg}(A) = 2$.

Consideriamo la matrice completa, e calcoliamo, il determinante con le prime due colonne e l'ultima:

$$\det(A^1|A^2|B) = \begin{vmatrix} -2 & -3 & k \\ 1 & -1 & 1+k \\ 2 & -2 & -2k \end{vmatrix} = -20k - 10\,.$$

Ragionando come sopra, se $k \neq -\frac{1}{2}$ il sistema non ammette soluzione, perché $\mathrm{rg}(A|B) = 3 > \mathrm{rg}(A) = 2$. Inoltre, se $k = -\frac{1}{2}$ la matrice completa diventa

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

Si vede subito che la terza riga è ancora il doppio della seconda, quindi $\mathrm{rg}(A|B) = 2 = \mathrm{rg}(A)$, ed il sistema ammette soluzioni; la varietà lineare soluzione ha dimensione 1 (come prima, si tratta di una retta non passante per l'origine).

APPENDICE

8. Sistemi a scala e loro applicazioni

Come abbiamo accennato in precedenza alla fine della Sezione 6, possiamo generalizzare il metodo di riduzione di Gauss per risolvere i sistemi quadrati, estendendolo a sistemi rettangolari generici, e, anche, usare questo metodo generale per ottenere altri risultati, come la determinazione del rango di una matrice generica.

Per fare tutto questo, è necessario ridurre il sistema ad uno equivalente ad esso, attraverso passaggi simili a quelli visti per la riduzione di Gauss; arriveremo alla fine ad una matrice che presenta una struttura cosiddetta “a scala”, che formalizziamo qui di seguito.

Definizione 5.10. Una matrice S reale di ordine $k \times n$ viene detta **a scala** se

- in ogni riga S_i, con $i = 1 \dots r$, le entrate sono nulle fino ad una determinata posizione, nella quale si trova un elemento detto **pivot** p_i non nullo (eventualmente, il pivot può essere il primo elemento);
- passando alla riga successiva, il pivot si trova in corrispondenza di una colonna successiva rispetto alla precedente
- solo le prime r righe sono diverse dal vettore nullo, con $r \leq k$, ossia, se $r < k$, le righe dalla $(r+1)-$esima alla fine sono tutte nulle:

$$S_i = \mathbf{0}_n \quad i = r+1, \dots k,\ r \leq k.$$

Compatibilmente con la definizione, una matrice a scala può avere tutte le righe diverse dal vettore nullo, o anche tutte coincidenti con il vettore nullo. Quest'ultima affermazione si sintetizza dicendo che la matrice nulla $O_{k\times n}$ è una matrice a scala con 0 pivot. In altre parole, la struttura di una matrice a scala S con r *pivot* è la seguente

$$B = \begin{pmatrix} 0 & \dots & 0 & p_1 & * & * & * & * & * & * & * & * \\ 0 & \dots & 0 & \dots & \dots & 0 & p_2 & * & * & * & * & * \\ 0 & \dots & 0 & \dots & \dots & \dots & \dots & 0 & p_3 & * & * & * \\ 0 & \dots & \vdots & & & & & & \vdots & & * & \vdots \\ 0 & \dots & \vdots & & & & & & & 0 & p_r & \vdots \\ 0 & \dots & 0 & \dots & & & & & & & & 0 \\ 0 & \dots & \vdots & & & & & & & & & \vdots \\ 0 & \dots & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{5.8}$$

dove gli asterischi $*$ indicano un qualunque valore (anche nullo) e tutti i pivot $p_1, \dots p_r$ sono non nulli.

Esempio 5.17. Sono matrici a scala:

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

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

di cui S ha 3 pivot, e M ha 4 pivot (evidenziati in rosso).

Qual vantaggio presenta una matrice a scala per determinarne le sue caratteristiche? La determinazione del suo rango risulta molto facile. Infatti, vale la seguente

Proprietà 5.10. *Il rango di una matrice a scala S coincide con il numero r dei* pivot *della matrice a scala.*

Dimostrazione. Risulta facile convincersi che lo spazio generato dalle colonne di una matrice a scala di ordine $k \times n$ con r pivot ($r \leq k$) è generato dai seguenti r vettori colonna:

$$\boldsymbol{u}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \boldsymbol{u}_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots \boldsymbol{u}_r = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \tag{5.9}$$

costruiti in modo che $\boldsymbol{u}_j = \boldsymbol{e}_j$, con $j = 1, \dots r$ e gli $\boldsymbol{e}_j$ che coincidono con i primi r vettori della base canonica di $\mathbb{R}^k$.

Infatti, se $r = k$, l'affermazione è ovvia; altrimenti, solo le prime r righe possono essere diverse dal vettore nullo, quindi ogni vettore colonna della matrice ha nulle le entrate dalla $(r+1)-$esima alla $n-$esima; i vettori $\boldsymbol{e}_{r+1}, \dots \boldsymbol{e}_n$ sono da escludere per generare $\mathrm{Span}(A^1, \dots A^n)$.

Poiché i vettori della base canonica sono linearmente indipendenti, togliendo alcuni vettori si ottiene comunque una lista di vettori linearmente indipendenti (Osservazione 2.27).

Quindi, i vettori elencati in 5.9 sono una base di $\mathrm{Span}(A^1, \dots A^n)$, e la sua dimensione è, pertanto, r. □

Osserviamo a questo punto che le operazioni che abbiamo eseguito per una riduzione di Gauss su una matrice o un sistema quadrati possono essere eseguite anche su una matrice (o un sistema) rettangolare, e queste operazioni non cambiano il rango (o l'insieme delle soluzioni di un sistema). Vale quindi la seguente

PROPOSIZIONE 5.11. *Sia $A \in \mathcal{M}_{\mathbb{R}}(k, n)$ una matrice reale di ordine $k \times n$, e sia $B \in \mathbb{R}^k$ un vettore colonna; sia inoltre $\mathcal{S}$ il sistema di k equazioni in n incognite la cui matrice dei coefficienti è A ed il vettore di termini noti B, ossia $A\,X = B$.*

(1) *Scambiando tra loro due righe della matrice A (o due righe del sistema $\mathcal{S}$) il rango di A (o, rispettivamente, l'insieme delle soluzioni di $\mathcal{S}$) non cambia.*
(2) *Aggiungendo ad una riga di A (di $\mathcal{S}$) una combinazione lineare delle altre righe, il rango di A (l'insieme delle soluzioni di $\mathcal{S}$) non cambia.*

Quindi, per determinare il rango di una matrice A possiamo procedere attraverso una sequenza di passaggi compatibili con le operazioni consentite dalla Proposizione 5.11, ed ottenere una matrice a scala; il numero r di pivot della matrice coincide con $\operatorname{rg} A$ in base alla Proprietà 5.10. Sintetizziamo di seguito il metodo da seguire.

ALGORITMO 5.12 (Determinazione del rango mediante riduzione a scala). Partiamo da una matrice A di ordine $k \times n$. Ad ogni passo, occorre procedere in modo che fino ad una certa riga, a partire dalla prima e via via a scendere, la struttura della matrice equivalente che viene trovata sia quella di una matrice a scala. In dettaglio, operiamo così.

- Passo 1a: si parte dalla prima riga. Se il primo elemento è nullo, si controlla se in una delle righe successive il primo elemento è diverso da 0; se ciò avviene, si scambia questa riga con la prima. Altrimenti, si ripete il controllo sul secondo elemento, e così via.
  Al termine del processo, la prima riga e tutte quelle sotto di essa (ossia, tutte le altre), hanno entrate nulle fino ad una certa colonna: sulla prima riga, questa corrisponderà alla colonna j_1 del primo pivot della matrice a scala equivalente che si vuole ottenere.
- Passo 1b: a questo punto, a tutte le righe successive alla prima, si sottrae una riga proporzionale alla prima, in modo che sulla colonna j_1 del primo pivot si ottenga un'entrata nulla. Al termine, tutte le entrate della colonna j_1 sono nulle al di sotto del pivot.
- Passo 2a: si ripete il passo 1a, solo che si parte dalla seconda riga, e si cerca un elemento non nullo a partire dalla posizione $j_1 + 1$, ossia dalla colonna successiva a quella in cui si trova il primo pivot. Si termina con un pivot sulla seconda riga posizionato in corrispondenza della colonna $j_2 > j_1$.
- Passo 2b: a questo punto a tutte le righe successive alla seconda si sottrae una riga proporzionale ad essa, in modo che sulla colonna j_2 del secondo pivot si ottenga un'entrata nulla.
- Si ripetono i passi a ed b sulla terza riga, ecc., con l'accorgimento che, di volta in volta, la ricerca del pivot avviene su una colonna successiva a quella del passo precedente.
- L'algoritmo termina se non ci sono più righe su cui lavorare, o se le righe sono tutte nulle a partire da un certo indice di riga.

Al termine, la struttura della matrice equivalente ottenuta è già a scala, e quindi si determina facilmente il rango.

L'algoritmo risulta di facile applicazione se la matrice è numerica; se invece essa contiene uno o più parametri reali, non sempre il metodo risulta efficace, perchè nei vari passaggi occorre tenere traccia dei possibili valori dei parametri che rendono non valida o difficile l'operazione; per esempio, alcuni valori dei parametri potranno fare annullare un'entrata, e, quindi, questa non potrà corrispondere ad un pivot in quel caso. Analoghe considerazioni valgono per i sistemi parametrici.

Per questo, lo studio del rango di una matrice o di un sistema parametrici si conduce più agevolmente come indicato negli esempi riportati nella Sezione 7.

Con modalità analoga, ridurremo il sistema $AX = B$ ad un sistema $SX = B'$ a scala, ossia un sistema in cui la matrice dei coefficienti S è a scala, equivalente a quello di partenza.

Fatto questo, le colonne dei pivot individuano le r variabili *dipendenti*, mentre le rimanenti $n-r$ quelle *libere*, ossia quelle che costituiranno parametri liberi nell'espressione generale della soluzione.

Il sistema viene riscritto eliminando le righe nulle, e spostando i termini in modo che quelli con le variabili libere siano collocati nella colonna dei termini noti, e si otterrà così un sistema quadrato, anzi *triangolare* per le r variabili dipendenti, che è determinato univocamente perché sulla diagonale della matrice triangolare dei coefficienti si vengono a trovare gli r pivot della matrice, per costruzione non nulli.

Risolvendo all'indietro, si otterrà l'espressione generale della soluzione del sistema in esame.

Esempio 5.18. Vediamo alcuni esempi che ci consentono di avere più chiaramente la visione della struttura con la quale opereremo.

(1) Risolviamo il sistema

$$\begin{cases} x + y - z = 1 \\ x - y - 3z = 5 \end{cases}$$

(Osserviamo che il sistema individua l'intersezione di due piani in $\mathbb{E}^3_O$, non paralleli; la sua soluzione sarà pertanto una retta in $\mathbb{E}^3_O$).

La matrice dei coefficienti del sistema è:

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

come abbiamo fatto per la riduzione di Gauss di sistemi quadrati, possiamo procedere sinteticamente operando sulla matrice completa del sistema; questa è:

$$\left(\begin{array}{ccc|c} 1 & 1 & -1 & 1 \\ 1 & -1 & -3 & 5 \end{array}\right)$$

La prima riga è a posto: il suo primo elemento sarà il primo pivot; possiamo sottrarre alla seconda riga la prima, ed otteniamo

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

A questo punto, le prime due colonne contengono i pivot (in rosso), e, quindi, la terza colonna (in blu) individua la variabile libera, z. Il sistema triangolare che otteniamo è

$$\begin{cases} x + y = z + 1 \\ -2y = 4 + 2z \end{cases} \qquad z \in \mathbb{R}$$

che si risolve facilmente all'indietro.

$$\begin{cases} y = -z - 2 \\ x - z - 2 = z + 1 \end{cases} \implies \begin{cases} x = 2z + 3 \\ y = -z - 2 \end{cases} \qquad z \in \mathbb{R}\,.$$

La soluzione del sistema può essere scritta in una forma particolarmente efficace per individuarne esplicitamente la struttura; se esplicitiamo con un'identità il fatto che z è libera

$$\begin{cases} x = 2z + 3 \\ y = -z - 2 \\ z = z \end{cases}$$

risulta immediato scrivere la soluzione in forma compatta vettoriale:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = z \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix} + \begin{pmatrix} 3 \\ -2 \\ 0 \end{pmatrix} \qquad z \in \mathbb{R}\,.$$

Confrontando questa scrittura con la struttura generale della soluzione (5.5), identifichiamo subito la soluzione particolare del sistema

$$X_0 = \begin{pmatrix} 3 \\ -2 \\ 0 \end{pmatrix}$$

ed il nucleo della applicazione associata alla matrice A

$$\operatorname{Ker} L_A = \left\{ \boldsymbol{v} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \,|\, \boldsymbol{v} = \alpha \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix} , \alpha \in \mathbb{R} \right\}.$$

In altre parole,

$$\operatorname{Ker} L_A = \operatorname{Span}\left(\begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix} \right),$$

o, anche: una base del nucleo è data dai vettori che sono moltiplicati per le variabili libere. In questo caso,

$$\mathcal{B}_{\operatorname{Ker}_{L_A}} = \left\{ \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix} \right\}.$$

(2) Risolviamo il sistema

$$\begin{cases} x + y - 2z + 2t + w = 1 \\ 2x + y + w = 2 \\ x + 2y - 6z + w = 1 \end{cases}$$

La matrice dei coefficienti del sistema è:

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

Come prima, procediamo direttamente con la matrice completa:

$$\left(\begin{array}{ccccc|c} 1 & 1 & -2 & 2 & 1 & 1 \\ 2 & 1 & 0 & 0 & 1 & 2 \\ 1 & 2 & -6 & 0 & 1 & 1 \end{array}\right) ;$$

la prima riga ha un termine non nullo in prima posizione: possiamo tenerla, e sottrarre alla seconda la prima moltiplicata per il fattore 2, ed alla terza riga sottriaiamo la prima:

$$\left(\begin{array}{ccccc|c} 1 & 1 & -2 & 2 & 1 & 1 \\ 0 & -1 & 4 & -4 & -1 & 0 \\ 0 & 1 & -4 & -2 & 0 & 0 \end{array}\right) ;$$

a questo punto la seocnda riga presenta un pivot in seconda posizione. Dobbiamo arrivare cancellare i termini su quella colonna nelle righe dopo, ossia nella terza; basta sommare alla terza riga la seconda:

$$\left(\begin{array}{ccccc|c} \color{red}{1} & 1 & \color{blue}{-2} & 2 & \color{blue}{1} & 1 \\ 0 & \color{red}{-1} & \color{blue}{4} & -4 & \color{blue}{-1} & 0 \\ 0 & 0 & \color{blue}{0} & \color{red}{-6} & \color{blue}{-1} & 0 \end{array}\right)$$

(come sopra, i pivot sono evidenziati in rosso, le colonne delle variabili libere in blu). Otteniamo quindi il sistema triangolare con i parametri z, w liberi:

$$\begin{array}{rrl} x & +y + 2t = & 1 + 2z \quad -w \\ & -y - 4t = & \quad\; - 4z \quad +w \\ & - 6t = & \qquad\qquad\; +w \end{array} \qquad z, w \in \mathbb{R} ,$$

che risolto, dà:

$$\begin{pmatrix} x \\ y \\ z \\ t \\ w \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + z \begin{pmatrix} -2 \\ 4 \\ 1 \\ 0 \\ 0 \end{pmatrix} + w \begin{pmatrix} -1/3 \\ -1/3 \\ 0 \\ -1/6 \\ 1 \end{pmatrix} \qquad z, w \in \mathbb{R} .$$

Pertanto, il nucleo della matrice A risulta essere:

$$\operatorname{Ker} L_A = \operatorname{Span}\left(\begin{pmatrix} -2 \\ 4 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -1/3 \\ -1/3 \\ 0 \\ -1/6 \\ 1 \end{pmatrix}\right).$$

o, anche, equivalentemente, la base del nucleo è:

$$\mathcal{B}_{\operatorname{Ker}_{L_A}} = \left\{\begin{pmatrix} 2 \\ -4 \\ -1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \\ 0 \\ 1 \\ -6 \end{pmatrix}\right\}.$$

CAPITOLO 6

Autovalori e diagonalizzazione

1. Introduzione ed esempi

Nel precedente Capitolo 4 abbiamo visto cosa si intende per applicazione lineare, e ne abbiamo studiato diversi esempi; in particolare, risultano interessanti gli operatori lineari di uno spazio vettoriale in sé, e ora ci focalizzeremo quindi su questo caso.

In generale, un operatore trasforma un vettore in un vettore; risulta utile chiedersi se e quando il vettore trasformato **non cambia direzione**. Per quanto astratta questa idea possa apparire, in realtà in molti problemi saper affrontare questa questione è utile dal punto di vista teorico e/o pratico.

Per esempio, nella meccanica dei continui, si può descrivere lo stato di stress (sforzo) cui è sottoposto un corpo mediante un'applicazione lineare che, fissata una porzione di superficie infinitesima, dà lo sforzo interno $\boldsymbol{f}$ (vettore!) agente sulla superficie come il risultato di un operatore lineare T che agisce sul versore $\hat{\boldsymbol{n}}$, perpendicolare alla superficie stessa:

$$\boldsymbol{f} = T(\hat{\boldsymbol{n}})\,;$$

tranne situazioni particolari, $\boldsymbol{f}$ ed $\hat{\boldsymbol{n}}$ *non* hanno la stessa direzione. Si riesce a mostrare, però, che lo sforzo è massimo per quegli orientamenti della superficie in corrispondenza dei quali $\boldsymbol{f} \in \operatorname{Span}(\hat{\boldsymbol{n}})$, cioè, esiste $\lambda \in \mathbb{R}$ tale che $\boldsymbol{f} = \lambda\hat{\boldsymbol{n}}$. Risulta, pertanto importante stabilire se esistono e quali siano le direzioni $\hat{\boldsymbol{n}}$ che soddisfano tale condizione.

Anche senza scomodare esempi così articolati, possiamo porci il problema dal punto di vista geometrico. Vediamo qualche esempio.

Esempio 6.1. Prendiamo l'operatore lineare L_1 di $\mathbb{E}^2_O$ corrispondente ad una rotazione di un angolo $\pi/4$ in senso antiorario nel piano; l'operatore può essere definito dando l'immagine dei vettori della base canonica:

$$L_1(\hat{\boldsymbol{\imath}}) = \frac{\sqrt{2}}{2}(\hat{\boldsymbol{\imath}} + \hat{\boldsymbol{\jmath}}) \tag{6.1a}$$

$$L_1(\hat{\boldsymbol{\jmath}}) = \frac{\sqrt{2}}{2}(-\hat{\boldsymbol{\imath}} + \hat{\boldsymbol{\jmath}}) \tag{6.1b}$$

Ci chiediamo se vi sono dei vettori che non cambiano direzione. Per fare questo, passiamo alla rappresentazione stessa dei vettori sulla base canonica $\mathcal{B}_0$; un vettore $\boldsymbol{u}$ di $\mathbb{E}^2_O$ sarà rappresentato da

$$[\boldsymbol{u}]_{\mathcal{B}_0} = \begin{pmatrix} x \\ y \end{pmatrix}\,,$$

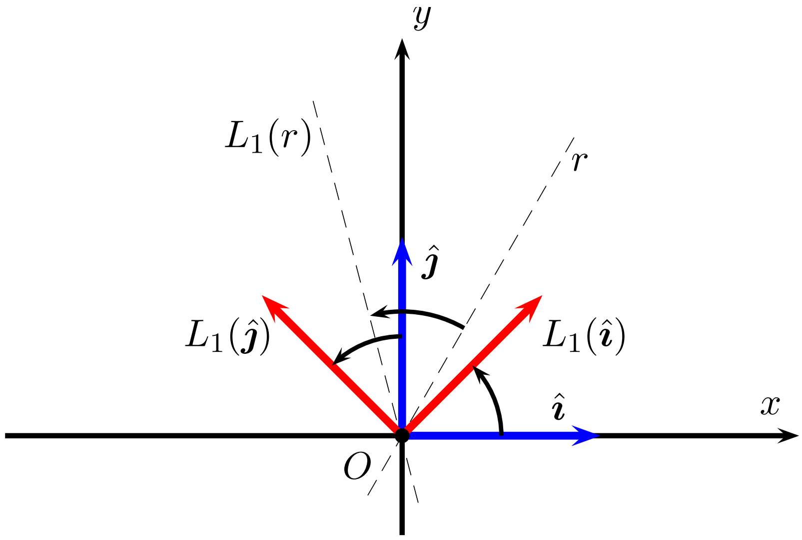

FIGURA 6.1. I vettori $\hat{\boldsymbol{\imath}}$ e $\hat{\boldsymbol{\jmath}}$ vengono ruotati di un angolo $\pi/4$ in senso antiorario dall'operatore lineare L_1; ongi retta r passante per l'origine viene ruotata di $\pi/4$, e nessuna direzione resta invariata. Solo l'origine O rimane invariata (come per ogni applicazione lineare, $L_1(\mathbf{0}) = \mathbf{0}$).

e le equazioni per le coordinate dell'immagine di un generico vettore saranno

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \frac{\sqrt{2}}{2} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}, \tag{6.2}$$

avendo esplicitato la rappresentazione di L_1 usando le equazioni 6.1. Ricordiamo che cerchiamo vettori che non cambiano direzione, cioè vettori immagine che sono nel sottospazio generato dal vettore di partenza. Quindi vogliamo sapere se, per qualche valore di $\lambda \in \mathbb{R}$ si ha:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \lambda \begin{pmatrix} x \\ y \end{pmatrix} \qquad \lambda \in \mathbb{R},$$

vale a dire se riusciamo a trovare qualche soluzione del sistema

$$\begin{cases} \frac{\sqrt{2}}{2}(x-y) & = \lambda x \\ \frac{\sqrt{2}}{2}(x+y) & = \lambda y \end{cases} \tag{6.3}$$

ottenuto esplicitando la (6.2).

Il sistema (6.3) ha una soluzione banale: $x = y = 0$ risolve il sistema per qualunque λ, ma questo è sempre vero: per ogni operatore lineare L in $\mathbb{R}^n$, si ha $L(\mathbf{0}) = \mathbf{0} = \lambda\mathbf{0}$. Nel caso sopra, però, se $\lambda = 0$, allora **solo** $x = y = 0$ risolve il sistema. Ci sono altre soluzioni quando $\lambda \neq 0$? Sommando e sottraendo membro a membro otteniamo:

$$\begin{cases} \sqrt{2}x & = \lambda(x+y) \\ \sqrt{2}y & = \lambda(y-x) \end{cases} \tag{6.4}$$

ossia, riusando le (6.3)

$$\begin{cases} x &= \lambda^2 y \\ y &= -\lambda^2 x \,. \end{cases} \tag{6.5}$$

Quindi, ricavando y dalla seconda e sostituendolo nella prima

$$x = -\lambda^4 x \,,$$

che ci dice $x = 0$, poiché $\lambda \neq 0$ (se fosse $x \neq 0$, semplificando x si nota che servirebbe un valore di λ tale che $1 = -\lambda^4$, ma non c'è perché $\lambda^4 \geq 0$...), e, quindi, $y = 0$. Esiste solo la soluzione banale, che non ci dà molta informazione (vale per ogni operatore lineare....).

Vediamo ora un altro esempio, in cui riusciremo invece a risolvere il nostro problema.

Esempio 6.2. Consideriamo l'operatore L_2 di $\mathbb{E}^2_O$ corrispondente ad una simmetria rispetto alla bisettrice del primo e terzo quadrante; esso risulta definito dalle immagini dei vettrori della base canonica:

$$L_2(\hat{\boldsymbol{\imath}}) = \hat{\boldsymbol{\jmath}} \tag{6.6a}$$

$$L_2(\hat{\boldsymbol{\jmath}}) = \hat{\boldsymbol{\imath}} \tag{6.6b}$$

Passando, come sopra alle rappresentazioni dei vettori di $\mathbb{E}^2_O$, ed impostando il sistema corripondente come nell'esempio precedente, otteniamo:

$$\begin{cases} y &= \lambda x \\ x &= \lambda y \end{cases} ; \tag{6.7}$$

sostituendo y dalla prima nella seconda equazione, otteniamo

$$x = \lambda^2 x \,.$$

Ora, questa presenta la solita soluzione banale $x = y = 0$ per ogni valore di λ. Ma è anche soddisfatta per ogni x se $\lambda^2 = 1$, ossia se $\lambda = \pm 1$.

Per $\lambda = 1$, sostituendo nella (6.7) otteniamo:

$$\begin{cases} y &= x \\ x &= y \end{cases} , \tag{6.8}$$

che è risolta da ogni vettore del tipo

$$\begin{pmatrix} x \\ y \end{pmatrix} = \alpha \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

con qualunque valore di $\alpha \in \mathbb{R}$.

Analogamente, per $\lambda = -1$, sostituendo nella (6.7) otteniamo:

$$\begin{cases} y &= -x \\ x &= -y \end{cases} , \tag{6.9}$$

che è risolta da ogni vettore del tipo

$$\begin{pmatrix} x \\ y \end{pmatrix} = \beta \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

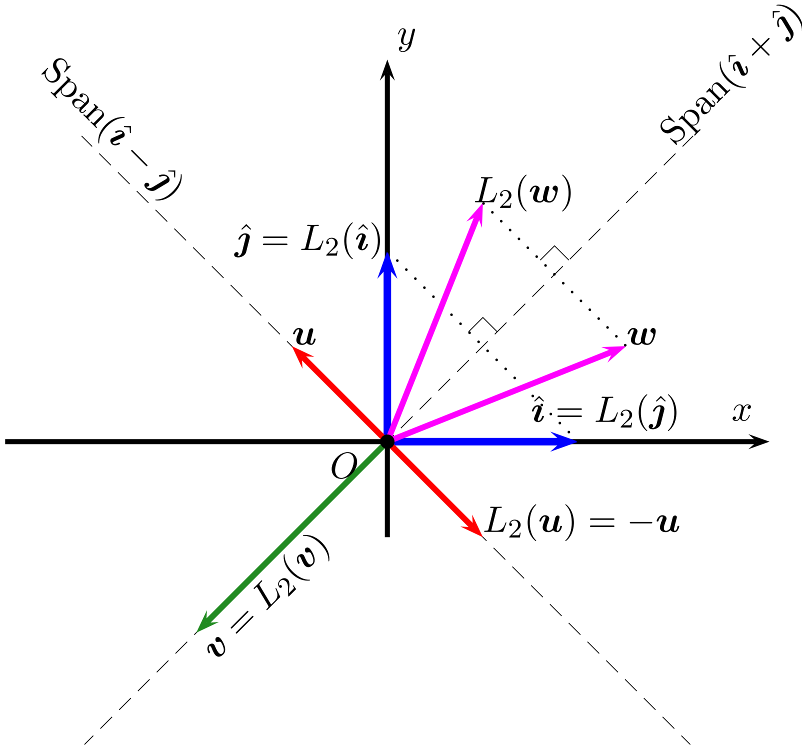

Figura 6.2. I vettori $\hat{\boldsymbol{\imath}}$ e $\hat{\boldsymbol{\jmath}}$ vengono trasformati nei loro simmetrici rispetto alla bisettrice del primo e terzo quadrande mediante l'operatore L_2. Le rette $y = \pm x$ rimangono invariate; un vettore $\boldsymbol{v} \in \mathrm{Span}(\hat{\boldsymbol{\imath}} + \hat{\boldsymbol{\jmath}})$ rimane inalterato: $L_2(\boldsymbol{v}) = \boldsymbol{v}$. Un vettore $\boldsymbol{u} \in \mathrm{Span}(\hat{\boldsymbol{\imath}} - \hat{\boldsymbol{\jmath}})$ viene trasformato nel suo opposto: $L_2(\boldsymbol{u}) = -\boldsymbol{u} \in \mathrm{Span}(\hat{\boldsymbol{\imath}} - \hat{\boldsymbol{\jmath}})$.

con qualunque valore di $\beta \in \mathbb{R}$. Vi sono quindi due rette nel piano $\mathbb{E}^2_O$ ($y = \pm x$) che corrispondono a direzioni lasciate invariate da L.

Esempio 6.3. Consideriamo ora un'applicazione lineare $L_3 \colon \mathbb{E}^2_O \to \mathbb{E}^2_O$ definita sulla base canonica dalla matrice

$$A_3 = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}, \tag{6.10}$$

ossia, per ogni vettore $\boldsymbol{v} = \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{E}^2_O$

$$L_3(\boldsymbol{v}) = A_3 \begin{pmatrix} x \\ y \end{pmatrix} = x \begin{pmatrix} 1 \\ 2 \end{pmatrix} + y \begin{pmatrix} 2 \\ 1 \end{pmatrix}. \tag{6.11}$$

Si verifica immediatamente che, anche in questo caso i vettori

$$\boldsymbol{u}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{e} \quad \boldsymbol{u}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

hanno l'immagine nel sottospazio generato da ciascuno, rispettivamente. Infatti (Fig. 6.3):

$$L_3(\boldsymbol{u}_1) = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \end{pmatrix} = 3 \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 3\boldsymbol{u}_1$$

e

$$L_3(\boldsymbol{u}_2) = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix} = -1 \begin{pmatrix} 1 \\ -1 \end{pmatrix} = -\boldsymbol{u}_2.$$

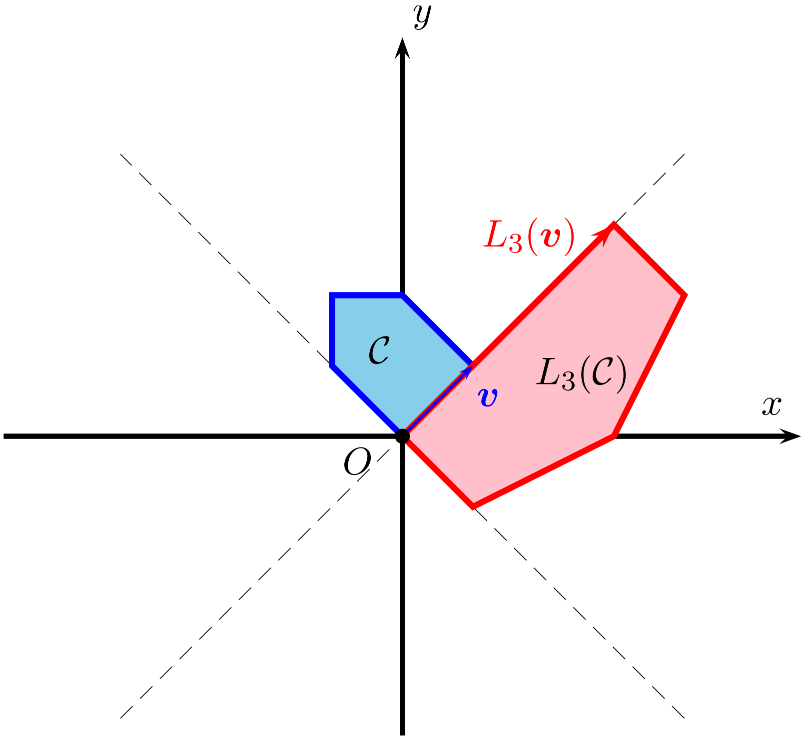

FIGURA 6.3. L'applicazione lineare $L_3 \colon \mathbb{E}^2_O \to \mathbb{E}^2_O$ definita dalla matrice dell'Eq. (6.10) lascia invariate le direzioni delle bisettrici dei quadranti $y = \pm x$. Un insieme di vettori corrispondenti ai punti della figura $\mathcal{C}$ viene trasformato dall'applicazione nella figura $L_3(\mathcal{C})$.

La Figura 6.3 suggerisce che se usiamo un diverso sistema di riferimento, in cui i nuovi assi coincidano con le bisettrici dei quadranti, l'espressione dell'applicazione potrebbe essere molto più semplice. In effetti, in questi nuovi assi, le direzioni dei versori associati agli assi restano invariate. Il cambio di base è semplice, e richiede che

$$\hat{\boldsymbol{\imath}}' = \frac{\sqrt{2}}{2}(\hat{\boldsymbol{\imath}} + \hat{\boldsymbol{\jmath}}) \tag{6.12a}$$

$$\hat{\boldsymbol{\jmath}}' = \frac{\sqrt{2}}{2}(-\hat{\boldsymbol{\imath}} + \hat{\boldsymbol{\jmath}})\,. \tag{6.12b}$$

Se scriviamo esplicitamente come agisce L_3 sulla nuova base, otteniamo

$$L_3(\hat{\boldsymbol{\imath}}') = L_3(\frac{\sqrt{2}}{2}(\hat{\boldsymbol{\imath}} + \hat{\boldsymbol{\jmath}})) = A_3\frac{\sqrt{2}}{2}\begin{pmatrix}1\\1\end{pmatrix} = \frac{\sqrt{2}}{2}\,3\begin{pmatrix}1\\1\end{pmatrix} = 3\hat{\boldsymbol{\imath}}' \tag{6.13a}$$

$$L_3(\hat{\boldsymbol{\jmath}}') = L_3(\frac{\sqrt{2}}{2}(-\hat{\boldsymbol{\imath}} + \hat{\boldsymbol{\jmath}})) = A_3\frac{\sqrt{2}}{2}\begin{pmatrix}-1\\1\end{pmatrix} = \frac{\sqrt{2}}{2}(-1)\begin{pmatrix}-1\\1\end{pmatrix} = -\hat{\boldsymbol{\jmath}}'\,. \tag{6.13b}$$

Osserviamo subito che le direzioni di $\hat{\boldsymbol{\imath}}'$ e $\hat{\boldsymbol{\jmath}}'$ rimangono invariate (anche se non sempre i vettori mantengono lo stesso orientamento: $\hat{\boldsymbol{\jmath}}'$ viene “ribaltato”); inoltre, nela primo caso cambia anche il modulo del vettore (viene triplicato, in questo caso); la matrice A'_3 che rappresenta L_3 in questa nuova base è, quindi

$$A'_3 = \begin{pmatrix}3 & 0\\0 & -1\end{pmatrix}\,. \tag{6.14}$$

Non vi sono altre direzioni invariate (verificare l'asserto, procedendo come negli esempi 6.1 e 6.2). Alternativamente, A'_3 si può ottenere tramite la matrice del cambio di base N come $A'_3 = N^{-1}\,A\,N$; le colonne della matrice del cambio di base N sono le rappresentazioni dei vettori della nuova base $\{\hat{\boldsymbol{\imath}}', \hat{\boldsymbol{\jmath}}'\}$ nella vecchia base (canonica):

$$N = \frac{\sqrt{2}}{2}\begin{pmatrix}1 & -1\\1 & 1\end{pmatrix}. \tag{6.15}$$

Viene lasciato allo studente il compito di verificare i calcoli.

Siamo ora in possesso di due descrizioni *della stessa applicazione lineare* L_3 *in basi differenti*, ed una di queste è molto più semplice. Se prendiamo l'insieme $\mathcal{C}$ usato nella Figura 6.3, possiamo osservare di questo; si veda la Figura 6.4 al riguardo. L'individuazione delle direzioni inalterate nella trasformazione risulta molto più facile usando il riferimento (x', y'), ossia la nuova base, e, conseguentemente, la descrizione dell'effetto dell'applicazione. Nell'esempio, si ha un ribaltamento nella direzione y' ed un allungamento di un fattore 3 nella direzione x'.

L'esempio 6.3 ci convince che conviene sapere risolvere anche il seguente problema: *esiste una base diversa da quella che sto usando in cui la matrice di rappresentazione del mio operatore è diagonale?*

Quando aumentano le dimensioni, il problema si fa un po' più arduo; vediamo un caso in $\mathbb{E}^3_O$.

Esempio 6.4. Consideriamo l'operatore L_M di $\mathbb{R}^3$ associato nella base canonica $\mathcal{B}_0$ alla seguente matrice

$$M = \begin{pmatrix}0 & -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2}\\ \frac{\sqrt{2}}{2} & \frac{1}{2} & \frac{1}{2}\\ -\frac{\sqrt{2}}{2} & \frac{1}{2} & \frac{1}{2}\end{pmatrix}\,. \tag{6.16}$$

Verifichiamo che il vettore

$$X := [\boldsymbol{v}]_{\mathcal{B}_0} = \begin{pmatrix}0\\1\\1\end{pmatrix} \tag{6.17}$$

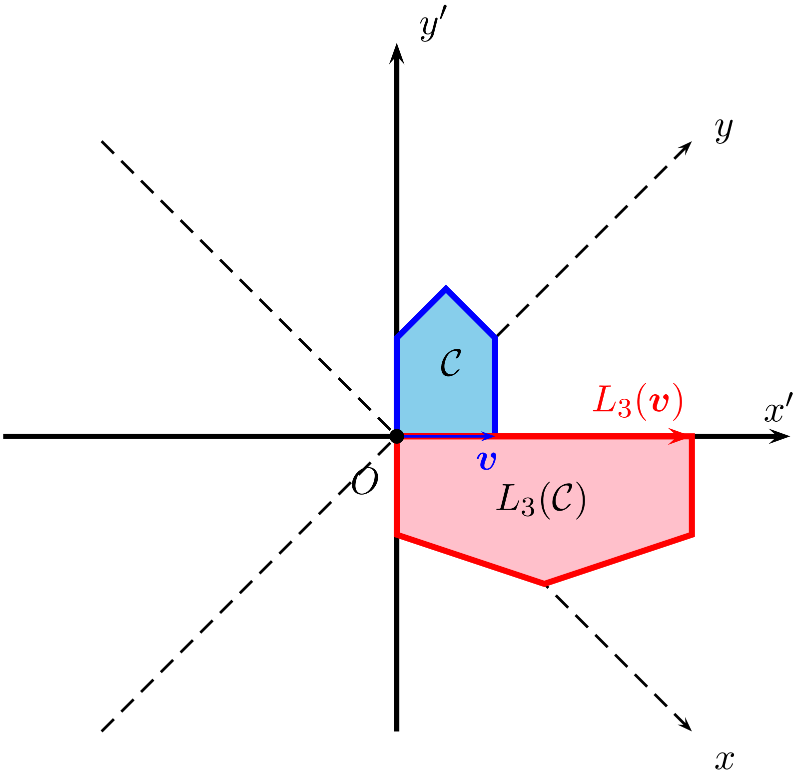

FIGURA 6.4. L'applicazione lineare $L_3 \colon \mathbb{E}^2_O \to \mathbb{E}^2_O$ è rappresentata dalla matrice dell'Eq. (6.14) nella base $\mathcal{B}' = \{\hat{\boldsymbol{\imath}}', \hat{\boldsymbol{\jmath}}'\}$ definita in 6.13 lascia invariate le direzioni degli assi coordinati x' e y'. L'immagine di ogni vettore $L_3(\boldsymbol{v})$ si ottiene rappresentando $\boldsymbol{v}$ sulla nuova base con $X' = [\boldsymbol{v}]_{\mathcal{B}'}$ e calcolando $A'\,X'$. L'immagine dell'insieme $\mathcal{C}$ è equivalente a quella descritta ottenuta mediante la matrice A_3 che rappresenta L_3 nella base canonica (Figura 6.3).

rimane invariato sotto l'azione di L_M.

$$
\begin{aligned}
M\,X &= \begin{pmatrix} 0 & -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{1}{2} & \frac{1}{2} \\ -\frac{\sqrt{2}}{2} & \frac{1}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \\
&= 0 \begin{pmatrix} 0 \\ \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} \end{pmatrix} + 1 \begin{pmatrix} -\frac{\sqrt{2}}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{pmatrix} + 1 \begin{pmatrix} \frac{\sqrt{2}}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{pmatrix} \\
&= \begin{pmatrix} -\frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2} \\ \frac{1}{2} + \frac{1}{2} \\ \frac{1}{2} + \frac{1}{2} \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} .
\end{aligned}
\tag{6.18}
$$

Quindi $L_M(\boldsymbol{v}) = \lambda \boldsymbol{v}$ con $\lambda = 1$.

Inoltre, non vi sono altre direzioni che rimangono invariate. Per convincersi del risultato, è necessario affrontare il problema come fatto negli esempi sopra, e studiare esplicitamente il sistema

$$MX = \lambda X\,.$$

Osserviamo che, per arrivare a trovare il vettore $\boldsymbol{v}$ proposto nella (6.17) aiuta sapere per quale λ effettuare la ricerca: quando $\lambda = 1$, il sistema diventa

$$\begin{cases} -\frac{\sqrt{2}}{2}y + \frac{\sqrt{2}}{2}z & = x \\ \frac{\sqrt{2}}{2}x + \frac{1}{2}y + \frac{1}{2}z & = y \\ -\frac{\sqrt{2}}{2}x + \frac{1}{2}y + \frac{1}{2}z & = z \end{cases}\,. \tag{6.19}$$

Sottraendo membro a membro le ultime due equazioni, abbiamo

$$\sqrt{2}x = y - z$$

cioè

$$x = \frac{\sqrt{2}}{2}(y - z)$$

e la prima può essere scritta come

$$-\frac{\sqrt{2}}{2}(y - z) = x\,.$$

Quindi, $x = -x$, e allora $x = 0$ e $y = z$.

Senza sapere con quale valore di λ operare, diventa tutto più complicato. Lo studente è invitato a provare a effettuare i calcoli relativi per convincersi delle complicazioni dei calcoli.

Ci serve, pertanto, un metodo sistematico per affrontare il problema.

2. Autovalori ed autovettori: definizioni

Gli esempi che abbiamo fatto nella Sezione 1 giustificano le definizioni che stiamo per dare e le conseguenti tecniche di calcolo.

Definizione 6.1. Sia $\mathcal{V}$ uno spazio vettoriale sul campo dei numeri reali $\mathbb{R}$ di dimensione n; sia $L\colon \mathcal{V} \to \mathcal{V}$ un operatore lineare. Diciamo che un vettore $\boldsymbol{v} \in \mathcal{V}$ è un'*autovettore* di L relativo all'*autovalore* λ se esiste $\lambda \in \mathbb{R}$ tale che

$$L(\boldsymbol{v}) = \lambda \boldsymbol{v}\,, \tag{6.20}$$

con $\boldsymbol{v} \neq \boldsymbol{0}_{\mathcal{V}}$

Conviene avere un termine per indicare l'insieme di tutti gli autovalori di un operatore L: viene detto *spettro* di L l'insieme di tutti i possibili autovalori, ossia

$$\operatorname{Spec}(L) = \{\lambda \in \mathbb{R} \mid \exists \boldsymbol{v} \in \mathcal{V} \colon L(\boldsymbol{v}) = \lambda \boldsymbol{v}\,, \boldsymbol{v} \neq \boldsymbol{0}_{\mathcal{V}}\}\,.$$

Alla luce della definizione data, possiamo rileggere gli esempi studiati in precedenza:

- nell'esempio 6.1 non esistono autovettori;

- nell'esempio 6.2 abbiamo trovato un autovalore $\lambda_1 = 1$ con relativo autovettore, rappresentato nella base canonica $\begin{pmatrix}1\\1\end{pmatrix}$ e l'autovalore $\lambda_2 = -1$, con relativo autovettore rappresentato da $\begin{pmatrix}1\\-1\end{pmatrix}$;
- nell'esempio 6.3 la matrice A'_3 che rappresenta la L_3 sulla nuova base è diagonale; la determinazione di autovalori ed autovettori è particolarmente semplice, ed otteniamo $\lambda_1 = 2$ con autovettore $\hat{\boldsymbol{\imath}}$ e $\lambda_2 = -1$, con autovettore $\hat{\boldsymbol{\jmath}}$;
- infine, nell'esempio 6.4 abbiamo trovato solo l'autovalore $\lambda = 1$ e relativo autovettore rappresentato da $\begin{pmatrix}0\\1\\1\end{pmatrix}$

Quando abbiamo un'operatore L sullo spazio vettoriale $\mathcal{V}$, la ricerca dei suoi possibili autovalori ed i corrispondenti autovettori, come abbiamo visto, può consentire di semplificare lo studio dell'effetto dell'operatore; questo costituisce, pertanto, un problema "standard".

Definizione 6.2. Sia $\mathcal{V}$ uno spazio vettoriale sul campo dei numeri reali $\mathbb{R}$ di dimensione n; sia $L\colon \mathcal{V} \to \mathcal{V}$ un operatore lineare.
Risolvere il *problema agli autovalori* per L significa trovare gli eventuali autovalori e caratterizzarne i corrispondenti autovettori.

Osservazione 6.5. La definizione 6.1 è valida per qualunque operatore lineare, si qualunque spazio vettoriale; in particolare, non si richiede che lo spazio sia finitamente generato. Tuttavia, la trattazione nel caso di spazi vettoriali di dimensione infinita esula dagli scopi del corso, e richiede tecniche diverse da quelle che saranno illustrate. Pertanto, **d'ora in avanti, intendiamo sempre che la dimensione dello spazio vettoriale sia finita e pari a n.**

Se ci limitiamo, quindi, a spazi di dimensione finita, notiamo che sappiamo risolvere problemi agli autovalori (6.2) una volta che passiamo alle rappresentazioni dei vettori su una base ed alle equazioni corrispondenti.

Noi lavoreremo quasi sempre negli spazi vettoriali reali di tipo $\mathbb{R}^n$, ma qualche volta negli spazi vettoriali complessi di tipo $\mathbb{C}^n$. Se consideriamo la rappresentazione dei vettori sula base canonica (o su un'altra base) potremo sempre associare una matrice reale (o complessa) ad un operatore lineare, e parleremo in tal caso degli autovettori (e autovalori) di una matrice. La definizione 6.1 data sopra si specializza come segue.

Definizione 6.3 (caso reale). Sia $A \in \mathcal{M}_{\mathbb{R}}(n)$ una matrice quadrata di ordine n a entrate reali. Diciamo che un $X \in \mathbb{R}^n$ è un'*autovettore* di A relativo all'*autovalore* λ se esiste $\lambda \in \mathbb{R}$ tale che

$$AX = \lambda X\,, \tag{6.21}$$

con $X \neq \mathbf{0}_n$.

Analogamente, chiameremo *spettro* della matrice A, indicato con $\operatorname{Spec}(A)$, l'insieme degli autovalori di A.

Osservazione 6.6. Alle condizioni della definizione 6.3, possiamo considerare l'operatore lineare $L_A\colon \mathbb{R}^n \to \mathbb{R}^n$ associato alla matrice A in una base $\mathcal{B}$ dello spazio vettoriale *reale* $\mathbb{R}^n$, ed $X = [\boldsymbol{v}]_{\mathcal{B}}$ coincidente con rappresentazione di un vettore $\boldsymbol{v} \in \mathbb{R}^n$ nella stessa base. Quando non specificato altrimenti, si intenderà che la base usata è quella canonica $\mathcal{B}_0$, ed $X = \boldsymbol{v} \in \mathbb{R}^n$. In altre parole, possiamo partire direttamente dalla matrice e sottintendere l'operatore ad essa associato.

Vediamo un esempio:

Esempio 6.7. Consideriamo la matrice

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \tag{6.22}$$

Cerchiamo di trovare se ci sono autovalori/autovettori; impostiamo le equazioni:

$$\begin{cases} -y & = \lambda x \\ x & = \lambda y \end{cases} \tag{6.23}$$

Si vede subito sostituendo la prima nella seconda che abbiamo:

$$\begin{cases} -y & = \lambda x \\ x & = -\lambda^2 x \end{cases} . \tag{6.24}$$

La seconda equazione del sistema ci dà

$$(1 + \lambda^2)x = 0\,, \tag{6.25}$$

e, poiché non esiste alcun numero reale che renda $1 + \lambda^2 = 0$ (somma di due quantità positive, di cui una strettamente positiva), esiste solo la soluzione banale $x = y = 0$.

Niente autovalori!

Le cose cambierebbero in campo complesso: si veda a tale proposito l'appendice al capitolo).

Assodato che le cose possono cambiare se gli scalari sono presi in $\mathbb{R}$ o in $\mathbb{C}$, **per ora, comunque, lavoreremo in $\mathbb{R}$, finché possiamo**.

Come abbiamo visto nell'esempio 6.3, avere una matrice diagonale semplifica notevolmente la ricerca degli autovettori e degli autovalori (anzi, la rende banale); d'altro canto, abbiamo visto nell'esempio 6.7 anche che non sempre possiamo trovare autovettori ed autovalori. Il nostro obiettivo è proprio quello di arrivare ad avere strumenti per capire quando potremo farlo (e come); in effetti, ci interessa sapere quando possiamo mettere in forma **diagonale** il nostro operatore. Cerchiamo, come al solito, di formalizzare la cosa. Cominciamo con un risultato generale semplice, ma importante.

Proposizione 6.1. *Sia $L\colon V \to V$ un operatore lineare di uno spazio vettoriale; sia $\mathcal{B}$ una base di V. $\mathcal{B}$ è composta di autovettori di V se e solo se la matrice che rappresenta L nella base $\mathcal{B}$ è diagonale.*

Dimostrazione. Sia $\mathcal{B} = \{\boldsymbol{v}_1, \dots \boldsymbol{v}_n\}$ la base, ed $A = [L]_{\mathcal{B}}$ la matrice che rappresenta l'operatore nella base. Se i vettori $\boldsymbol{v}_i$ sono autovettori, relativi agli autovalori λ_i rispettivamente, possiamo dire che la $j-$esima colonna di A è data da

$$A^j = [L(\boldsymbol{v}_j)]_{\mathcal{B}} = \lambda_i[(\boldsymbol{v}_j)]_{\mathcal{B}} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \lambda_j \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

dove l'unico termine non nullo è proprio λ_j, nella $j-$esima riga. Quindi la matrice A è diagonale

$$A = \begin{pmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & & & \ddots & 0 \\ 0 & & \cdots & 0 & \lambda_n \end{pmatrix}. \tag{6.26}$$

Viceversa, se la matrice è diagonale come nella (6.26), la j-esima colonna di A è l'immagine del vettore $\boldsymbol{v}_j$ della base, rappresentato nella base stessa:

$$[L(\boldsymbol{v}_j)]_{\mathcal{B}} = A^j = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \lambda_j \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \lambda_i[(\boldsymbol{v}_j)]_{\mathcal{B}},$$

e, quindi, $L(\boldsymbol{v}_j) = \lambda \boldsymbol{v}_j$ per l'isomorfismo di rappresentazione, (4.12), cioè $\boldsymbol{v}_j$ è un autovettore di L relativo all'autovalore λ_j.

(Naturalmente, non necessariamente tutti gli autovalori devono essere diversi, né è unico l'ordine con il quale gli autovalori sono disposti sulla diagonale, il quale dipende dall'ordine con cui sono messi i vettori della base). □

Le considerazioni fatte sulle matrici simili consentono di esprimere quanto detto in altra maniera:

Corollario 6.2. *sia A una matrice quadrata di ordine n, e sia L_A l'operatore associato ad A (o, viceversa, sia L un operatore ed A la matrice associata). Allora le seguenti affermazioni sono equivalenti:*

(1) *esiste una base di autovettori per lo spazio dell'operatore;*
(2) *A è simile ad una matrice diagonale;*
(3) *la classe di similitudine di A contiene una matrice diagonale.*

Dimostrazione. Ricordiamo che, se esiste una base di autovettori, la matrice invertibile $N \in \mathcal{GL}(n, \mathbb{R})$ di cambio di base trasforma la matrice A nella matrice $N^{-1}AN$, che è diagonale, quindi A è simile ad una matrice diagonale. Viceversa, se A è simile ad una matrice diagonale D, esiste una matrice invertibile $N \in \mathcal{GL}(n, \mathbb{R})$ tale che $D = N^{-1}AN$, quindi la matrice N rappresenta un cambio di base: le sue colonne sono le rappresentazioni dei vettori della nuova base $\mathcal{B}' = \{\boldsymbol{v}_1, \dots \boldsymbol{v}_n\}$ in quella di partenza, e questi vettori sono autovettori; nella nuova base, infatti $Y' = DX'$, e se $X' = [\boldsymbol{v}_i]_{\mathcal{B}'}$, $Y' = [L_D(\boldsymbol{v}_i)]_{\mathcal{B}'} = d_{i,i}[\boldsymbol{v}_i]_{\mathcal{B}'}$. In altre parole, il vettore $[\boldsymbol{v}_i]_{\mathcal{B}'}$ è

$$\begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

con l'unico elemento diverso da 0 nella riga $i-$esima. $[L_D(\boldsymbol{v}_i)]_{\mathcal{B}'}$, quindi, è l'$i-$esima colonna della matrice D: ma questa è diagonale, quindi l'unico elemento non nullo è quello nella $i-$esima riga. Notiamo, quindi, che $d_{i,i}$ è l'autovalore dell'autovettore $\boldsymbol{v}_i$. Lasciamo lo studente convincersi dell'equivalenza del secondo punto con il terzo punto. □

Ecco finalmente che possiamo cominciare a tirare le somme, e dare senso alla seguente definizione:

Definizione 6.4 (Diagonalizzabilità). Un operatore lineare di V (o una matrice A quadrata di ordine n, cui è associabile un operatore) è **diagonalizzabile** se esiste una base dello spazio in cui opera formata da suoi autovettori.

Cioè, **se riusciamo ad effettuare un cambio di base in cui la matrice di rappresentazione dell'operatore è diagonale; inoltre, sulla diagonale si trovano gli autovalori**.

Come possiamo trovare gli autovalori di un operatore L, esplicitamente? Fissiamo una base $\mathcal{B}$ e rappresentiamo i vettori dello spazio con la n-upla

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n \,;$$

detta A la matrice che rappresenta l'operatore L su $\mathcal{B}$, dobbiamo risolvere il sistema

$$(A - \lambda I_n)X = \mathbf{0}_n. \tag{6.27}$$

trovando soluzioni diverse dal vettore nullo (che non va bene come autovettore).

Sappiamo che il sistema ha soluzioni non banali se e solo se

$$\det(A - \lambda I_n) = 0\,. \tag{6.28}$$

Sorge però un dubbio: le soluzioni che troveremo in generale dipendono dalla base scelta; chi ci garantisce che, se un certo operatore è rappresentato da una matrice A nella base $\mathcal{B}$ e da una matrice A' nella base $\mathcal{B}'$, gli autovalori non dipendono dalla base scelta? Dobbiamo studiare il problema, e convincersi che i valori di λ che permettono di soddisfare la 6.28 non dipendono dalla base (in questo modo avremo modo di trovare gli eventuali autovalori, e, corrispondentemente, caratterizzare gli autovettori). Questo viene fatto dal seguente

TEOREMA 6.3. *Sia $L\colon \mathbb{R}^n \to \mathbb{R}^n$ un operatore lineare. Sia $\mathcal{B}$ una base di $\mathbb{R}^n$. Sia A la matrice che rappresenta L sulla base $\mathcal{B}$. Allora, valgono le seguenti affermazioni.*

(1) *La funzione $p_A(\lambda)$ usata nell'equazione (6.28)*
$$p_A(\lambda) = \det(A - \lambda I_n)$$
non dipende dalla base, e, quindi, è caratteristica dell'operatore L. Inoltre, $p_A(\lambda)$ è un polinomio in λ (detto polinomio caratteristico*).*

(2) *Dire che un certo valore di λ_0 è autovalore di A equivale a dire che λ_0 è una radice del polinomio caratteristico, ossia:*
$$p_A(\lambda_0) = 0 \iff \lambda_0 \in \operatorname{Spec}(A)\,.$$

(3) *Il polinomio $p_A(\lambda)$ è di grado n. Il termine di grado massimo del polinomio è $(-1)^n\lambda^n$, quello di grado $n-1$ è $(-1)^{n-1}\lambda^{n-1}\operatorname{tr}(A)$, ed il termine noto è $\det(A)$.*

DIMOSTRAZIONE. Cominciamo con il primo punto. Intanto, sicuramente quando calcoliamo $\det(A - \lambda I_n)$, combiniamo le entrate secondo prodotti e somme, quindi otteniamo un polinomio in λ. Il polinomio caratteristico nella base $\mathcal{B}$ è $p_A(\lambda) = \det(A - \lambda I_n)$. Sia ora A' la matrice che rappresenta L in una nuova base $\mathcal{B}'$; deve esistere una matrice di cambio di base $N \in \mathcal{GL}(n, \mathbb{R})$ invertibile, tale che
$$A' = N^{-1}AN\,.$$
Calcolando il polinomio caratteristico in questa base, si ottiene
$$\begin{aligned} p_{A'}(\lambda) &= \det(A' - \lambda I_n) = \det(N^{-1}AN - \lambda I_n) \\ &= \det(N^{-1}AN - \lambda N^{-1}I_nN) = \det(N^{-1}(A - \lambda I_n)N) \\ &= \det(N^{-1})\det(A - \lambda I_n)\det(N) \\ &= \det(A - \lambda I_n) = p_A(\lambda) \end{aligned} \tag{6.29}$$
dove abbiamo sfruttato il teorema di Binet e le proprietà dell'identità e dell'inversa per dire che $N^{-1}I_n\,N = N^{-1}N = I_n$; inoltre,
$$\det(N^{-1}) = \frac{1}{\det(N)}.$$
Il polinomio caratteristico, quindi, non dipende dalla base, quindi ha senso denotarlo $p_A(\lambda)$, indicando la sola matrice (tanto, se rappresentiamo l'operatore associato a quella matrice in un'altra base, il polinomio è lo stesso), ma potremmo anche indicarlo con $p_L(\lambda)$, dove L è l'operatore associato ad A.

Per il secondo punto, osserviamo che se $p_A(\lambda_0) = 0$ il rango della matrice $A - \lambda_0 I_n$ è minore di n, quindi il suo nucleo non è banale, quindi esiste un vettore non nullo tale che $(A-\lambda_0 I_n)X=\mathbf{0}_n$, cioè $AX=\lambda_0 X$, ossia X è autovettore con autovalore λ_0.

Viceversa, se λ_0 è autovalore, esiste una soluzione non banale del sistema $(A-\lambda_0 I_n)X$, quindi deve essere $\det(A-\lambda_0 I_n)=p_A(\lambda_0)=0$.

Veniamo al terzo punto: anzitutto, osserviamo che il termine noto del polinomio si ottiene ponendo $\lambda=0$, quindi sarà dato $\boldsymbol{p}_A(0)=\det(A)$, come affermato. Per il resto dell'enunciato, potremo procedere per induzione. Per ora, mostriamo la validità dell'enunciato per $n=2$; posto

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

abbiamo:

$$\begin{aligned} p_A(\lambda) &= \begin{vmatrix} a_{1,1}-\lambda & a_{1,2} \\ a_{2,1} & a_{2,2}-\lambda \end{vmatrix} = (a_{1,1}\lambda)(a_{2,2}-\lambda)-a_{1,2}a_{2,1} \\ &= \lambda^2-(a_{1,2}+a_{2,1})\lambda+a_{1,1}a_{2,2}-a_{1,2}a_{2,1}=\lambda^2-\lambda\operatorname{tr}(A)+\det(A) \end{aligned} \tag{6.30}$$

che è quanto affermato.

Possiamo ora procedere con l'induzione; sia ora vero l'enunciato fino ad $n-1$, vediamo che è vero per n. Calcoliamo il polinomio caratteristico:

$$\begin{aligned} p_A(\lambda)=\det(A-\lambda I_n) &= \begin{vmatrix} a_{1,1}-\lambda & a_{1,2} & \cdots & \\ a_{2,1} & a_{2,2}-\lambda & \cdots & \\ \vdots & & \ddots & \\ \vdots & & & a_{n,n}-\lambda \end{vmatrix} \\ &= (a_{1,1}-\lambda)\begin{vmatrix} a_{2,2}-\lambda & \cdots & \\ \vdots & \ddots & \\ \vdots & & a_{n,n}-\lambda \end{vmatrix} + q_{n-2}\,, \end{aligned} \tag{6.31}$$

dove q_{n-2} indica il resto dello sviluppo del determinante secondo la prima colonna: questo termine ha grado al più $n-2$, perché nelle sottomatrici in cui si calcola il complemento algebrico, mancano due volte i termini del tipo $a_{k,k}-\lambda$ (una volta per la prima riga, un'altra per la riga $i-$esima corrispondente al termine $a_{i,1}$ della prima colonna nel calcolo), e questi complementi algebrici verranno moltiplicati per gli elementi della colonna dalla seconda riga in giù, dove λ non compare più.

Osserviamo, adesso, che, per calcolare $p_A(\lambda)$ nella 6.31, ci siamo ricondotti a calcolare il determinante di una matrice $(n-1)\times(n-1)$ della forma $\det(B-\lambda I_{n-1})$, dove B è la sottomatrice ottenuta da A cancellando la prima riga e la prima colonna. Per ipotesi induttiva (visto che abbiamo cioè supposto che il teorema fosse vero fino a $n-1$), possiamo scrivere:

$$\det(A-\lambda I_n)=(a_{1,1}-\lambda)[(-1)^{n-1}\lambda^{n-1}+(-1)^{n-2}\lambda^{n-2}\operatorname{tr}(B)+\cdots+\det(B)]\,. \tag{6.32}$$

Sviluppando i calcoli, troviamo, riordinando i termini ed esplicitando solo i termini di grado utile (non ci interessano i termine finali di grado $n-2$, $n-3$ ecc. –neanche il termine noto, tanto abbiamo già stabilito la sua sorte):

$$
\begin{aligned}
p_A(\lambda) &= \det(A - \lambda I_n) \\
&= a_{1,1}[(-1)^{n-1}\lambda^{n-1} + (-1)^{n-2}\lambda^{n-2}\operatorname{tr}(B) + \cdots + \det(B)] \\
&\quad - \lambda[(-1)^{n-1}\lambda^{n-1} + (-1)^{n-2}\lambda^{n-2}\operatorname{tr}(B) + \cdots + \det(B)] \\
&= (-1)(-1)^{n-1}\lambda\lambda^{n-1} + a_{1,1}(-1)^{n-1}\lambda^{n-1} + (-\lambda)(-1)^{n-2}\lambda^{n-2}\operatorname{tr}(B) \\
&\quad + \dots \\
&= (-1)^n\lambda^n + a_{1,1}(-1)^{n-1}\lambda^{n-1} + (-1)(-1)^{n-2}\lambda\lambda^{n-2}\operatorname{tr}(B) + \dots \\
&= (-1)^n\lambda^n + a_{1,1}(-1)^{n-1}\lambda^{n-1} + (-1)^{n-1}\lambda^{n-1}(a_{2,2} + \cdots + a_{n,n}) + \dots \\
&= (-1)^n\lambda^n + (-1)^{n-1}\lambda^{n-1}(a_{1,1} + a_{2,2} + \cdots + a_{n,n}) + \dots \\
&= (-1)^n\lambda^n + (-1)^{n-1}\lambda^{n-1}\operatorname{tr}(A) + \dots
\end{aligned}
\tag{6.33}
$$

e i termini dopo i puntini hanno grado inferiore a $n-1$, quindi non ci interessano.

Dunque, anche per n è vero che il termine di grado massimo è $(-1)^n\lambda^n$, e quello immediatamente sotto è $(-1)^{n-1}\lambda^{n-1}\operatorname{tr}(A)$; allora è vero sempre (per $n=2$ vale, allora vale per $n=3$; allora vale per $n=4$, ecc.). □

Ora che abbiamo un metodo per identificare gli autovalori; cerchiamo di analizzare quello che possiamo dire per gli autovettori. In base alle considerazioni appena fatte circa il polinomio caratteristico, quando servirà potremo limitarci a lavorare esplicitamente nella rappresentazione su una data base $\mathcal{B}$, ossia a considerare vettori $X \in \mathbb{R}^n$.

Dobbiamo cercare di caratterizzare gli insiemi formati da autovettori relativi ad un dato autovalore.

Osservazione 6.8. Una volta trovato un autovettore, tutti i vettori nel sottospazio generato da esso sono ancora autovettori. In altre parole, sia $\boldsymbol{v}$ un autovettore di un operatore $L\colon \mathcal{V} \to \mathcal{V}$; allora, $\alpha\boldsymbol{v}$ è ancora autovettore di L; ossia $\operatorname{Span}(\boldsymbol{v})$ è formato da autovettori di L relativi all'autovalore λ (e dal vettore nullo).

Dimostrazione. Banalmente, essendo L lineare

$$L(\alpha\boldsymbol{v}) = \alpha L(\boldsymbol{v}) = \alpha\lambda\boldsymbol{v} = \lambda(\alpha\boldsymbol{v})\,,$$

usando le proprietà della moltiplicazione per uno scalare □

L'osservazione precedente porta a chiederci: *dato un autovalore λ, com'è fatto l'insieme dei suoi autovettori?* La risposta viene nella seguente definizione e con l'analisi conseguente.

Definizione 6.5. Sia A una matrice reale di ordine n; sia $\lambda \in \operatorname{Spec}(A)$. Si chiama *autospazio* associato a λ (o, semplicemente, di λ) l'insieme V_λ di tutti i vettori che soddisfano l'equazione (6.27):

$$V_\lambda = \{X \in \mathbb{R}^n \mid A\,X = \lambda X\}\ .$$

Osservazione 6.9. In questa maniera, il vettore nullo è incluso in V_λ: $A\,\mathbf{0}_n = \mathbf{0}_n = \lambda\mathbf{0}_n$, quindi, per definizione $\mathbf{0}_n \in V_\lambda$. Un po' impropriamente si usa dire che V_λ è l'insieme degli autovettori di λ. A rigor di logica, in base alla definizione, il vettore nullo non può essere annoverato fra gli autovettori.

L'inclusione del vettore nullo nell'autospazio V_λ è importante, perché consente di affermare che

Proposizione 6.4. *L'autospazio associato ad un autovalore è un sottospazio vettoriale (da cui appunto il nome), di dimensione $n - \operatorname{rg}(A)$.*

Dimostrazione. Basta osservare che

$$AX = \lambda X = \lambda I_n X,$$

ossia

$$(A - \lambda I_n)X = \mathbf{0}_n$$

quindi l'insieme V_λ coincide con il nucleo della matrice $(A - \lambda I_n)$, ed è pertanto un autospazio (4.4).

Sfruttando il Teorema delle Dimensioni 4.8 visto nel Capitolo 4, siamo in grado subito di affermare che

$$\dim(V_\lambda) = n - \operatorname{rg}(A - \lambda I_n)\,. \tag{6.34}$$

□

Osservazione 6.10. La dimensione di V_{λ_i} è compresa fra 1 ed n per ogni $i = 1, \dots h$:

$$1 \leq \dim(V_{\lambda_i}) \leq n \qquad \forall\, i = 1, \dots h\,;$$

Dimostrazione. La dimensione è almeno 1, perché, se esiste un autovalore λ, deve esistere un autovettore $X \neq \mathbf{0}_n$, e allora $\operatorname{Span}(X) \subseteq V_\lambda$ (osservazione 6.8). La dimensione non può superare quella dello spazio $\mathbb{R}^n$, di cui V_λ è sottoinsieme. □

3. Autospazi vettoriali e somma diretta

Può uno stesso vettore stare in autospazi associati ad autovalori differenti? no, basta considerare la seguente proposizione:

Proposizione 6.5. *Se λ_1 e λ_2 sono autovalori distinti di un operatore lineare, allora l'intersezione fra i due autospazi associati contiene solo il vettore nullo.*

Dimostrazione. Semplicemente, se $\boldsymbol{u} \in V_{\lambda_1}$, allora

$$L(\boldsymbol{u}) = \lambda_1 \boldsymbol{u};$$

analogamente, se $\boldsymbol{u} \in V_{\lambda_2}$

$$L(\boldsymbol{u}) = \lambda_2 \boldsymbol{u}.$$

Un vettore nell'intersezione soddisfa entrambe le equazioni; sottraendo membro a membro, otteniamo:

$$L(\boldsymbol{u}) - L(\boldsymbol{u}) = L(\boldsymbol{u} - \boldsymbol{u}) = L(\mathbf{0}_\mathcal{V}) = \mathbf{0}_\mathcal{V}\,,$$

ma anche

$$L(\boldsymbol{u}) - L(\boldsymbol{u}) = \lambda_1 \boldsymbol{u} - \lambda_2 \boldsymbol{u} = (\lambda_1 - \lambda_2)\boldsymbol{u}\,.$$

Cioè, $(\lambda_1 - \lambda_2)\boldsymbol{u} = \mathbf{0}_{\mathcal{V}}$: siccome i due autovalori sono diversi, questo implica che $\boldsymbol{u} = \mathbf{0}_{\mathcal{V}}$. □

Risulta a questo punto immediato il seguente

Corollario 6.6. *Se λ_1 e λ_2 sono autovalori distinti di un operatore lineare, allora i due autospazi relativi, V_{λ_1} e V_{λ_2} sono in somma diretta.*

Dimostrazione. Infatti, l'intersezione fra i due spazi contiene solo il vettore nullo, in virtù della Proposizione 6.5. □

A questo punto, costa poca fatica convincersi che *tutti* gli autospazi sono in somma diretta. Per la dimostrazione, ci viene utile un risultato semplice, ma efficace sugli spazi in somma diretta, in generale.

Lemma 6.7. *Siano $V_1, V_2, \dots V_h$ h sottospazi vettoriali di uno spazio vettoriale su campo $\mathbb{R}$ qualunque; se i sottospazi sono in somma diretta, allora l'insieme $\mathcal{S} = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \dots \boldsymbol{u}_h\}$ costruito prendendo $\boldsymbol{u}_i \in V_i$, con $\boldsymbol{u}_i \neq \mathbf{0}_{\mathcal{V}}$ per $i = 1 \dots h$ (cioè "pescando" un vettore in ciascun sottospazio in modo che ciascuno non sia nullo) è un sistema di vettori linerarmente indipendenti. Viceversa, se, comunque si costruisce l'insieme $\mathcal{S}$, come sopra si ottengono vettori linearmente indipendenti, allora gli spazi sono in somma diretta.*

Dimostrazione. Possiamo procedere per induzione. Per un vettore è ovvio. Supponiamo che sia vero se prendiamo $h-1$ vettori; vediamo che è vero per h vettori. Consideriamo una combinazione lineare dei vettori di $\mathcal{S}$ che produca il vettore nullo:

$$\alpha_1 \boldsymbol{u}_1 + \alpha_2 \boldsymbol{u}_2 + \dots \alpha_h \boldsymbol{u}_h = \mathbf{0}_{\mathcal{V}} \,. \tag{6.35}$$

Supponiamo per assurdo che esista una combinazione non banale; allora c'è almeno un coefficiente α_i non nullo nella (6.35); senza perdere in generalità, possiamo sempre dire che questo è l'ultimo. Possiamo in tale caso scrivere

$$\boldsymbol{u}_h = -\frac{\alpha_1}{\alpha_h}\boldsymbol{u}_1 + \dots - \frac{\alpha_{h-1}}{\alpha_h}\boldsymbol{u}_{h-1} \,. \tag{6.36}$$

Ora, $\boldsymbol{u}_h \neq \mathbf{0}$, per ipotesi; inoltre, $\boldsymbol{u}_h$ risulta essere una combinazione lineare di $h-1$ vettori presi negli altri $h-1$ sottospazi in somma diretta. Quindi, si avrebbe $\boldsymbol{u}_h \in (V_1 \oplus V_2 + \dots V_{h-1})$. Daltra parte, $\boldsymbol{u}_h \in V_h$: avremmo trovato un vettore diverso da quello nullo nell'intersezione:

$$\boldsymbol{u}_h \in (V_1 \oplus V_2 + \dots V_{h-1}) \cap V_h \,,$$

ma ciò è impossibile, perché per ipotesi i sottospazi sono in somma diretta, quindi sicuramente $(V_1 \oplus V_2 + \dots V_{h-1}) \cap V_h = \{\mathbf{0}_{\mathcal{V}}\}$ L'assurdo nasce dall'avere ipotizzato che la combinazione lineare della (6.35) potesse essere non banale. Allora l'unica combinazione che produce il vettore nullo è quella banale, ossia i vettori sono linearmente indipendenti.

Vediamo ora l'implicazione opposta: Cominciamo con due sottospazi, V_1 e V_2; per ipotesi, ogni sistema di vettori non nulli $\mathcal{S} = \{\boldsymbol{u}_1, \boldsymbol{u}_2\}$ con $\boldsymbol{u}_1 \in V_1$ e $\boldsymbol{u}_2 \in V_2$ è formato da vettori linearmente indipendenti. Supponiamo che V_1 e V_2 non siano in somma diretta: esisterebbe un vettore $\boldsymbol{v}$ non nullo nell'intersezione; quindi, avremmo trovato $\boldsymbol{u}_1 = \boldsymbol{v} \in V_1$ e $\boldsymbol{u}_2 = \boldsymbol{v} \in V_2$ tali che $\boldsymbol{u}_1 - \boldsymbol{u}_2 = \mathbf{0}$, ossia due vettori linearmente dipendenti, che non può essere. Allora, per forza $V_1 + V_2 = V_1 \oplus V_2$.

Sia ora vero l'asserto per $h-1$. Se V_h non fosse in somma diretta con $(V_1 \oplus V_2 \oplus \dots V_{h-1})$, dovremmo trovare $\boldsymbol{v}$ non nullo nell'intersezione:

$$\boldsymbol{v} \in (V_1 \oplus V_2 \oplus \dots V_{h-1}) \cap V_h \,,$$

ossia $\boldsymbol{u}_h = \boldsymbol{u}_1 + +\boldsymbol{u}_2 + \dots + \boldsymbol{u}_{h-1}$, con $\boldsymbol{u}_i \in V_i$ per $i = 1 \dots h$. Allora, $\boldsymbol{u}_1 + +\boldsymbol{u}_2 + \dots + \boldsymbol{u}_{h-1} - \boldsymbol{u}_h = \mathbf{0}$, e, come sopra, avremmo una combinazione di vettori non banale che produce il vettore nullo, che non può essere, visto che stiamo supponendo che pescando un vettore non nullo in ciascun spazio, otteniamo h vettori linearmente indipendenti. Quindi

$$(V_1 \oplus V_2 \oplus \dots V_{h-1}) + V_h = V_1 \oplus V_2 \oplus \dots V_{h-1} \oplus V_h \,;$$

l'asserto è vero anche per h. Allora, per induzione è vero per ogni h. □

Una conseguenza immediata è questa

Corollario 6.8. *Siano $V_1, V_2, \dots V_h$ in somma diretta; se*

$$\boldsymbol{v}_1 + \dots + \boldsymbol{v}_h = \mathbf{0}$$

con $\boldsymbol{v}_i \in V_i$ per $i = 1 \dots h$, allora tutti i vettori $\boldsymbol{v}_i$ devono essere nulli.

Dimostrazione. Se così non fosse, ce ne sarebbe almeno uno, e quindi, almeno due non nulli; pertanto, potrei ricavare almeno uno dei vettori come combinazione lineare degli altri: il che va contro il risultato del Lemma 6.7 precedente. □

Possiamo ora dimostrare il risultato che volevamo:

Teorema 6.9. *Sia A una matrice reale di ordine n; sia $\operatorname{Spec}(A) = \{\lambda_1, \dots \lambda_h\}$ con $i = 1, \dots h$ lo spettro degli h autovalori distinti di A (se non vuoto). Allora gli autospazi relativi sono in somma diretta*

$$V_{\lambda_1} \oplus V_{\lambda_2} \oplus \dots + V_{\lambda_h} \,.$$

Dimostrazione. Per due autospazi relativi ad autovalori distinti è vero (Corollario 6.6); cominciamo a considerare la somma diretta $V_{\lambda_1} \oplus V_{\lambda_2}$. Mostriamo che V_{λ_3} è in somma diretta con questa somma. Basta vedere che V_{λ_3} è in somma diretta con $V_{\lambda_1} \oplus V_{\lambda_2}$, ossia che l'intersezione di questi contiene solo il vettore nullo.

Sia $X \in V_{\lambda_3}$; deve essere

$$AX = \lambda_3 X \,. \tag{6.37}$$

Se X appartiene anche a $V_{\lambda_1} \oplus V_{\lambda_2}$, possiamo scrivere $X = Y_1 + Y_2$ (decomposizione unica), con $Y_1 \in V_{\lambda_1}$ e $Y_2 \in V_{\lambda_2}$, cioè

$$AX = \lambda_1 Y_1 + \lambda_2 Y_2 \,. \tag{6.38}$$

Confrontando le due equazioni (6.37) e (6.37), abbiamo

$$\lambda_3 X = \lambda_3(Y_1 + Y_2) = \lambda_1 X_1 + \lambda_2 X_2 \,,$$

ossia

$$(\lambda_1 - \lambda_3)X_1 + (\lambda_2 - \lambda_3)X_2 = \mathbf{0}_n \,. \tag{6.39}$$

Per ipotesi, V_1 e V_2 sono in somma diretta; poichè la combinazione lineare nella (6.39) è non banale (per ipotesi gli autovalori sono distinti): per il Lemma 6.7, se Y_1 e Y_2 sono non nulli, sono linearmente indipendenti, e una loro combinazione lineare non banale non può essere il vettore nullo. In altre parole, deve essere $Y_1 = Y_2 = \mathbf{0}$, e quindi $Y_3 = X = X_1 + X_2 + \mathbf{0}$, ossia V_{λ_3} è in somma diretta con $(V_{\lambda_1} \oplus V_{\lambda_2})$.

A questo punto, seguendo la stessa strada, possiamo convincerci che

$$(V_{\lambda_1} \oplus V_{\lambda_2} \oplus V_{\lambda_3}) \cap V_{\lambda_4} = \{\mathbf{0}_n\} \ ,$$

quindi i 4 autospazi sono in somma diretta

$$V_{\lambda_1} \oplus V_{\lambda_2} \oplus V_{\lambda_3} \oplus V_{\lambda_4}$$

e così via, fino a V_{λ_h} (formalmente, possiamo rendere più elegante la dimostrazione usando il principio di induzione). □

Una conseguenza immediata di questo Teorema è la seguente Proposizione.

Proposizione 6.10. *Sia A una matrice quadrata di ordine n. Possono esistere al più n autovalori distinti.*

Dimostrazione. È una conseguenza diretta del Teorema 6.9 e dell'Osservazione 6.10, che ci dice che la dimensione di ciascun autospazio è almeno 1. Se ci fossero più di n spazi in somma diretta, la somma delle loro dimensioni, che coincide con la dimensione dello spazio somma, sarebbe maggiore di n. Ma ciò è impossibile, poichè tutti insieme non possono dare uno spazio di dimensione maggiore di $\mathbb{R}^n$, di cui sono sottoinsiemi. □

Definizione 6.6. Sia $A \in \mathcal{M}_{\mathbb{K}}(n)$ una matrice quadrata di ordine n a entrate nel campo $\mathbb{R}$; sia L_A l'operatore lineare dello spazio vettoriale $\mathbb{R}^n$ su $\mathbb{R}$, associato alla matrice A usando la base canonica $\mathcal{B}_0 = \{\boldsymbol{e}_1, \dots \boldsymbol{e}_n\}$, con $\boldsymbol{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$, $\boldsymbol{e}_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}$, ecc.

Diciamo che L_A (o, direttamente A) ha *tutti gli autovalori in* $\mathbb{R}$ se il suo polinomio caratteristico $p_A(\lambda)$ ha n radici (non necessariamente distinte) nel campo (contando le molteplicità), e, quindi, si fattorizza in $\mathbb{R}$ in termini lineari in λ. Altrimenti detto, se possiamo scrivere

$$p_A(\lambda) = (\lambda_1 - \lambda)(\lambda_2 - \lambda) \dots (\lambda_n - \lambda) \,, \tag{6.40}$$

con tutti i λ_i in $\mathbb{R}$, per ogni $i = 1 \dots n$.

Teorema 6.11. *Sia $A \in \mathcal{M}_{\mathbb{K}}(n)$ una matrice quadrata di ordine n a entrate nel campo $\mathbb{R}$; se A tutti gli autovalori in $\mathbb{R}$, allora, detti $\lambda_1, \lambda_2, \dots \lambda_n$ tali autovalori, si ha*

$$\det(A) = \lambda_1 \cdot \lambda_2 \cdot \ \dots \ \cdot \lambda_n \,, \tag{6.41a}$$

$$\operatorname{tr}(A) = \lambda_1 + \lambda_2 + \dots + \lambda_n \,. \tag{6.41b}$$

Dimostrazione. Se A ha tutti gli autovalori nel campo, il polinomio caratteristico si può scrivere come

$$p_A(\lambda) = \det(A - \lambda I_n) = (\lambda_1 - \lambda)(\lambda_2 - \lambda) \dots (\lambda_n - \lambda) \,, \tag{6.42}$$

e, ponendo $\lambda = 0$ si trova

$$\det(A) = \lambda_1 \cdot \lambda_2 \cdot \dots \cdot \lambda_n \,.$$

Inoltre, convinciamoci della seconda equazione delle (6.41). Se $n = 2$,

$$p_A(\lambda) = (\lambda_1 - \lambda)(\lambda_2 - \lambda) = \lambda^2 - (\lambda_1 + \lambda_2)\lambda + \lambda_1\lambda_2 \,,$$

e, siccome il termine di grado $n-1=1$ del polinomio carateristico deve essere $(-\lambda)^{n-1}\operatorname{tr}(A)$, questo ci dice che $\operatorname{tr}(A)=(\lambda_1+\lambda_2)$.

Supponiamo, ora, vero l'asserto per $n-1$; Proviamo che vale per n. Il polinomio caratteristico si può calcolare in qualunque base: scegliamo una base $\mathcal{B}'$ che abbia come ultimo vettore $\boldsymbol{v}_n$ un autovettore dell'ultimo autovalore λ_n (ossia, completiamo la base lasciando $\boldsymbol{v}_n$ per ultimo anziché per primo): questa scelta migliora in chiarezza la notazione. Chiamiamo A' la matrice che rappresenta l'operatore nella base $\mathcal{B}'$, prima associato nella base originale alla matrice A. L'ultima colonna di A' ha tutte le entrate nulle, tranne l'ultima, che corrisponde a λ_n. Calcoliamo $p_{A'}(\lambda)=\det(A'-\lambda I_n)$ svilupando secondo l'ultima colonna. Risulterà

$$p_{A'}(\lambda)=(\lambda_n-\lambda)\det(B-\lambda I_{n-1}),$$

dove B è la sottomatrice di A ottenuta cancellando ultima riga e ultima colonna. Siccome B è di ordine $(n-1)\times(n-1)$, possiamo dire che per essa vale l'enunciato. Inoltre,

$$\det(B-\lambda I_{n-1})=(\lambda_1-\lambda)(\lambda_2-\lambda)\dots(\lambda_{n-1}-\lambda)\dots$$

riassumendo, possiamo scrivere

$$\begin{aligned}
p_A(\lambda)&=[(\lambda_1-\lambda)(\lambda_2-\lambda)\dots(\lambda_n-\lambda)]\\
&=[(\lambda_1-\lambda)(\lambda_2-\lambda)\dots(\lambda_{n-1}-\lambda)](\lambda_n-\lambda)\\
&=[(-\lambda)^{n-1}+(-\lambda)^{n-2}(\lambda_1+\lambda_2+\dots+\lambda_{n-1})+\dots](\lambda_n-\lambda)\\
&=(-\lambda)^{n-1}(-\lambda)+(-\lambda)^{n-1}\lambda_n+(-\lambda)(-\lambda)^{n-2}(\lambda_1+\lambda_2+\dots+\lambda_{n-1})\\
&\quad+\dots\\
&=(-\lambda)^n+(-\lambda)^{n-1}[(\lambda_1+\lambda_2+\dots+\lambda_{n-1})+\lambda_n]+\dots\\
&=(-\lambda)^n+(-\lambda)^{n-1}(\lambda_1+\lambda_2+\dots+\lambda_{n-1}+\lambda_n)+\dots,
\end{aligned}\tag{6.43}$$

dove abbiamo omesso di scrivere i termini di grado inferiore a quello necessario.

Ma il risultato ottenuto è proprio quello che volevamo, perché il termine di grado $n-1$ deve essere $(-\lambda)^{n-1}\operatorname{tr}(A)$. □

4. Molteplicità algebrica e geometrica

Abbiamo già osservato in precedenza quando abbiamo dato la definizione di diagonalizzabilità che, una volta cambiata la base su cui rappresentiamo l'operatore, un determinato autovalore può ricorrere più di una volta sulla diagonale. Inoltre, avendo ormai ben chiaro che gli autovalori sono tutte e sole le radici del polinomio caratteristico, ci rendiamo conto che una radice può non essere singola, ma doppia, tripla, ecc.

Infine, abbiamo anche capito che la dimensione di un autospazio non necessariamente è unitaria: possono esistere 2 o anche più vettori indipendenti che sono autovettori, e allora la dimensione dell'autospazio aumenta. A parte esempi banali (la matrice identità di ordine n ha tutto lo spazio come autospazio, con autovalore $\lambda=1$) vi sono casi meno ovvi.

Poniamo quindi le seguenti

Definizione 6.7. Viene detta *molteplicità algebrica* μ di un autovalore λ la molteplicità di λ come radice del polinomio caratteristico.

Definizione 6.8. Viene detta *molteplicità geometrica* ν di un autovalore λ la dimensione dell'autospazio associato, ossia $\nu = n - \text{rg}(A - \lambda I_n)$, se n è la dimensione dello spazio dell'operatore ed A la matrice che lo rappresenta sulla base scelta.

Consideriamo il seguente esempio:

Esempio 6.11. Consideriamo la seguente matrice di ordine 3:

$$A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & -1 \\ -1 & 1 & 2 \end{pmatrix} \tag{6.44}$$

e determiniamone gli autovalori e gli autospazi relativi.

Il polinomio caratteristico è

(6.45)

$$\begin{aligned} p_A(\lambda) &= \begin{vmatrix} -\lambda & 1 & 1 \\ 1 & -\lambda & -1 \\ -1 & 1 & 2-\lambda \end{vmatrix} \\ &= \begin{vmatrix} 1-\lambda & 1-\lambda & 0 \\ 1 & -\lambda & -1 \\ -1 & 1 & 2-\lambda \end{vmatrix} \qquad \text{sommando alla prima riga la seconda} \\ &= \begin{vmatrix} 1-\lambda & 1-\lambda & 0 \\ 0 & 1-\lambda & 1-\lambda \\ -1 & 1 & 2-\lambda \end{vmatrix} \qquad \text{sommando alla seconda riga la terza} \\ &= (1-\lambda)[(1-\lambda)(2-\lambda) - (1-\lambda)] - (1-\lambda)(1-\lambda) \\ &= (1-\lambda)^2[(2-\lambda) - 1] - (1-\lambda)^2 \\ &= (1-\lambda)^2[(2-\lambda) - 1 - 1] = -\lambda(1-\lambda)^2 . \end{aligned}$$

Quindi gli autovalori sono

- $\lambda_1 = 1$ con molteplicità algebrica $\mu_1 = 2$, e
- $\lambda_2 = 0$ con molteplicità algebrica $\mu_2 = 1$.

Per determinare la molteplicità geometrica, osserviamo che, per $\lambda = 1$

$$A - \lambda I_n = A - I_n = \begin{pmatrix} -1 & 1 & 1 \\ 1 & -1 & -1 \\ -1 & 1 & 1 \end{pmatrix} ;$$

si vede subito che la prima e la terza riga coincidono, e la seconda è l'opposto della prima; pertanto, $\text{rg}(A - I_n) = 1$, cioè $\nu_1 = 3 - 1 = 2$. L'autospazio V_1 è caratterizzato dall'equazione cartesiana

$$x - y - z = 0 .$$

Per $\lambda = 0$, si ha

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

la terza colonna è ottenibile sottraendo la seconda alla prima colonna. Si trova subito che $\mathrm{rg}(A) = 2$, cioè $\nu_1 = 3 - 2 = 1$. L'autospazio V_1 è caratterizzato dal sistema di equazioni cartesiane

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

Sono sempe uguali o sono scorrelate le due molteplicità? vediamo una maggiorazione con il seguente

Teorema 6.12. *La molteplicità geometrica di un autovalore λ_* non può superare la molteplicità geometrica.*

$$\nu(\lambda_*) \le \mu(\lambda_*) .$$

Dimostrazione. Supponiamo che l'autospazio V_λ associato a λ abbia dimensione ν. Consideriamo una base di V_λ:

$$\mathcal{B}_{V_{\lambda_*}} = \{\boldsymbol{u}_1, \dots \boldsymbol{u}_\nu\} ,$$

e completiamola con $n - \nu$ vettori indipendenti per avere una base dello spazio su cui agisce l'operatore:

$$\mathcal{B}_V = \left\{\boldsymbol{u}_1, \dots \boldsymbol{u}_\nu, \boldsymbol{u}'_{\nu+1} \dots \boldsymbol{u}'_n\right\} .$$

Su questa base, la matrice che rappresenta l'operatore ha una struttura a blocchi:

$$A = \left(\begin{array}{c|c} \lambda_* I_\nu & A'' \\ \hline 0_{n-\nu,\nu} & A' \end{array}\right) \tag{6.46}$$

dove la matrice A' ha formato $(n - \nu) \times (n - \nu)$.

Per il polinomio caratteristico, otteniamo

$$\det(A - \lambda I_n) = \left|\begin{array}{c|c} (\lambda_* - \lambda) I_\nu & A'' \\ \hline 0_{n-\nu,\nu} & A' - \lambda I_{n-\nu} \end{array}\right| . \tag{6.47}$$

Se svolgiamo i calcoli lungo le prime ν colonne di seguito, grazie alle entrate nulle che incontriamo nelle colonne, otteniamo:

$$(\lambda_* - \lambda)^\nu \det(A' - \lambda I) . \tag{6.48}$$

Ora, questo mostra che la molteplicità algebrica μ è almeno pari a ν. Infatti, se $\det(A' - \lambda_* I_n) \neq 0$, $\mu = \nu$; invece, se $\det(A' - \lambda_* I_n) = 0$, il termine $(\lambda_* - \lambda)$ si deve fattorizzare ancora almeno una volta (Proprietà 0.7 , Capitolo 0), e, quindi,

$$\mu(\lambda_*) \ge \nu(\lambda_*) .$$

□

Ma esistono matrici in cui la molteplicità algebrica e quella geometrica di un autovalore sono diverse? Vediamo un caso elementare:

Esempio 6.12. Consideriamo la matrice

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} . \tag{6.49}$$

Si verifica subito che gli autovalori di sono

- $\lambda_1 = 1$ con molteplicità algebrica $\mu_1 = 2$, e
- $\lambda_2 = 0$ con molteplicità algebrica $\mu_2 = 1$.

Infatti, la matrice è triangolare, quindi il calcolo del polinomio caratteristico è velocissimo:

$$\det(A - \lambda I) = \begin{vmatrix} 1-\lambda & 1 & 0 \\ 0 & 1-\lambda & 0 \\ 0 & 0 & -\lambda \end{vmatrix} = -\lambda(1-\lambda)^2 ; \tag{6.50}$$

in altre parole, se la matrice è triangolare (e ricordiamo che una matrice diagonale è un caso particolare di questo), i numeri sulla diagonale sono gli autovalori, e la molteplicità algebrica di ciascuno coincide con il numero di volte che lo troviamo sulla diagonale.

Però, con $\lambda = 1$ per la molteplicità geometrica otteniamo:

$$A - \lambda I = A - I = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} . \tag{6.51}$$

Vi sono due colonne indipendenti (le ultime due), e immediatamente abbiamo che

$$\nu(1) = 3 - \mathrm{rg}(A - I) = 3 - 2 = 1 \neq \mu(1) = 2 .$$

Ora come ora, siamo in grado di avere una condizione sufficiente per diagonalizzare una matrice:

Teorema 6.13. *Se un'operatore di uno spazio V $n-$dimensionale ha n autovalori distinti, è sicuramente diagonalizzabile.*

Dimostrazione. Semplicemente, prendiamo un sistema di n autovettori estraendo da ciascun autospazio un vettore. Siccome gli autospazi sono in somma diretta (Teorema 6.9), questo deve essere formato da vettori linearmente indipendenti (Lemma 6.7). Quindi formano una base di un sottospazio di dimensione pari a quella di V, che, pertanto, deve coincidere con V.

In altre parole, abbiamo trovato una base di autovettori, l'operatore è diagonalizzabile, per definizione (6.4) □

Riusciamo a dare un criterio di diagonalizzabilità più generale?

L'idea che sta sotto, in sostanza è abbastanza semplice: mi serve una base per tutto lo spazio di autovettori: siccome gli autospazi sono in somma diretta, l'unione di basi dei singoli autospazi mi darà la base dello spazio somma, e non posso fare molto meglio di questo: quindi, devo in tutto avere tanti vettori da "pescare" nelle basi indipendenti pari a n, dimensione dello spazio originario. Contando tutte le molteplicità, se va bene ho n autovalori. Se ne ho meno, non riesco a diagonalizzare; avrei la somma delle μ inferiore a n, e la somma delle dimensioni degli autospazi, ossia delle molteplicità geometriche, non può superare la somma delle μ, quindi non arriva ad n.

Sono costretto al "bilancio dettagliato", e non posso perdere un "pezzo" da nessuna parte. Ogni autospazio deve vere la dimensione "massima", ossia quella data dalla molteplicità algebrica. Diamo un nome ad un autovalore che "si comporta bene" in questo senso:

DEFINIZIONE 6.9. Sia λ un autovalore di un operatore L di uno spazio vettoriale n (o di una matrice A quadrata di ordine n). Diciamo che l'autovalore è *regolare* se le sue molteplicità algebrica μ e geometrica ν coincidono, ossia se

$$\mu(\lambda) = \nu(\lambda) = n - \mathrm{rg}(A - \lambda I_n)$$

Dobbiamo, come al solito, rendere formale questa intuizione. Lo facciamo con l'ultimo risultato fondamentale.

TEOREMA 6.14. *Sia A una matrice quadrata di ordine n, e sia L_A l'operatore associato ad A in uno spazio vettoriale $\mathcal{V}$ sul campo $\mathbb{R}$ (o, viceversa, sia L un operatore ed A la matrice associata nella base canonica $\mathcal{B}_0$, per comodità, ma va bene una base qualunque). Allora le seguenti affermazioni sono equivalenti:*

- *A è diagonalizzabile;*
- *A ha tutti gli autovalori in $\mathbb{R}$, e questi sono tutti regolari. Ossia, indicando con μ_k e ν_k la molteplicità algebrica e geometrica dell'autovalore λ_k, rispettivamente:*

  $$\mu(\lambda_k) = \nu(\lambda_k) \quad \forall\, k = 1 \dots h\,,$$

  dove h è il numero di autovalori distinti, e $\mu_1 + \dots \mu_h = n$, visto che devo avere tutti gli autovalori in $\mathbb{R}$
- *Con la notazione introdotta al punto di prima:*

  $$\nu_1 + \nu_2 + \dots \nu_h = n$$

DIMOSTRAZIONE. Per vedere che sono tutte equivalenti, basta vedere che ognuna implica quella dopo e l'ultima implica la prima. Cominciamo il tour de force finale: scopriremo che ormai è tutto in discesa.

Se A è diagonalizzabile, esiste una base di autovettori

$$\mathcal{B} = \{\boldsymbol{v}_1, \boldsymbol{v}_2 \dots \boldsymbol{v}_n\}\ ;$$

nella nuova base l'operatore è associato ad una matrice A', legata alla matrice A dalla matrice invertibile di cambio di base N, con $N \in \mathcal{GL}(n, \mathbb{R})$ (in generale, se il campo non è quello reale, sarà $N \in \mathcal{GL}(n, \mathbb{K})$). Ricordiamo che $N^i = [\boldsymbol{v}_i]_{\mathcal{B}_0}$, cioè l'$i-$esima colonna di N è la rappresentazione del vettore $i-$esimo $\boldsymbol{v}_i$ della nuova base nella base originaria. Ordiniamo i vettori in $\mathcal{B}'$ in modo che prima vi siano tutti quelli relativi al primo autovalore λ_1, poi quelli relativi al secondo

λ_2 e via così fino a λ_h;

$$A' = N^{-1}AN = \begin{pmatrix} \lambda_1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & \lambda_1 & 0 & \cdots & & \vdots \\ \vdots & & \ddots & & \vdots & \\ \vdots & & & \lambda_2 & & \vdots \\ \vdots & & & & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & 0 & \lambda_h \end{pmatrix} \tag{6.52}$$

come osservato nell'esempio 6.12, la matrice diagonale A', riporta μ_1 volte λ_1 nelle prime μ_1 righe, μ_2 volte λ_2 nelle successive μ_2 righe, ecc., fino alle ultime μ_h righe, dove riporta λ_h per μ_h volte. Così facendo, il polinomio caratteristico diventa sicuramente

$$p_{A'}(\lambda) = \det(A' - \lambda I_n) = (\lambda_1 - \lambda)^{\mu_1}(\lambda_2 - \lambda)^{\mu_2} \ldots (\lambda_h - \lambda)^{\mu_h}\,, \tag{6.53}$$

ogni radice ricorre con la corretta molteplicità algebrica, e **la matrice ha tutti gli autovalori in $\mathbb{R}$**.

Qual è la molteplicità geometrica ν_k dei diversi autovalori? dobbiamo studiare la matrice $A' - \lambda_k I_n$, poiché $\nu_k = \nu(\lambda_k) = n - \operatorname{rg}(A' - \lambda_k I_n)$. Per il primo autovalore avremo

$$A' - \lambda_1 I_n = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & \ddots & 0 & \cdots & & \vdots \\ \vdots & & 0 & & \vdots & \\ \vdots & & & \lambda_2 - \lambda_1 & & \vdots \\ \vdots & & & & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & 0 & \lambda_h - \lambda_1 \end{pmatrix}, \tag{6.54}$$

cioè le prime μ_1 righe sono nulle, e dalla successiva in poi troviamo sempre un numero sulla diagonale diverso da 0, perché gli altri autovalori sono differenti (abbiamo raggruppato apposta e contato le radici, per dire che $\lambda_i \neq \lambda_j$ se $i \neq j$).

Il rango della matrice è sicuramente $n - \mu_1$: ha μ_1 righe nulle, e fra le altre basta prendere la sottomatrice $(n-\mu_1) \times (n-\mu_1)$ con le ultime righe e colonne, che individua un minore non nullo, visto che è diagonale e nessun elemento rimasto sulla diagonale è 0 ($\lambda_1 \neq \lambda_2$, $\lambda_1 \neq \lambda_3$, ...). Allora,

$$\nu_1 = n - \operatorname{rg}(A' - \lambda_1 I_n) = n - (n - \mu_1) = \mu_1\,.$$

Il primo autovalore è regolare.

Stesso procedimento per il secondo, il terzo ecc.: basta individuare ogni volta la sottomatrice che si ottiene buttando via le righe corrispondenti ad ogni autovalore che si sta "indagando", e le colonne giuste, per avere una matrice diagonale con determinante non nullo, quindi per tutti i $k = 1 \ldots h$

$$\nu_k = n - \operatorname{rg}(A' - \lambda_k I_n) = n - (n - \mu_k) = \mu_k\,.$$

Tutti gli autovalori sono regolari (e ricordiamo che i risultati non dipendono dalla base, quindi valgono anche per A).

Vediamo ora la seconda implicazione. Supponiamo che A abbia tutti gli autovalori in $\mathbb{R}$, e questi siano tutti regolari.

Allora, banalmente,

$$\nu_1 + \nu_2 + \dots \nu_h = \mu_1 + \mu_2 + \dots \mu_h = n\,.$$

Infine, supponiamo che la somma delle molteplicità geometriche sia n. Sappiamo che per questo significa che la somma delle dimensioni degli autospazi è n. Siccome gli autospazi sono in somma diretta (Teorema 6.9), ciò significa che la dimensione dello spazio somma è n. Infatti, le dimensioni delle intersezioni "incapsulate" sono nulle (V_1 deve essere in somma diretta con $(V_2 \oplus \dots V_h)$, la dimensione della somma è $\nu_1 + \dim(V_2 \oplus \dots V_h)$; ma ora V_2 deve essere in somma diretta con $(V_3 \oplus \dots V_h)$, quindi la dimensione totale è $\nu_1 + \nu_2 + \dim(V_3 \oplus \dots V_h)$, ecc.). Allora, lo spazio somma è lo spazio $\mathcal{V}$ dell'operatore associato alla matrice.

Possiamo costruire una base di $\mathcal{V}$ prendendo

- ν_1 vettori che sono una base di V_{λ_1},
- ν_2 vettori che sono una base di V_{λ_2},
- ecc.
- ν_h vettori che sono una base di V_{λ_h}.

Questa base è quindi formata tutta da autovettori: la base di un autospazio è composta solo da vettori non nulli dell'autospazio, quindi, per definizione da autovettori.

E abbiamo finito! □

APPENDICE

5. Autovalori in campo complesso

Le stesse definizioni date nel Capitolo per gli spazi vettoriali reali si possono "duplicare" in campo complesso (avremmo potuto scrivere tutto tranquillamente in un campo $\mathbb{K}$ generico, in effetti).

Vediamo come cambia la Definizione 6.3:

DEFINIZIONE 6.10 (caso complesso). Sia $A \in \mathcal{M}_{\mathbb{C}}(n)$ una matrice quadrata di ordine n a entrate complesse; sia $L_A\colon \mathbb{C}^n \to \mathbb{C}^n$ l'operatore lineare associato alla matrice per lo spazio vettoriale *complesso* $\mathbb{C}^n$ nella base $\mathcal{B}$; sia infine $Z = [\boldsymbol{v}]_{\mathcal{B}}$ la rappresentazione di un vettore $\boldsymbol{v}$ di $\mathbb{C}^n$ nella stessa base. Diciamo che un Z che è un'*autovettore* di A relativo all'*autovalore* λ se esiste $\lambda \in \mathbb{C}$ tale che

$$AZ = \lambda Z\,, \tag{6.55}$$

con $Z \neq \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$.

OSSERVAZIONE 6.13. La definizione è sostanzialmente la stessa data prima (6.3); se adesso cerchiamo di risolvere il problema dell'esempio 6.7, perveniamo ancora all'equazione (6.25) (l'algebra nel campo complesso è la stessa che nel campo reale): solo che stavolta possiamo trovare soluzioni non banali: per $\lambda = \pm i$, l'equazione è soddisfatta!

Allora, esistono soluzioni non banali: per $\lambda = i$, $y = -\lambda x = -ix$, e quindi gli autovettori avranno la forma

$$\begin{pmatrix} x \\ y \end{pmatrix} = \alpha \begin{pmatrix} 1 \\ -i \end{pmatrix}\,,$$

con $\alpha \in \mathbb{C}$.
Analogamente, per $\lambda = -i$, $y = -\lambda x = ix$, e troviamo gli autovettori

$$\begin{pmatrix} x \\ y \end{pmatrix} = \beta \begin{pmatrix} 1 \\ i \end{pmatrix}$$

con $\beta \in \mathbb{C}$.

Analogamente a quanto fatto in precedenza, diremo che L_A (o, direttamente A) ha *tutti gli autovalori in* $\mathbb{C}$ se il suo polinomio caratteristico $p_A(\lambda)$ ha n radici (non necessariamente distinte) nel campo (contando le molteplicità), e, quindi, si fattorizza in $\mathbb{C}$ in termini lineari in λ. Altrimenti detto, se possiamo scrivere

$$p_A(\lambda) = (\lambda_1 - \lambda)(\lambda_2 - \lambda)\dots(\lambda_n - \lambda)\,, \tag{6.56}$$

con tutti i λ_i in $\mathbb{C}$, per ogni $i = 1 \dots n$.

Conseguenza immediata del Teorema Fondamentale dell'Algebra (0.8) è che se il campo è quello complesso, allora, l'operatore ha tutti gli autovalori nel campo. Quindi, risulta vero il seguente

Corollario 6.15. *Sia* $A \in \mathcal{M}_{\mathbb{C}}(n)$ *una matrice quadrata complessa di ordine* n*. Sia* L_A *l'operatore lineare di* $\mathbb{C}^n$ *associato alla matrice* A*; ossia, scelta la base canonica* $\mathcal{B}_0$ *di* $\mathbb{C}^n$*, e detti* $X = [\boldsymbol{u}]_{\mathcal{B}_0}$ *ed* $Y = [\boldsymbol{v}]_{\mathcal{B}_0}$ *le rappresentazioni di due vettori* $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{C}^n$*,* L_A *è definita in modo che*

$$L_A(\boldsymbol{u}) = \boldsymbol{v} \iff Y = A\,X\,.$$

Allora A *ha tutti gli autovalori in* $\mathbb{C}$*.*

CAPITOLO 7

Struttura metrica in $\mathbb{R}^n$

1. Prodotto scalare

Vogliamo ora estendere allo spazio vettoriale $\mathbb{R}^n$, delle n-uple ordinate di numeri reali, le proprietà metriche, come distanze ed angoli, che abbiamo studiato nel piano e nello spazio. Abbiamo visto che lo strumento essenziale per trattare questioni metriche nello spazio è il prodotto scalare. Ricordiamo che dati nello spazio due vettori $\overrightarrow{u}$ e $\overrightarrow{v}$, applicati nello stesso punto O, il loro prodotto scalare è il seguente numero reale:

$$\langle \overrightarrow{u}, \overrightarrow{v} \rangle = \| \overrightarrow{u} \| \, \| \overrightarrow{v} \| \cos\vartheta,$$

dove ϑ è l'angolo convesso formato dai vettori $\overrightarrow{u}$ e $\overrightarrow{v}$.
Abbiamo osservato che il prodotto scalare è strettamante legato alla proiezione ortogonale di un vettore lungo una retta. Abbiamo inoltre verificato che gode delle seguenti proprietà:

(1) **simmetria**: $\langle \overrightarrow{u}, \overrightarrow{v} \rangle = \langle \overrightarrow{v}, \overrightarrow{u} \rangle, \ \forall \ \overrightarrow{u}, \overrightarrow{v} \in \mathbb{E}^3_O$;
(2) **bilinearità** :
  (a) linearità rispetto alla prima variabile:

$$\langle a\overrightarrow{u}_1 + b\overrightarrow{u}_2, \overrightarrow{v} \rangle = a\langle \overrightarrow{u}_1, \overrightarrow{v} \rangle + \ b\langle \overrightarrow{u}_2, \overrightarrow{v} \rangle, \quad \forall \overrightarrow{u}_1, \overrightarrow{u}_2, \overrightarrow{v} \in \mathbb{E}^3_O, \ \forall a, b \in \mathbb{R};$$

  (b) linearità rispetto alla seconda variabile:

$$\langle \overrightarrow{u}, a\overrightarrow{v}_1 + b\overrightarrow{v}_2 \rangle = a\langle \overrightarrow{u}, \overrightarrow{v}_1 \rangle + \ b\langle \overrightarrow{u}, \overrightarrow{v}_2 \rangle, \quad \forall \overrightarrow{u}, \overrightarrow{v}_1, \overrightarrow{v}_2 \in \mathbb{E}^3_O, \quad \forall a, b \in \mathbb{R};$$

(3) **positività**:

$$\langle \overrightarrow{u}, \overrightarrow{u} \rangle \geq 0 \quad \forall \overrightarrow{u} \in \mathbb{E}^3_O,$$

$$\langle \overrightarrow{u}, \overrightarrow{u} \rangle = 0 \iff \overrightarrow{u} = \overrightarrow{0}.$$

Fissiamo ora nello spazio un sistema di riferimento cartesiano ortogonale monometrico $\mathcal{R}(O, \hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}, \hat{\boldsymbol{k}})$ di origine O, ciò equivale a fissare nello spazio $\mathbb{E}^3_O$ la base *ortonormale* $\mathcal{B} = \left\{ \hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}, \hat{\boldsymbol{k}} \right\}$ data dai versori degli assi. Possiamo associare ad ogni vettore $\overrightarrow{v}$ applicato in O il vettore di $\mathbb{R}^3$ delle coordinate di $\overrightarrow{v}$ rispetto alla base $\mathcal{B}$:

$$[\overrightarrow{v}]_{\mathcal{B}} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \iff \overrightarrow{v} = x\hat{\boldsymbol{\imath}} + y\hat{\boldsymbol{\jmath}} + z\hat{\boldsymbol{k}}.$$

Siano $\overrightarrow{u}$ e $\overrightarrow{v}$ due vettori applicati in O, supponiamo di conoscere i vettori delle loro coordinate rispetto alla base $\mathcal{B}$:

$$[\overrightarrow{u}]_{\mathcal{B}} = \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} \quad [\overrightarrow{v}]_{\mathcal{B}} = \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix},$$

applicando le proprietà di bilinearità, possiamo ricavare il prodotto scalare di $\overrightarrow{u}$ e $\overrightarrow{v}$ in funzione delle loro coordinate:

$$\langle \overrightarrow{u}, \overrightarrow{v} \rangle = x_1x_2 + y_1y_2 + z_1z_2. \tag{7.1}$$

Ricaviamo quindi la formula che consente di calcolare il modulo di un vettore $\overrightarrow{v}$ in funzione delle sue coordinate:

$$\|\overrightarrow{v}\| = \sqrt{\langle \overrightarrow{v}, \overrightarrow{v} \rangle} = \sqrt{x^2 + y^2 + z^2},$$

e la ben nota formula della distanza euclidea tra due punti A e B, in funzione delle loro coordinate

$$d(A, B) = \|\overrightarrow{AB}\| = \sqrt{(x_B - x_A)^2 + (y_B - y_A)^2 + (z_B - z_A)^2}.$$

Nel seguito, identifichiamo lo spazio vettoriale $\mathbb{R}^n$ con lo spazio vettoriale $\mathcal{M}_{\mathbb{R}}(n, 1)$ dei vettori colonna di tipo $n, 1$. Se $X \in \mathbb{R}^n$, la matrice trasposta di X è il vettore riga X^T. Introduciamo ora il prodotto scalare in $\mathbb{R}^n$ generalizzando la formula 7.1:

Definizione 7.1. Dati due vettori $X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ e $Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$ in $\mathbb{R}^n$, chiamiamo **prodotto scalare standard** dei vettori X e Y il seguente numero reale

$$\langle X, Y \rangle = x_1y_1 + x_2y_2 + ... + x_ny_n = X^T \cdot Y.$$

Si noti che la scrittura compatta $X^T \cdot Y$ indica il prodotto riga per colonna del vettore riga X^T per il vettore colonna Y.

Esempio 7.1. Calcoliamo il prodotto scalare tra i seguenti vettori di $\mathbb{R}^4$:

$$X = \begin{pmatrix} 1 \\ -1 \\ 2 \\ 1 \end{pmatrix} \quad Y = \begin{pmatrix} -1 \\ 0 \\ 3 \\ -1 \end{pmatrix}.$$

Abbiamo:

$$\langle X, Y\rangle = X^T \cdot Y = \begin{pmatrix} 1 & -1 & 2 & 1 \end{pmatrix} \cdot \begin{pmatrix} -1 \\ 0 \\ 3 \\ -1 \end{pmatrix} = -1 + 6 - 1 = 4.$$

Verifichiamo ora che il prodotto scalare che abbiamo ora definito gode delle stesse proprietà del prodotto scalare di $\mathbb{R}^3$.

Proposizione 7.1. *Il prodotto scalare standard di $\mathbb{R}^n$ gode delle seguenti proprietà:*

(1) ***Proprietà di simmetria:***

$$\langle X, Y\rangle = \langle Y, X\rangle \quad \forall X, Y \in \mathbb{R}^n;$$

(2) ***Proprietà di bilinearità:***

$$\langle aX_1 + bX_2, Y\rangle = a\langle X_1, Y\rangle + \ b\langle X_2, Y\rangle, \quad \forall X_1, X_2, Y \in \mathbb{R}^n, \forall a, b \in \mathbb{R};$$
$$\langle X, aY_1 + bY_2\rangle = a\langle X, Y_1\rangle + \ b\langle X, Y_2\rangle, \quad \forall X, Y_1, Y_2 \in \mathbb{R}^n, \forall a, b \in \mathbb{R};$$

(3) ***Proprietà di positività:***

$$\langle X, X\rangle \geq 0 \quad \forall X \in \mathbb{R}^n,$$
$$\langle X, X\rangle = 0 \iff X = \mathbf{0}_n.$$

Dimostrazione. Verifichiamo le proprietà.
(1) Siano $X, Y \in \mathbb{R}^n$, osserviamo che risulta:

$$\langle X, Y\rangle = \ X^T \cdot Y = x_1y_1 + x_2y_2 + ... + x_ny_n,$$
$$\langle Y, X\rangle = \ Y^T \cdot X = y_1x_1 + y_2x_2 + ... + y_nx_n,$$

poiché $x_iy_i = y_ix_i$ per ogni $i = 1, ., n$, le due espressioni scritte coincidono.
(2) Per la proprietà di simmetria basta provare la prima. Siano X_1, X_2 e Y vettori di $\mathbb{R}^n$, osserviamo che dalle proprietà del prodotto matriciale risulta:

$$\langle aX_1 + bX_2, Y\rangle = \ (aX_1 + bX_2)^T \cdot Y = (a\ X_1^T + b\ X_2^T) \cdot Y =$$
$$= a\ X_1^T \cdot Y + b\ X_2^T \cdot Y = a\langle X_1, Y\rangle + b\langle X_2, Y\rangle.$$

(3) Sia $X \in \mathbb{R}^n$, si ha:

$$\langle X, X\rangle = \ X^T \cdot X = x_1^2 + x_2^2 + ... + x_n^2 \geq 0,$$

inoltre $\langle X, X\rangle = 0$ se e solo se $x_1 = x_2 = ... = x_n = 0$, cioè X è il vettore nullo di $\mathbb{R}^n$. □

Alla luce di questa nuova operazione che abbiamo definito possiamo rileggere il modo di calcolare gli elementi della matrice prodotto.

Osservazione 7.2. Siano $A \in \mathcal{M}_{\mathbb{R}}(m, p)$ e $B \in \mathcal{M}_{\mathbb{R}}(p, n)$. L'elemento della $i-$esima riga e della $j-$esima colonna della matrice prodotto AB è dato dal prodotto scalare standard fra la trasposta del vettore riga A_i ed il vettore colonna B^j. Infatti, tale elemento è il "vettore" 1×1 ottenuto moltiplicando il vettore riga A_i, di ordine $1 \times p$ per il vettore colonna B^j, di ordine $p \times 1$:

$$(AB)_{i,j} = a_{i,1}b_{1,j} + a_{i,2}b_{2,j} + \cdots + a_{i,p}b_{p,j} = A_iB^j = \langle {A_i}^{\mathrm{T}}, B^j\rangle\,. \tag{7.2}$$

Estendiamo ora ai vettori di $\mathbb{R}^n$ la nozione di lunghezza di un vettore:

Definizione 7.2. Fissato in $\mathbb{R}^n$ il prodotto scalare standard,

(1) la **norma** di un vettore $X \in \mathbb{R}^n$ è il seguente numero reale non negativo:
$$\|X\| = \sqrt{\langle X, X\rangle} = \sqrt{x_1^2 + x_2^2 + ... + x_n^2};$$
(2) un vettore $X \in \mathbb{R}^n$ è detto **versore** se $\|X\| = 1$.

Esempio 7.3. Calcoliamo la norma dei seguenti vettori di $\mathbb{R}^5$:
$$X = \begin{pmatrix} 1 \\ -1 \\ 2 \\ 1 \\ 1 \end{pmatrix} \quad Y = \begin{pmatrix} -1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$
Abbiamo:
$$\begin{aligned} \|X\| &= \sqrt{\langle X, X\rangle} = \sqrt{x_1^2 + x_2^2 + ... + x_n^2} \\ &= \sqrt{1^2 + (-1)^2 + 2^2 + 1^2 + 1^2} = \sqrt{8} = 2\sqrt{2}; \end{aligned}$$
$$\begin{aligned} \|Y\| &= \sqrt{\langle Y, Y\rangle} = \sqrt{y_1^2 + y_2^2 + ... + y_n^2} \\ &= \sqrt{(-1)^2 + 0^2 + 0^2 + 0^2 + 0^2} = \sqrt{1} = 1. \end{aligned}$$
Il vettore Y è un versore di $\mathbb{R}^5$.

Osservazione 7.4.
- Osserviamo che, dalla proprietà di positività, l'unico vettore $X \in \mathbb{R}^n$ con norma nulla è il vettore nullo di $\mathbb{R}^n$:
$$\|X\| = 0 \iff \langle X, X\rangle = 0 \iff X = \mathbf{0}_n.$$
- Sia $X \in \mathbb{R}^n$ un vettore non nullo, osserviamo che per la linearità del prodotto scalare si ha:
$$\|\alpha X\| = \sqrt{\langle \alpha X, \alpha X\rangle} = \sqrt{\alpha^2\langle X, X\rangle} = |\alpha|\sqrt{\langle X, X\rangle} = |\alpha|.\|X\|,$$
in particolare se $\alpha > 0$ allora risulta $\|\alpha X\| = \alpha.\|X\|$.
- Sia $X \in \mathbb{R}^n$ un vettore non nullo, osserviamo che il vettore $Y = \frac{1}{\|X\|}X$ è un versore. Infatti, applicando la proprietà precedente si ha:
$$\|Y\| = \frac{1}{\|X\|}\|X\| = 1.$$

Le seguenti proprietà della norma sono fondamentali:

PROPOSIZIONE 7.2. *Siano X e Y due vettori di $\mathbb{R}^n$, valgono le seguenti proprietà:*

(1) ***Disuguaglianza di Cauchy-Schwarz:***

$$\langle X, Y\rangle^2 \leq \|X\|^2\|Y\|^2,$$

e vale l'uguale se e solo se i vettori X e Y sono linearmente dipendenti;

(2) ***Disuguaglianza triangolare:***

$$|\ \|X\| - \|Y\|\ | \leq \|X+Y\| \leq \|X\| + \|Y\|.$$

DIMOSTRAZIONE.
(1) Osserviamo che la disuguaglianza è verificata se Y è il vettore nullo di $\mathbb{R}^n$: infatti entrambi i membri della disuguaglianza sono nulli:

$$\langle X, \mathbf{0}_n\rangle = \ X^T \cdot \mathbf{0}_n = 0, \quad \|\mathbf{0}_n\| = 0.$$

Siano $Y \in \mathbb{R}^n$ un vettore non nullo e $t \in \mathbb{R}$, consideriamo il vettore $X+tY$ e calcoliamo $\langle X+tY, X+tY\rangle$. Per la bilinearità del prodotto scalare si ha:

$$\begin{aligned}\langle X+tY, X+tY\rangle &= \langle X, X\rangle + \langle X, tY\rangle + \langle tY, X\rangle + \langle tY, tY\rangle = \\ &= \langle X, X\rangle + t\langle X, Y\rangle + t\langle Y, X\rangle + t^2\langle Y, Y\rangle\end{aligned}$$

poiché $\langle X, Y\rangle = \langle Y, X\rangle$, si ha:

$$\langle X+tY, X+tY\rangle = \langle X, X\rangle + 2t\langle X, Y\rangle + t^2\langle Y, Y\rangle. \tag{7.3}$$

Per la positività del prodotto scalare, risulta

$$\langle X+tY, X+tY\rangle \geq 0, \quad \forall t \in \mathbb{R},$$

possiamo quindi concludere che il trinomio di secondo grado in t in 7.3 soddisfa la seguente disequazione

$$\langle X, X\rangle + 2t\langle X, Y\rangle + t^2\langle Y, Y\rangle \geq 0, \quad \forall t \in \mathbb{R}.$$

Quindi il discriminante Δ del trinomio risulta necessariamente non positivo:

$$\Delta = \langle X, Y\rangle^2 - \langle X, X\rangle\langle Y, Y\rangle \leq 0,$$

sostituendo $\langle X, X\rangle = \|X\|^2$ e $\langle Y, Y\rangle = \|Y\|^2$ otteniamo la disuguaglianza (1).
Osserviamo infine che vale l'uguaglianza nella formula se e solo $\Delta = 0$, cioè il trinomio 7.3 ha due radici reali coincidenti $t_1 = t_2$. Per tale valore, risulta $\langle X+t_1Y, X+t_1Y\rangle = 0$, per la positività del prodotto scalare il vettore $X + t_1Y$ è nullo: i vettori X e Y sono quindi linearmente dipendenti.

(2) Siano $X, Y \in \mathbb{R}^n$, ponendo $t = 1$ nell'espressione 7.3 otteniamo:

$$\|X+Y\|^2 = \|X\|^2 + 2\langle X, Y\rangle + \|Y\|^2.$$

Applichiamo la disuguaglianza di Cauchy,

$$-\|X\| \cdot \|Y\| \leq \langle X, Y\rangle \leq \|X\| \cdot \|Y\|,$$

otteniamo quindi:

$$\|X\|^2 - 2\|X\| \cdot \|Y\| + \|Y\|^2 \leq \|X+Y\|^2 \leq \|X\|^2 + 2\|X\| \cdot \|Y\| + \|Y\|^2,$$

cioè

$$(\|X\| - \|Y\|)^2 \leq \|X+Y\|^2 \leq (\|X\| + \|Y\|)^2,$$

da cui ricaviamo la disuguaglianza triangolare

$$|\ \|X\| - \|Y\|\ | \leq \|X+Y\| \leq \|X\| + \|Y\|.$$

□

La disuguaglianza triangolare estende il classico risultato della geometria euclidea il quale afferma che in un triangolo ogni lato è minore della somma degli altri due e maggiore della loro differenza.

Osserviamo che, se X e Y sono vettori non nulli di $\mathbb{R}^n$, la disuguaglianza di Cauchy può essere anche scritta nella seguente forma:

$$-1 \leq \frac{\langle X, Y \rangle}{\|X\| \cdot \|Y\|} \leq 1, \tag{7.4}$$

esiste quindi un unico angolo ϑ, $0 \leq \vartheta \leq \pi$, che soddisfa la relazione seguente:

$$\cos\vartheta = \frac{\langle X, Y \rangle}{\|X\| \cdot \|Y\|}. \tag{7.5}$$

Estendiamo ora ai vettori di $\mathbb{R}^n$ le nozioni di angolo formato da due vettori e di perpendicolarità tra vettori:

Definizione 7.3. Fissato in $\mathbb{R}^n$ il prodotto scalare standard,

(1) l'angolo convesso ϑ definito dalla 7.5 è detto **angolo formato dai vettori** X e Y;
(2) due vettori X e Y di $\mathbb{R}^n$ si dicono **ortogonali**, in simboli $X \perp Y$, se e solo se il loro prodotto scalare è nullo:
$$X \perp Y \iff \langle X, Y \rangle = 0.$$

Esempio 7.5. Verifichiamo che i seguenti vettori di $\mathbb{R}^4$ sono perpendicolari:

$$X = \begin{pmatrix} 1 \\ -1 \\ 2 \\ 1 \end{pmatrix} \quad Y = \begin{pmatrix} -1 \\ 0 \\ 1 \\ -1 \end{pmatrix}.$$

Calcoliamo il loro prodotto scalare:

$$\langle X, Y \rangle = X^T \cdot Y = \begin{pmatrix} 1 & -1 & 2 & 1 \end{pmatrix} \cdot \begin{pmatrix} -1 \\ 0 \\ 1 \\ -1 \end{pmatrix} = -1 + 2 - 1 = 0,$$

poiché risulta $\langle X, Y \rangle = 0$ i vettori sono perpendicolari.

Osservazione 7.6.
Osserviamo che il vettore nullo di $\mathbb{R}^n$ è perpendicolare a tutti i vettori X di $\mathbb{R}^n$:

$$\langle X, \mathbf{0}_n \rangle = X^T.\mathbf{0}_n = x_1.0 + x_2.0 + .. + x_n.0 = 0, \quad \forall X \in \mathbb{R}^n.$$

Verifichiamo che è l'unico vettore di $\mathbb{R}^n$ perpendicolare a tutti i vettori. Supponiamo che $Y \in \mathbb{R}^n$ sia un vettore che verifica la proprietà suddetta:

$$\langle X, Y \rangle = 0 \quad \forall X \in \mathbb{R}^n,$$

in particolare se $X = Y$ risulta $\langle Y, Y \rangle = 0$, che implica $Y = \mathbf{0}_n$.

Osservazione 7.7.
Osserviamo, inoltre, che per costruire la geometria euclidea in $\mathbb{R}^n$ abbiamo usato solo le proprietà di simmetria, bilinearità e positività del prodotto scalare standard, quindi è possibile introdurre una struttura metrica in $\mathbb{R}^n$ a partire da una qualsiasi applicazione

$$\langle \cdot, \cdot \rangle \colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$$

che soddisfi le proprietà di simmetria, bilinearità e positività.

2. Basi ortogonali

Nello studio della geometria dello spazio abbiamo visto come sia vantaggioso fissare un sistema di riferimento cartesiano ortogonale $\mathcal{R}(O, \hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}, \hat{\boldsymbol{k}})$ di origine O, ciò equivale a fissare nello spazio $\mathbb{E}^3_O$ la base *ortonormale* $\mathcal{B} = \left\{\hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}, \hat{\boldsymbol{k}}\right\}$ data dai versori degli assi. Ogni vettore $\overrightarrow{v} = \overrightarrow{OP} \in \mathbb{E}^3_O$ si scrive in uno e un solo modo nella forma seguente:

$$\overrightarrow{v} = x\hat{\boldsymbol{\imath}} + y\hat{\boldsymbol{\jmath}} + z\hat{\boldsymbol{k}},$$

dove gli scalari x, y, e z si ottengono dalle proiezioni ortogonali del punto P sugli assi coordinati. Più precisamente si ha che

$$x = \langle \overrightarrow{v}, \hat{\boldsymbol{\imath}} \rangle \quad y = \langle \overrightarrow{v}, \hat{\boldsymbol{\jmath}} \rangle \quad z = \langle \overrightarrow{v}, \hat{\boldsymbol{k}} \rangle. \tag{7.6}$$

Ci proponiamo ora di estendere allo spazio $\mathbb{R}^n$ tali proprietà.

Definizione 7.4. Fissato in $\mathbb{R}^n$ il prodotto scalare standard,

(1) un insieme di vettori non nulli $\{Y_1, ..., Y_k\}$, $k \geq 2$, è detto **ortogonale** se i vettori sono a due a due perpendicolari, cioè

$$\langle Y_i, Y_j \rangle = 0 \quad \forall i \neq j;$$

(2) un insieme di vettori non nulli $\{Y_1, ..., Y_k\}$, $k \geq 2$, è detto **ortonormale** se è formato da versori a due a due perpendicolari, cioè

$$\langle Y_i, Y_j \rangle = 0 \quad \forall i \neq j, \quad \langle Y_i, Y_i \rangle = 1, \quad \forall i = 1, .., k.$$

Osservazione 7.8. Osserviamo che è molto semplice trasformare un sistema ortogonale in un sistema ortonormale, basta infatti *normalizzare* ciascun vettore. Più precisamente, siano $\{Y_1, ..., Y_k\}$ un sistema ortogonale di vettori di $\mathbb{R}^n$, allora $\forall i = 1, .., k$ il vettore $\frac{1}{\|Y_i\|} Y_i$ è un versore, inoltre per la bilinearità del prodotto scalare risulta

$$\langle \frac{1}{\|Y_i\|} Y_i, \frac{1}{\|Y_j\|} Y_j \rangle = \frac{1}{\|Y_i\|} \frac{1}{\|Y_j\|} \langle Y_i, Y_j \rangle = 0, \quad \forall i \neq j$$

poiché $Y_i \perp Y_j$, essendo il sistema ortogonale. Possiamo quindi concludere che i vettori

$$\left\{ \frac{1}{\|Y_1\|} Y_1, ..., \frac{1}{\|Y_k\|} Y_k \right\}$$

sono un sistema ortonormale.

Esempio 7.9.

(1) I vettori $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3, \boldsymbol{e}_4\}$ della base canonica di $\mathbb{R}^4$ sono un insieme ortonormale di $\mathbb{R}^4$. Infatti, risulta:

$$\langle \boldsymbol{e}_1, \boldsymbol{e}_2 \rangle = \langle \boldsymbol{e}_1, \boldsymbol{e}_3 \rangle = \langle \boldsymbol{e}_1, \boldsymbol{e}_4 \rangle = 0,$$
$$\langle \boldsymbol{e}_2, \boldsymbol{e}_3 \rangle = \langle \boldsymbol{e}_2, \boldsymbol{e}_4 \rangle = \langle \boldsymbol{e}_3, \boldsymbol{e}_4 \rangle = 0,$$
$$\langle \boldsymbol{e}_1, \boldsymbol{e}_1 \rangle = \langle \boldsymbol{e}_2, \boldsymbol{e}_2 \rangle = \langle \boldsymbol{e}_3, \boldsymbol{e}_3 \rangle = \langle \boldsymbol{e}_4, \boldsymbol{e}_4 \rangle = 1.$$

(2) I vettori $\{\boldsymbol{e}_1, \boldsymbol{e}_2, ..., \boldsymbol{e}_n\}$ della base canonica di $\mathbb{R}^n$ sono un sistema ortonormale di $\mathbb{R}^n$, infatti si ha:

$$\langle \boldsymbol{e}_i, \boldsymbol{e}_j \rangle = 0 \qquad i \neq j \qquad \langle \boldsymbol{e}_i, \boldsymbol{e}_i \rangle = 1, \quad \forall i = 1, 2, .., n.$$

(3) I vettori

$$\boldsymbol{u}_1 = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}, \quad \boldsymbol{u}_2 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \quad \boldsymbol{u}_3 = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$$

sono un insieme *ortogonale* in $\mathbb{R}^3$. Per ottenere un insieme ortonormale, basta normalizzare ogni vettore dividendolo per la sua norma; otteniamo il sistema di vettori formato da:

$$\boldsymbol{w}_1 = \begin{pmatrix} \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{pmatrix}, \quad \boldsymbol{w}_2 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ -\frac{1}{\sqrt{2}} \end{pmatrix}, \quad \boldsymbol{w}_3 = \begin{pmatrix} \frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \end{pmatrix}.$$

Ricordiamo che tre vettori nello spazio, a due e due ortogonali, sono linearmente indipendenti. Questa proprietà si estende ai sistemi ortogonali di $\mathbb{R}^n$:

Proprietà 7.3. *Se $\{Y_1, \dots, Y_k\}$ sono un insieme ortogonale di $\mathbb{R}^n$, con $k \geq 2$, allora i vettori sono linearmente indipendenti.*

Dimostrazione. Supponiamo che esista una combinazione lineare dei vettori $Y_1, .., Y_k$ che dà il vettore nullo:

$$a_1 Y_1 + a_2 Y + ... + a_k Y_k = \mathbf{0}_n.$$

Osserviamo che per ogni $i = 1, .., k$ risulta

$$\langle a_1 Y_1 + a_2 Y + ... + a_k Y_k, Y_i \rangle = \langle \mathbf{0}_n, Y_i \rangle = 0,$$

poiché il vettore nullo è perpendicolare a tutti i vettori. Inoltre, per le proprietà di bilinearità del prodotto scalare, per ogni $i = 1, ..., k$, risulta:

$$\langle a_1 Y_1 + ... + a_k Y_k, Y_i \rangle = a_1 \langle Y_1, Y_i \rangle + ... + a_i \langle Y_i, Y_i \rangle + ... + a_k \langle Y_k, Y_i \rangle,$$

ma poiché $Y_j \perp Y_i$, per ogni $j \neq i$, si ha $\langle Y_j, Y_i \rangle = 0$ e quindi risulta:

$$0 = \langle a_1 Y_1 + ... + a_k Y_k, Y_i \rangle = a_i \langle Y_i, Y_i \rangle.$$

Poiché $Y_i \neq \mathbf{0}_n$, si ha $\langle Y_i, Y_i \rangle \neq 0$, quindi risulta necessariamente $a_i = 0$, $\forall i = 1, .., k$. Otteniamo che tutti gli scalari $a_1, .., a_k$ sono nulli e possiamo concludere che i vettori sono linearmente indipendenti. □

Osservazione 7.10. Un sistema ortogonale di n vettori di $\mathbb{R}^n$ è una base di $\mathbb{R}^n$!!!

Estendiamo infine la nozione di base ortogonale e ortonormale ai sottospazi di $\mathbb{R}^n$:

Definizione 7.5. Fissato in $\mathbb{R}^n$ il prodotto scalare standard, sia $V \subset \mathbb{R}^n$ un sottospazio non nullo.

(1) una **base ortogonale** di V è una base di V che è un sistema ortogonale di $\mathbb{R}^n$;
(2) una **base ortonormale** di V è una base di V che è un sistema ortonormale di $\mathbb{R}^n$.

Osservazione 7.11. Osserviamo che la base canonica $\{\boldsymbol{e}_1, .., \boldsymbol{e}_n\}$ di $\mathbb{R}^n$ è una base ortonormale di $\mathbb{R}^n$, rispetto al prodotto scalare standard.

Esempio 7.12. Fissato in $\mathbb{R}^4$ il prodotto scalare standard, consideriamo il sottospazio $V \subset \mathbb{R}^4$ di equazioni:

$$x - y = z + t = 0.$$

Sia $Y \in \mathbb{R}^4$, si ha:

$$Y = \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \in V \iff y = x \quad t = -z.$$

Risulta quindi

$$V = \left\{ Y = \begin{pmatrix} x \\ x \\ z \\ -z \end{pmatrix} \mid x \in \mathbb{R}, z \in \mathbb{R} \right\}.$$

Osserviamo che ogni vettore $Y \in V$ può essere scritto nella forma:

$$Y = x \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + z \begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix}, \quad x \in \mathbb{R}, \quad z \in \mathbb{R},$$

per cui possiamo concludere che

$$V = \text{Span}(Y_1, Y_2), \quad Y_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad Y_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix}.$$

Verifichiamo che $\{Y_1, Y_2\}$ sono una base ortogonale per V. Poiché risulta

$$\langle Y_1, Y_2 \rangle = \ {Y_1}^T . Y_2 = \begin{pmatrix} 1 & 1 & 0 & 0 \end{pmatrix} . \begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix} = 1.0 + 1.0 + 0.1 + 0.(-1) = 0,$$

i vettori sono perpendicolari, quindi sono una base ortogonale di V. Calcoliamo la norma dei vettori della base:

$$\|Y_1\| = \sqrt{1^2 + 1^2 + 0^2 + 0^2} = \sqrt{2},$$

$$\|Y_2\| = \sqrt{0^2 + 0^2 + 1^2 + (-1)^2} = \sqrt{2}.$$

Una base ortonormale di V si ottiene normalizzando i vettori Y_1 e Y_2:

$$\left\{ \frac{1}{\sqrt{2}} Y_1, \frac{1}{\sqrt{2}} Y_2 \right\}.$$

Fissato il prodotto scalare standard in $\mathbb{R}^n$, sia $V \subset \mathbb{R}^n$ un sottospazio non nullo. Vediamo ora quali sono i vantaggi nel fissare una base ortogonale o ortonormale in V. Sia $\mathcal{B}_V = \{Y_1, Y_2, .., Y_h\}$ una base ortogonale di V. Ci proponiamo di determinare le coordinate di un vettore $Y \in V$, rispetto alla base $\mathcal{B}_V$. Sappiamo che Y ammette un 'unica scrittura come combinazione lineare dei vettori $Y_1, .., Y_h$:

$$Y = a_1 Y_1 + a_2 Y_2 + ... + a_h Y_h,$$

con $a_i \in \mathbb{R}$, $i = 1, .., h$. Per ogni $i = 1, .., h$, calcoliamo il prodotto scalare $\langle Y, Y_i \rangle$, per la linearità del prodotto scalare risulta:

$$\langle Y, Y_i \rangle = a_1 \langle Y_1, Y_i \rangle + ... + a_i \langle Y_i, Y_i \rangle + .. + a_h \langle Y_h, Y_i \rangle = a_i \langle Y_i, Y_i \rangle,$$

poiché $Y_j \perp Y_i$, per $j \neq i$. Possiamo quindi ricavare la $i-$esima coordinata di Y, $i = 1, .., h$:

$$a_i = \frac{\langle Y, Y_i \rangle}{\langle Y_i, Y_i \rangle}, \tag{7.7}$$

detta anche *coefficiente di Fourier* del vettore Y rispetto alla base ortogonale $\mathcal{B}_V$.

Il vettore $\frac{\langle Y, Y_i \rangle}{\langle Y_i, Y_i \rangle} Y_i$ è detto *vettore proiezione ortogonale* del vettore Y su $\text{Span}(Y_i)$, per ogni $i = 1, ..h$.

Abbiamo quindi verificato la seguente proprietà:

Proprietà 7.4. *Fissato in* $\mathbb{R}^n$ *il prodotto scalare standard, sia*

$$\mathcal{B}_V = \{Y_1, Y_2, .., Y_h\}$$

una base ortogonale di un sottospazio $V \subset \mathbb{R}^n$*. Allora* $\forall\ Y\ \in V$, Y *ammette la seguente scrittura:*

$$Y = \frac{\langle Y, Y_1\rangle}{\langle Y_1, Y_1\rangle} Y_1 + \frac{\langle Y, Y_2\rangle}{\langle Y_2, Y_2\rangle} Y_2 + ... + \frac{\langle Y, Y_h\rangle}{\langle Y_h, Y_h\rangle} Y_h, \tag{7.8}$$

cioè Y *è la sommma delle sue proiezioni ortogonali ripettivamente lungo ciascun vettore* Y_i *della base,* $i = 1, .., h$*.*

In particolare, se fissiamo in V una base ortonormale, otteniamo:

Corollario 7.5. *Fissato in* $\mathbb{R}^n$ *il prodotto scalare standard, sia* $\mathcal{B}_V = \{Y_1, Y_2, .., Y_h\}$ *una base ortonormale di un sottospazio* $V \subset \mathbb{R}^n$*. Allora ogni vettore* $Y \in V$ *ammette la seguente scrittura:*

$$Y = \langle Y, Y_1\rangle Y_1 + \langle Y, Y_2\rangle Y_2 + ... + \langle Y, Y_h\rangle Y_h. \tag{7.9}$$

Se fissiamo una base ortonormale in V, possiamo generalizzare le regole di calcolo della geometria euclidea. Abbiamo infatti la seguente:

Proprietà 7.6. *Fissato in* $\mathbb{R}^n$ *il prodotto scalare standard, siano* $V \subset \mathbb{R}^n$ *un sottospazio e* $\mathcal{B}_V = \{Y_1, Y_2, .., Y_h\}$ *una base ortonormale di* V*.*

(1) *se* $Y = a_1Y_1 + ... + a_hY_h$ *e* $Z = b_1Y_1 + ... + b_hY_h$*, allora si ha:*
$$\langle Y, Z\rangle = a_1b_1 + ... + a_hb_h;$$
(2) *se* $Y = a_1Y_1 + ... + a_hY_h$*,*
$$\|Y\|^2 = a_1^2 + a_2^2 + ... + a_h^2,$$
detta Teorema di Pitagora generalizzato.

Dimostrazione.
(1) Calcoliamo il prodotto scalare e applichiamo le proprietà di bilinearità:

$$\langle Y, Z\rangle = \langle a_1Y_1 + ... + a_hY_h, b_1Y_1 + ... + b_hY_h\rangle = \sum_{i,j=1}^{h} a_ib_j\langle Y_i, Y_j\rangle,$$

poiché la base è ortonormale risulta

$$\langle Y_i, Y_j\rangle = 0 \quad \forall i \neq j, \quad \langle Y_i, Y_i\rangle = 1, \forall i = 1, .., h;$$

sostituendo nel prodotto scalare $\langle Y, Z\rangle$ otteniamo:

$$\langle Y, Z\rangle = a_1b_1 + a_2b_2 + .. + a_hb_h.$$

(2) La formula di Pitagora si ottiene semplicemente ponendo $Y = Z$. □

ESEMPIO 7.13. Consideriamo la base canonica $\mathcal{B} = \{\boldsymbol{e}_1, .., \boldsymbol{e}_n\}$ di $\mathbb{R}^n$. Abbiamo verificato che è una base ortonormale rispetto al prodotto scalare standard di $\mathbb{R}^n$. Sia $Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{R}^n$, allora Y ammette l'unica scrittura

$$Y = y_1\boldsymbol{e}_1 + y_2\boldsymbol{e}_2 + ... + y_n\boldsymbol{e}_n,$$

come combinazione lineare dei vettori della base $\mathcal{B}$. Per quanto abbiamo visto sopra, risulta:

$$y_i = \langle Y, \boldsymbol{e}_i\rangle,$$

ed il vettore $y_i\boldsymbol{e}_i$ è la proiezione ortogonale di Y lungo $\mathrm{Span}(\boldsymbol{e}_i)$.

A questo punto è naturale chiedersi, fissati il prodotto scalare standard in $\mathbb{R}^n$ e un sottospazio non nullo $V \subset \mathbb{R}^n$, *esiste sempre una base ortogonale in V?*

La risposta alla domanda è affermativa ed è conseguenza dell'esistenza di un algoritmo che permette di trasformare un sistema di vettori indipendenti di $\mathbb{R}^n$ in un sistema ortogonale di vettori di $\mathbb{R}^n$, conservando i sottospazi generati da tali vettori.

TEOREMA 7.7. *(Metodo di ortogonalizzazione di Gram-Schmidt)*
Fissato in $\mathbb{R}^n$ il prodotto scalare standard, siano $\{Y_1, Y_2, ..., Y_h\}$, $2 \leq h \leq n$, vettori linearmente indipendenti in $\mathbb{R}^n$. Esistono h vettori $\{X_1, .., X_h\} \in \mathbb{R}^n$ con le seguenti proprietà:

(1) $\mathrm{Span}\ (Y_1, ..., Y_i) = \mathrm{Span}\ (X_1, ..., X_i), \qquad \forall\ i = 1, 2, ..., h;$
(2) $X_i \perp X_j \quad \forall\ i \neq j.$

Per ogni $i = 2, .., h$, il vettore X_i si ottiene sottraendo dal vettore Y_i le sue proiezioni ortogonali rispettivamente sui vettori X_1,..,X_{i-1}.

Dimostriamo come funziona l'algoritmo per $h = 2$ e $h = 3$.
Se $h = 2$, abbiamo due vettori linearmente indipendenti $\{Y_1, Y_2\}$. Cerchiamo una coppia di vettori perpendicolari $\{X_1, X_2\}$ tali che:

(1) $X_1 \in \mathrm{Span}(Y_1)$,
(2) $\mathrm{Span}(X_1, X_2) = \mathrm{Span}(Y_1, Y_2)$.

Possiamo scegliere

$$X_1 = Y_1,$$

ora cerchiamo $X_2 \in \mathrm{Span}(Y_1, Y_2) = \mathrm{Span}(X_1, Y_2)$, tale che $X_2 \perp X_1$. Osserviamo che il sottospazio $\mathrm{Span}(X_1, Y_2)$ ha dimensione 2, quindi è isomorfo a $\mathbb{R}^2$: sappiamo che nel piano esiste un'unica direzione ortogonale a X_1. Poiché Y_2 è somma delle sue proiezioni ortogonali lungo la direzione di X_1 e la direzione ortogonale, sapendo che la proiezione ortogonale di Y_2 su $\mathrm{Span}(X_1)$ è la seguente:

$$\frac{\langle Y_2, X_1\rangle}{\langle X_1, X_1\rangle} X_1,$$

possiamo concludere che il vettore

$$Y_2 - \frac{\langle Y_2, X_1\rangle}{\langle X_1, X_1\rangle} X_1 \in \mathrm{Span}(X_1, Y_2),$$

ha la direzione cercata. Infatti, per la linearità del prodotto scalare risulta:

$$\begin{aligned}\left\langle Y_2 - \frac{\langle Y_2, X_1\rangle}{\langle X_1, X_1\rangle}X_1, X_1\right\rangle &= \langle Y_2, X_1\rangle - \frac{\langle Y_2, X_1\rangle}{\langle X_1, X_1\rangle}\langle X_1, X_1\rangle \\ &= \langle Y_2, X_1\rangle - \langle Y_2, X_1\rangle = 0.\end{aligned}$$

Possiamo scegliere quindi

$$X_2 = Y_2 - \frac{\langle Y_2, X_1\rangle}{\langle X_1, X_1\rangle}X_1,$$

e ottenere un sistema ortogonale che soddisfa (1) e (2).

Se $h = 3$, abbiamo tre vettori linearmente indipendenti $\{Y_1, Y_2, Y_3\}$. Cerchiamo un sistema di vettori $\{X_1, X_2, X_3\}$, a due a due perpendicolari, tali che:

(1) $X_1 \in \mathrm{Span}(Y_1)$,
(2) $\mathrm{Span}(X_1, X_2) = \mathrm{Span}(Y_1, Y_2)$,
(3) $\mathrm{Span}(X_1, X_2, X_3) = \mathrm{Span}(Y_1, Y_2, Y_3)$.

Per il passo precedente, possiamo scegliere:

$$X_1 = Y_1, \quad X_2 = Y_2 - \frac{\langle Y_2, X_1\rangle}{\langle X_1, X_1\rangle}X_1 \in \mathrm{Span}(X_1, Y_2),$$

$\{X_1, X_2\}$ è un sistema ortogonale e soddisfa le condizioni (1) e (2). Osserviamo che il sottospazio $\mathrm{Span}(X_1, X_2, Y_3)$ ha dimensione 3, quindi è isomorfo a $\mathbb{R}^3$: sappiamo che nello spazio esiste un'unica direzione ortogonale ad un piano assegnato. Il vettore Y_3 è la somma delle sue proiezioni ortogonali rispettivamente su $\mathrm{Span}(X_1, X_2)$ e sulla direzione ortogonale, la proiezione ortogonale di Y_3 su $\mathrm{Span}(X_1, X_2)$ si ottiene sommando le proiezioni ortogonali su $\mathrm{Span}(X_1)$ e $\mathrm{Span}(X_2)$:

$$\frac{\langle Y_3, X_1\rangle}{\langle X_1, X_1\rangle}X_1 + \frac{\langle Y_3, X_2\rangle}{\langle X_2, X_2\rangle}X_2.$$

Possiamo quindi concludere che il vettore

$$Y_3 - \left(\frac{\langle Y_3, X_1\rangle}{\langle X_1, X_1\rangle}X_1 + \frac{\langle Y_3, X_2\rangle}{\langle X_2, X_2\rangle}X_2\right),$$

ha la direzione cercata. Infatti, per la linearità del prodotto scalare risulta:

$$\left\langle Y_3 - \frac{\langle Y_3, X_1\rangle}{\langle X_1, X_1\rangle}X_1 - \frac{\langle Y_3, X_2\rangle}{\langle X_2, X_2\rangle}X_2, X_1\right\rangle = \langle Y_3, X_1\rangle - \frac{\langle Y_3, X_1\rangle}{\langle X_1, X_1\rangle}\langle X_1, X_1\rangle = 0,$$

$$\left\langle Y_3 - \frac{\langle Y_3, X_1\rangle}{\langle X_1, X_1\rangle}X_1 - \frac{\langle Y_3, X_2\rangle}{\langle X_2, X_2\rangle}X_2, X_2\right\rangle = \langle Y_3, X_2\rangle - \frac{\langle Y_3, X_2\rangle}{\langle X_2, X_2\rangle}\langle X_2, X_2\rangle = 0.$$

Possiamo quindi scegliere

$$X_3 = Y_3 - (\frac{\langle Y_3, X_1\rangle}{\langle X_1, X_1\rangle}X_1 + \frac{\langle Y_3, X_2\rangle}{\langle X_2, X_2\rangle}X_2),$$

e ottenere un sistema ortogonale che soddisfa (1), (2) e (3).

Esempio 7.14. Fissato in $\mathbb{R}^4$ il prodotto scalare standard, consideriamo il seguente sottospazio di $\mathbb{R}^4$:

$$V = \left\{ Y = \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \in \mathbb{R}^4 \colon x + y + z + t = 0 \right\}.$$

Osserviamo che i vettori

$$Y_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix}, \quad Y_2 = \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix}, \quad Y_3 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ -1 \end{pmatrix}$$

costituiscono una base $\mathcal{B}$ di V. Vogliamo costruire a partire da $\mathcal{B}$ una base ortogonale di V. Poniamo:

$$X_1 = Y_1.$$

Calcoliamo la proiezione ortogonale di Y_2 lungo X_1. Calcoliamo dapprima i seguenti prodotti scalari:

$$\langle X_1, X_1 \rangle = 1^2 + (-1)^2 = 2 \quad \langle Y_2, X_1 \rangle = 1.0 + (-1).1 + 0.(-1) + 0.0 = -1,$$

la proiezione ortogonale di Y_2 lungo X_1 è la seguente:

$$\frac{\langle Y_2, X_1 \rangle}{\langle X_1, X_1 \rangle} X_1 = -\frac{1}{2} X_1.$$

Allora il secondo vettore della base ortogonale si ottiene sottraendo a Y_2 la proiezione ortogonale lungo X_1:

$$X_2 = Y_2 + \frac{1}{2} X_1 = \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ -1 \\ 0 \end{pmatrix}.$$

Calcoliamo ora le proiezioni ortogonali di Y_3 lungo X_1 e X_2. Calcoliamo dapprima i seguenti prodotti scalari:

$$\langle Y_3, X_1 \rangle = 1.0 + (-1).1 + 0.0 + 0.(-1) = -1,$$

$$\langle Y_3, X_2 \rangle = 0.(\frac{1}{2}) + 1.(\frac{1}{2}) + 0.(-1) + (-1).0 = \frac{1}{2},$$

$$\langle X_2, X_2 \rangle = \frac{1}{4} + \frac{1}{4} + 1 = \frac{6}{4}.$$

La proiezione ortogonale di Y_3 lungo X_1 è la seguente:

$$\frac{\langle Y_3, X_1 \rangle}{\langle X_1, X_1 \rangle} X_1 = -\frac{1}{2} X_1.$$

La proiezione ortogonale di Y_3 lungo X_2 è la seguente:

$$\frac{\langle Y_3, X_2 \rangle}{\langle X_2, X_2 \rangle} X_2 = \frac{1}{3} X_2.$$

Infine il terzo vettore della base ortogonale di V si ottiene sottraendo da Y_3 le proiezioni del vettore su X_1 e X_2:

$$X_3 = Y_3 + \frac{1}{2} X_1 - \frac{1}{3} X_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ -1 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix} - \frac{1}{3} \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ -1 \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{3} \\ \frac{1}{3} \\ \frac{1}{3} \\ -1 \end{pmatrix}.$$

Verifichiamo che i vettori $\{X_1, X_2, X_3\}$ sono un sistema ortogonale di $\mathbb{R}^4$. Risulta:

$$\langle X_1, X_2\rangle = \frac{1}{2} - \frac{1}{2} = 0,$$

$$\langle X_1, X_3\rangle = \frac{1}{3} - \frac{1}{3} = 0,$$

$$\langle X_2, X_3\rangle = \frac{1}{6} + \frac{1}{6} - \frac{1}{3} = 0.$$

Abbiamo trovato una base ortogonale di V.

Se vogliamo determinare una base ortonormale per V, basta normalizzare i vettori della base ortogonale. Calcoliamo a tal fine la norma dei vettori della base:

$$\|X_1\| = \sqrt{\langle X_1, X_1\rangle} = \sqrt{2},$$

$$\|X_2\| = \sqrt{\langle X_2, X_2\rangle} = \sqrt{\frac{3}{2}};$$

$$\|X_3\| = \sqrt{\langle X_3, X_3\rangle} = \sqrt{\frac{1}{9} + \frac{1}{9} + \frac{1}{9} + 1} = \sqrt{\frac{4}{3}}.$$

Una base ortonormale per V è la seguente:

$$\left\{\frac{1}{\sqrt{2}}X_1, \sqrt{\frac{2}{3}}X_2, \sqrt{\frac{3}{4}}X_3\right\}.$$

Come conseguenza immediata del processo di ortogonalizzazione di Gram-Schmidt si hanno i seguenti risultati:

Corollario 7.8. *Fissato in $\mathbb{R}^n$ il prodotto scalare standard, sia $V \subset \mathbb{R}^n$ un sottospazio vettoriale non nullo, esiste in V una base ortonormale.*

Corollario 7.9. *Fissato in $\mathbb{R}^n$ il prodotto scalare standard, sia $\{Y_1, .., Y_r\}$ un sistema ortogonale di $r \leq n$ vettori di $\mathbb{R}^n$. Esiste una base ortogonale di $\mathbb{R}^n$ contenente i vettori $\{Y_1, .., Y_r\}$.*

Dimostrazione. Aggiungiamo $n - r$ vettori a $\{Y_1, .., Y_r\}$ fino ad ottenere una base di $\mathbb{R}^n$:

$$\{Y_1, .., Y_r, Y_{r+1}, .., Y_n\}.$$

Applichiamo ora il processo di ortogonalizzazione di Gram-Schmidt, abbiamo una base ortogonale di $\mathbb{R}^n$

$$\{X_1, ..., X_n\}.$$

Poiché per ogni $i = 2, .., n$, il vettore X_i si ottiene sottraendo dal vettore Y_i le sue proiezioni ortogonali rispettivamente sui vettori X_1,..,X_{i-1}, risulta:

$$X_i = Y_i, \quad \forall i = 1, .., r,$$

quindi la base trovata contiene i vettori $\{Y_1, .., Y_r\}$. $\square$

3. Matrici ortogonali.

Come abbiamo visto nel paragrafo 2, l'uso di basi ortonormali può aiutare a semplificare i calcoli (Proprietà 7.6), oltre che ad essere utile per analisi teoriche. Ci chiediamo ora se il cambio di base in questo caso presenta alcune strutture speciali. Per rispondere a questo, analizzeremo quella particolare classe di matrici invertibili che risponde al nostro scopo.

Per esempio, consideriamo lo spazio $\mathbb{E}^2_O$ e la sua base canonica $\mathcal{B} = \{\hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}\}$; possiamo cambiare la base "ruotando" i vettori, ma mantenendo la loro ortogonalità.

Per esempio, possiamo usare la nuova base ortogonale $\mathcal{B}' = \{\overrightarrow{u}_1, \overrightarrow{u}_2\}$, con $\overrightarrow{u}_1 = \hat{\boldsymbol{\imath}} + \hat{\boldsymbol{\jmath}}$ e $\overrightarrow{u}_2 = -\hat{\boldsymbol{\imath}} + \hat{\boldsymbol{\jmath}}$, si verifica facilmente che $\overrightarrow{u}_1 \perp \overrightarrow{u}_2$. Per avere una base ortonormale, basta dividere ciascun vettore $\overrightarrow{u}_i$ per il suo modulo: $\mathcal{B}'' = \left\{\overrightarrow{u}_1/\sqrt{2}, \overrightarrow{u}_2/\sqrt{2}\right\}$, e si osserva subito che la nuova base si ottiene ruotando la base originale in senso antiorario di $\pi/4$.

La matrice del cambio di base fra $\mathcal{B}''$ e $\mathcal{B}$ si ottiene mettendo in colonna le coordinate dei vettori di $\mathcal{B}''$ nella base $\mathcal{B}$:

$$A = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix}; \tag{7.10}$$

le coordinate di un qualunque vettore $\overrightarrow{v}$ di $\mathbb{E}^2_O$ sono date dalla seguente relazione

$$X = AX'', \tag{7.11}$$

dove $X = [\overrightarrow{v}]_{\mathcal{B}}$ e $X'' = [\overrightarrow{v}]_{\mathcal{B}''}$. La matrice inversa si può ottenere osservando che la base $\mathcal{B}$ è ruotata di un angolo $\pi/4$ in senso orario rispetto a $\mathcal{B}'$; si ottiene

$$A' = A^{-1} = \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix}. \tag{7.12}$$

Ripetendo lo stesso percorso con un angolo di rotazione differente, per esempio $\pi/6$, otteniamo le due matrici B e B' di cambio di base diretta e inversa:

$$B = \begin{pmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix} \qquad B' = \begin{pmatrix} \frac{\sqrt{3}}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix} \tag{7.13}$$

Osserviamo che la matrice inversa di trasfomazione si ottiene molto facilmente da quella originaria: in effetti, è la sua trasposta; ma per potere essere sicuri del risultato, e per estenderlo anche a spazi n-dimensionali, dobbiamo, come al solito, procedere formalizzando il problema.

Iniziamo a caratterizzare quelle matrici che si comportano come sopra, ossia quelle la cui trasposta coincide con l'inversa. Questo si ottiene con la seguente

Definizione 7.6. Una matrice $Q \in \mathcal{GL}(n, \mathbb{R})$ é detta *matrice ortogonale di ordine n* se

$$Q^{\mathrm{T}} = Q^{-1}, \tag{7.14}$$

ossia se

$$Q\,Q^{\mathrm{T}} = Q^{\mathrm{T}}\,Q = I_n\,. \tag{7.15}$$

L'insieme delle matrici ortogonali di ordine n è indicato con $\mathcal{O}(n)$.

Vale la pena di notare subito alcune proprietà caratteristiche.

OSSERVAZIONE 7.15. La matrice identità I_n è ortogonale per qualunque n.

OSSERVAZIONE 7.16. Se $Q \in \mathcal{O}(n)$, allora $Q^{-1} = Q^{\mathrm{T}} \in \mathcal{O}(n)$.

DIMOSTRAZIONE. Banali entrambe le affermazioni: la prima si verifica immediatamente, la seconda è conseguenza immediata del fatto che $(Q^{\mathrm{T}})^{\mathrm{T}} = Q$, per qualunque matrice Q. □

OSSERVAZIONE 7.17. L'insieme $\mathcal{O}(n)$ è un gruppo rispetto all'operazione di moltiplicazione fra matrici, ed è pertanto un sottogruppo di $\mathcal{GL}(n,\mathbb{R})$

DIMOSTRAZIONE. L'identità è nell'insieme (vedi osservazione 7.15); quindi $\mathcal{O}(n)$ contiene l'elemento neutro; inoltre, per ogni elemento il suo opposto si trova ancora nell'insieme (vedi osservazione 7.16). Basta vedere che l'insieme è chiuso rispetto alla moltiplicazione; se $Q_1, Q_2 \in \mathcal{O}(n)$, allora il loro prodotto P è in $\mathcal{O}(n)$. Infatti

$$\begin{aligned} PP^{\mathrm{T}} &= (Q_1Q_2)(Q_1Q_2)^{\mathrm{T}} = (Q_1Q_2)(Q_2^{\mathrm{T}}Q_1^{\mathrm{T}}) \\ &= Q_1(Q_2Q_2^{\mathrm{T}})Q_1^{\mathrm{T}} = Q_1 I_n Q_1^{\mathrm{T}} = Q_1Q_1^{\mathrm{T}} = I_n \end{aligned}$$

e lo stesso vale per $P^{\mathrm{T}}P$. □

Le matrici A, A', B, B' riportate sopra (7.10,7.12,7.13) sono tutte matrici ortogonali. Vediamo anche altri esempi.

ESEMPIO 7.18.

(1) La matrice

$$Q_1 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

è ortogonale di ordine 3.

(2) La matrice

$$Q_2 = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{\sqrt{3}}{2} \\ 0 & 0 & \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}$$

è ortogonale di ordine 4.

(3) La matrice dell'esempio 6.4 nel capitolo 6

$$Q_3 = \begin{pmatrix} 0 & -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{1}{2} & \frac{1}{2} \\ -\frac{\sqrt{2}}{2} & \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

è ortogonale di ordine 3.

Lasciamo allo studente il compito di verificare le affermazioni.

Una condizione necessaria caratterizza tutte le matrici ortogonali:

PROPOSIZIONE 7.10. *Se* $Q \in \mathcal{O}(n)$ *allora* $\det(Q) = \pm 1$.

DIMOSTRAZIONE. È una conseguenza immediata del teorema di Binet (3.16) e del teorema del determinante della matrice trasposta. Infatti, se Q è ortogonale:

$$1 = \det(I_n) = \det(QQ^{\mathrm{T}}) = \det(Q)\det(Q^{\mathrm{T}}) = \det(Q)\det(Q) = (\det(Q))^2$$

Da cui, estraendo la radice, $\det(Q) = \pm 1$. □

La condizione **non** è affatto sufficiente: basta considerare la matrice

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

si verifica immediatamente che $\det(A) = 1$ (è triangolare), ma

$$AA^{\mathrm{T}} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \neq I_n,$$

quindi A *non* è ortogonale.

Una condizione necessaria, per quanto utile in molti casi, non ci basta. Vogliamo una caratterizzazione più forte, quindi una condizione necessaria e sufficiente, ossia un'equivalenza. Questa ci viene fornita dalla seguente

PROPRIETÀ 7.11. *Sia* $Q \in \mathcal{GL}(n, \mathbb{R})$ *una matrice reale invertibile di ordine* n*. Allora* Q *è ortogonale se e solo se le sue colonne sono un insieme ortonormale di vettori di* $\mathbb{R}^n$*.*

DIMOSTRAZIONE. Calcoliamo il prodotto $Q^{\mathrm{T}}Q$. Ricordiamo che l'elemento $(Q^{\mathrm{T}}Q)_{i,j}$, di posto i, j, della matrice prodotto $Q^{\mathrm{T}}Q$ si ottiene moltiplicando la riga i-esima di Q^{T} con la colonna j-esima di Q. Poiché la i-esima riga di Q^{T} è la trasposta della i-esima colonna di Q, l'elemento $(Q^{\mathrm{T}}Q)_{i,j}$ si ottiene calcolando il prodotto scalare standard dei vettori colonna Q^i e Q^j:

$$(Q^{\mathrm{T}}Q)_{i,j} = (Q^{\mathrm{T}})_i Q^j = (Q^i)^{\mathrm{T}} Q^j = \langle Q^i, Q^j \rangle. \tag{7.16}$$

Se Q è ortogonale, allora $Q^{\mathrm{T}}Q = I_n$, quindi

$$(Q^{\mathrm{T}}Q)_{ij} = (I_n)_{i,j} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} \tag{7.17}$$

ossia, i vettori colonna sono un'insieme ortonormale. Viceversa, se i vettori colonna sono ortonormali, vale la relazione 7.17, quindi $QQ^{\mathrm{T}} = I_n$, ossia Q è ortogonale. □

Analogo procedimento, applicato a QQ^{T} permette di mostrare anche un'affermazione analoga sulle righe:

PROPRIETÀ 7.12. *Sia* $Q \in \mathcal{GL}(n, \mathbb{R})$ *una matrice reale invertibile di ordine* n*. Allora* Q *è ortogonale se e solo se le sue righe sono un insieme ortonormale in* $\mathbb{R}^n$*.*

Conseguenza immediata della proprietà 7.11

COROLLARIO 7.13. *Sia* $Q \in \mathcal{O}(n)$*, allora le colonne di* Q *sono una base ortonormale di* $\mathbb{R}^n$*.*

Possiamo ora esprimere il legame fra le matrici ortogonali e le matrici di cambio di base fra basi ortonormali.

PROPOSIZIONE 7.14. *Sia $Q \in \mathcal{GL}(n, \mathbb{R})$ una matrice reale invertibile di ordine n. Allora*

(1) *Q è ortogonale se se solo se realizza un cambiamento di basi ortonormali in $\mathbb{R}^n$.*
(2) *Q è ortogonale se se solo se conserva il prodotto scalare standard, cioè*

$$\langle QX, QY \rangle = \langle X, Y \rangle \quad \forall X, Y \ \in \mathbb{R}^n.$$

DIMOSTRAZIONE.

(1) Verifichiamo che se Q realizza il cambio di base dell'enunciato, allora $Q \in \mathcal{O}(n)$. Siano $\mathcal{B} = \{\boldsymbol{w}_1, \dots, \boldsymbol{w}_n\}$ e $\mathcal{B}' = \{\boldsymbol{w}'_1, \dots, \boldsymbol{w}'_n\}$ due basi ortonormali di $\mathbb{R}^n$; sia $Q \in \mathcal{GL}(n, \mathbb{R})$ la matrice di cambio di base, ossia la i-esima colonna di Q è $[\boldsymbol{w}'_i]_{\mathcal{B}}$, vale a dire la rappresentazione dell'i-esimo vettore della base $\mathcal{B}'$ nella base $\mathcal{B}$:

$$\boldsymbol{w}'_i = q_{1i}\boldsymbol{w}_1 + q_{2i}\boldsymbol{w}_2 + \dots q_{ni}\boldsymbol{w}_n \,. \tag{7.18}$$

Usando la (7.18) per il vettore $\boldsymbol{w}'_j \in \mathcal{B}'$, e sfruttando la formula per il calcolo del prodotto scalare in una base ortonormale, 7.6, possiamo scrivere, se $i \neq j$:

$$\begin{aligned} 0 &= \langle \boldsymbol{w}'_i, \boldsymbol{w}'_j \rangle = q_{1i}q_{1j} + q_{2i}q_{2j} + \dots q_{ni}q_{nj} \\ &= \langle Q^i, Q^j \rangle \,; \end{aligned} \tag{7.19}$$

inoltre, se $i = j$, la precedente scrittura si modifica in

$$\begin{aligned} 1 &= \langle \boldsymbol{w}'_i, \boldsymbol{w}'_i \rangle = q_{1i}q_{1i} + q_{2i}q_{2i} + \dots q_{ni}q_{ni} \\ &= q_{1i}^2 + q_{2i}^2 + \dots q_{ni}^2 \\ &= \langle Q^i, Q^i \rangle \,. \end{aligned} \tag{7.20}$$

In altre parole, le colonne di Q sono un insieme ortonormale, e, pertanto Q è ortogonale, in base alla condizione espressa dalla 7.11.

Viceversa, se Q è ortogonale, le sue colonne sono una base ortonormale di $\mathbb{R}^n$ (corollario 7.13), ossia Q è la matrice di cambio di base fra la base canonica, che è ortonormale, e $\mathcal{B}'$, anch'essa ortonormale.

(2) La proprietà enunciata si riferisce all'operatore lineare $\mathbb{R}^n \to \mathbb{R}^n$ definito dalla matrice Q: $X \to Q.X$. Sia $Q \in \mathcal{O}(n)$, per qualunque X, Y in $\mathbb{R}^n$, si ha:

$$\langle QX, QY \rangle = (QX)^{\mathrm{T}}QY = X^{\mathrm{T}}Q^{\mathrm{T}}QY = X^{\mathrm{T}}I_nY = X^{\mathrm{T}}, Y = \langle X, Y \rangle. \tag{7.21}$$

Viceversa, se il prodotto scalare viene conservato, consideriamo i vettori della base standard di $\mathbb{R}^n$, $\mathcal{B}_0 = \{\boldsymbol{e}_1, \boldsymbol{e}_2 \dots \boldsymbol{e}_n\}$; possiamo interpretare i vettori colonna della matrice Q come i trasformati secondo l'operatore lineare associato alla matrice stessa dei vettori di $\mathcal{B}_0$:

$$Q^i = [L_Q(\boldsymbol{e}_i)]_{\mathcal{B}_0}.$$

Ora, il prodotto scalare si conserva, perciò

$$(Q^i)^{\mathrm{T}}Q^j = \langle Q^i, Q^j \rangle = \langle Q(\boldsymbol{e}_i), \langle Q(\boldsymbol{e}_j) \rangle = \langle \boldsymbol{e}_i, \boldsymbol{e}_j \rangle,$$

cioè i vettori colonna sono un insieme ortonormale, quindi Q è ortogonale (proposizione 7.11). □

Tornando agli esempi fatti all'inizio del paragrafo, ci chiediamo ora se sia possibile caratterizzare ulteriormente le matrici ortogonali di ordine 2 e 3; il maggiore interesse verso queste risiede nel fatto che descrivere applicazioni lineari in $\mathbb{E}^2_O$ ed $\mathbb{E}^3_O$, che si prestano ad essere visualizzate geometricamente. Cominceremo con le matrici di ordine 2: per esse, vale la seguente

Proposizione 7.15. *Le matrici $\mathcal{O}(2)$ sono tutte e sole quelle che hanno una delle seguenti forme:*

$$Q_p = \begin{pmatrix} \cos\vartheta & -\sin\vartheta \\ \sin\vartheta & \cos\vartheta \end{pmatrix} \qquad Q_m = \begin{pmatrix} \cos\vartheta & \sin\vartheta \\ \sin\vartheta & -\cos\vartheta \end{pmatrix} \tag{7.22}$$

con $\vartheta \in [0, 2\pi)$.

Dimostrazione. Verifichiamo che se le matrici hanno la forma della (7.22) sono ortogonali. Anzitutto, osserviamo che

$$\det(Q_p) = \cos^2\vartheta - (-\sin^2\vartheta) = \cos^2\vartheta + \sin^2\vartheta = 1$$

e

$$\det(Q_m) = \cos^2\vartheta - \sin^2\vartheta = -1$$

per qualunque valore di ϑ; la condizione necessaria (7.10) è verificata. Inoltre,

$$\langle (Q_p)^1, (Q_p)^2 \rangle = \cos\vartheta(-\sin\vartheta) + \sin\vartheta\,\cos\vartheta = 0 \tag{7.23a}$$

$$\langle (Q_p)^1, (Q_p)^1 \rangle = (\cos\vartheta)^2 + (\sin\vartheta)^2 = 1 \tag{7.23b}$$

$$\langle (Q_p)^2, (Q_p)^2 \rangle = (\cos\vartheta)^2 + (-\sin\vartheta)^2 = 1\,, \tag{7.23c}$$

ossia, le colonne di Q_p sono un insieme di vettori ortonormali, quindi $Q_p \in \mathcal{O}(2)$ (proprietà 7.11); analogamente, otteniamo che anche Q_m è ortogonale.

Vediamo l'altra implicazione; sia

$$Q = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \in \mathcal{O}(2)$$

una matrice ortogonale. Le condizioni della proprietà 7.11 continuano a valere; queste impongono

$$\langle Q^1, Q^2 \rangle = a_{11}a_{12} + a_{21}a_{22} = 0 \tag{7.24a}$$

$$\langle Q^1, Q^1 \rangle = (a_{11})^2 + (a_{21})^2 = 1 \tag{7.24b}$$

$$\langle Q^2, Q^2 \rangle = (a_{12})^2 + (a_{22})^2 = 1 \tag{7.24c}$$

Le ultime due equazioni (7.24b,7.24c) permettono di affermare che esistono due angoli ϑ e φ tali che

$$a_{11} = \cos\vartheta \qquad a_{21} = \sin\vartheta \tag{7.25a}$$

$$a_{12} = \cos\varphi \qquad a_{22} = \sin\varphi \tag{7.25b}$$

inoltre, la prima equazione (7.24a) diventa

$$\cos\vartheta\cos\varphi + \sin\vartheta\sin\varphi = \cos(\varphi - \vartheta) = 0\,,$$

ossia, $\varphi = \vartheta + k\frac{\pi}{2}$ con $k \in \mathbb{Z}$. Per quanto riguarda la struttura della matrice Q non perdiamo in generalità se ci limitiamo ai valori $\varphi = \vartheta \pm \frac{\pi}{2}$.

Ora, con $\varphi = \vartheta + \frac{\pi}{2}$, abbiamo

$$Q = \begin{pmatrix} \cos\vartheta & \cos(\vartheta + \frac{\pi}{2}) \\ \sin\vartheta & \sin(\vartheta + \frac{\pi}{2}) \end{pmatrix} = \begin{pmatrix} \cos\vartheta & -\sin\vartheta \\ \sin\vartheta & \cos\vartheta \end{pmatrix}\,,$$

che è la struttura della matrice Q_1. Invece, con $\varphi = \vartheta - \frac{\pi}{2}$, abbiamo

$$Q = \begin{pmatrix} \cos\vartheta & \cos(\vartheta - \frac{\pi}{2}) \\ \sin\vartheta & \sin(\vartheta - \frac{\pi}{2}) \end{pmatrix} = \begin{pmatrix} \cos\vartheta & \sin\vartheta \\ \sin\vartheta & -\cos\vartheta \end{pmatrix} = Q_m \, .$$

Quindi, se Q è ortogonale, ha la forma data nella (7.22) □

A questo punto risulta immediata l'interpretazione geometrica:

Osservazione 7.19. Dal punto di vista geometrico:

(1) sia $Q_p \in \mathcal{O}(2)$ come in (7.22); la matrice Q_p descrive un'operatore di rotazione di un angolo ϑ antiorario (se $\vartheta > 0$) dei vettori. Notiamo che in questo caso $\det(Q_p) = 1$.

(2) Sia $Q_n \in \mathcal{O}(2)$ come in (7.22); la matrice descrive un'operatore di simmetria rispetto alla retta $\mathrm{Span}(\overrightarrow{u})$, con $\overrightarrow{u}$ che forma un angolo $\frac{\vartheta}{2}$ rispetto a $\hat{\boldsymbol{\imath}}$ (l'angolo viene contato positivamente in senso antiorario). Alternativamente, si può interpretare l'operatore come una rotazione di un angolo ϑ come nel primo caso, composta con una inversione rispetto alla direzione del vettore $\hat{\boldsymbol{\imath}}$ ruotato. Notiamo che stavolta $\det(Q_m) = -1$.

Quest'ultima osservazione ci mostra che il segno del determinante gioca un ruolo importante. In effetti, le matrici ortogonali a determinante positivo formano un sottogruppo: ha senso quindi dare la seguente

Definizione 7.7. Una matrice $Q \in \mathcal{O}(n)$ viene detta *ortogonale speciale* se $\det(Q) = +1$. L'insieme delle matrici ortogonali speciale viene denotato con il simbolo $\mathcal{SO}(n)$, ossia

$$\mathcal{SO}(n) = \{Q \in \mathcal{O}(n) \mid \det(Q) = +1\} \, .$$

Lasciamo allo studente il compito di verificare che effettivamente $\mathcal{SO}(n)$ è un sottogruppo di $\mathcal{O}(n)$. Osserviamo comunque che le matrici con determinante negativo **non** sono un sottogruppo: infatti, se $\det(Q_1) = -1$ e $\det(Q_2) = -1$, il loro prodotto ha determinante $\det(Q_1)\det(Q_2) = (-1)(-1) = +1$, e l'insieme non è chiuso rispetto alla moltiplicazione.

Inoltre, per le matrici ortogonali di ordine 2, abbiamo un legame molto forte fra la diagonalizzabilità ed il segno del determinante. Vediamolo nel dettaglio:

Proposizione 7.16. *Siano $Q_p, Q_n \in \mathcal{O}(2)$ le matrici ortogonali delle equazioni (7.22) introdotte nella Proposizione 7.15. La matrice Q_n ha sempre due autovalori distinti reali $\lambda_{1,2} = \pm 1$, e, quindi è sempre diagonalizzabile. La matrice Q_p ha autovalori reali se e solo se $\vartheta = 0$ oppure $\vartheta = \pi$; in tale caso la matrice è diagonalizzabile, altrimenti, non ammette autovalori reali.*

Dimostrazione. Basta effettuare il calcolo. Cominciamo con Q_n; calcoliamo il polinomio caratteristico

$$\begin{aligned} p_{Q_n}(\lambda) = \det(Q_n - \lambda I) &= \begin{vmatrix} \cos\vartheta - \lambda & \sin\vartheta \\ \sin\vartheta & -\cos\vartheta - \lambda \end{vmatrix} \\ &= -(\cos\vartheta + \lambda)(\cos\vartheta - \lambda) - \sin^2\vartheta = -(\cos^2\vartheta - \lambda^2) - \sin^2\vartheta \\ &= \lambda^2 - \cos^2\vartheta - \sin^2\vartheta \\ &= \lambda^2 - 1 \, . \end{aligned} \tag{7.26}$$

Chiaramente, per qualunque valore di ϑ, il polinomio caratteristico ha due radici, $\lambda_1 = -1$ e $\lambda_2 = +1$, entrambe con molteplicità algebrica unitaria. La matrice Q_n è di ordine 2, quindi questo basta per dire che è diagonalizzabile (Teorema 6.13: 2 autovalori distinti).

Vediamo invece il caso di Q_p. Calcoliamo anche per essa il polinomio caratteristico

$$p_{Q_p}(\lambda) = \det(Q_p - \lambda I) = \begin{vmatrix} \cos\vartheta - \lambda & -\sin\vartheta \\ \sin\vartheta & \cos\vartheta - \lambda \end{vmatrix} \tag{7.27}$$
$$+ (\cos\vartheta - \lambda)^2 + \sin^2\vartheta\,;$$

ora, osserviamo che $p_{Q_p}(\lambda)$ è la somma di due quantità al quadrato, quindi in campo reale si può annullare solo quando entrambe le quantità sono nulle. Ciò siginifica che il polinomio caratteristico può avere radici reali solo se $\sin\vartheta = 0$, ossia se $\vartheta = 0$ oppure $\vartheta = \pi$; nel primo caso, le radici di $p_{Q_p}(\lambda)$ si trovano per $(\cos(0) - \lambda_a) = 0$, ossia per $\lambda_a = 1$. Nel secondo caso, per $(\cos(\pi) - \lambda_b) = 0$, ossia per $\lambda_b = -1$.

In entrambi i casi, la radice è doppia; tuttavia, la matrice Q_p diventa, nei due casi:

$$Q_p(\vartheta = 0) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad Q_p(\vartheta = \pi) = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}\,; \tag{7.28}$$

cioè la matrice è diagonale, quindi banalmente diagonalizzabile. □

A proposito degli autovalori di una matrice ortogonale qualunque, possiamo subito riportare la seguente

Osservazione 7.20. Sia A una matrice ortogonale di ordine n. Se $\alpha \in \mathbb{R}$ è un autovalore di A, allora $|\alpha| = 1$.

Dimostrazione. Sia $X \in \mathbb{R}^n$ un autovettore di A associato all'autovalore α, poiché A è ortogonale si ha:

$$\langle X, X\rangle = \langle AX, AX\rangle = \langle \alpha X, \alpha X\rangle = \alpha^2\langle X, X\rangle.$$

Otteniamo

$$(1 - \alpha^2)\langle X, X\rangle = 0,$$

poiché X è un autovettore, $\langle X, X\rangle \neq 0$, per cui necessariamente

$$\alpha^2 = 1.$$

□

Possiamo dire qualcosa anche sulle matrici ortogonali di ordine 3? La struttura non è più così semplice, si veda l'appendice.

4. Appendice: Complemento ortogonale di un sottospazio

Studiando la geometria euclidea abbiamo parlato di piano ortogonale ad una retta nello spazio, o di retta normale ad un piano. Ci proponiamo di generalizzare tali concetti allo spazio $\mathbb{R}^n$, introducendo il complemento ortogonale di un sottospazio.

DEFINIZIONE 7.8. Fissato in $\mathbb{R}^n$ il prodotto scalare standard, sia $S \subset \mathbb{R}^n$ un sottoinsieme non vuoto. Chiamiamo **insieme ortogonale** di S il seguente sottoinsieme di $\mathbb{R}^n$:

$$S^\perp = \{Y \in \mathbb{R}^n : \ Y \perp X \ , \ \ \forall \ X \in S \ \} = \{Y \in \mathbb{R}^n : \ \ \langle Y, X \rangle = 0, \ \ \forall \ X \in S \ \}.$$

ESEMPIO 7.21.

(1) Se $S = \{\mathbf{0}_n\}$, allora $S^\perp = \mathbb{R}^n$, perché ogni vettore $X \in \mathbb{R}^n$ è perpendicolare al vettore nullo.
(2) Se $S = \mathbb{R}^n$, allora $S^\perp = \{\mathbf{0}_n\}$, perché l'unico vettore ortogonale a tutti i vettori di $\mathbb{R}^n$ è il vettore nullo.
(3) Siano $\boldsymbol{v} \in \mathbb{R}^3$ un vettore non nullo e $S = \{\boldsymbol{v}\}$. L'insieme $S^\perp$ è il piano per l'origine perpendicolare a $\boldsymbol{v}$. Sia $V = \mathrm{Span}(\boldsymbol{v})$, la retta generata da $\boldsymbol{v}$, allora risulta $V^\perp = S^\perp$.
(4) Sia $S = \{\boldsymbol{u}_1, \boldsymbol{u}_2\}$, formato da due vettori linearmente indipendenti in $\mathbb{R}^3$. L' insieme $S^\perp$ è la retta per l'origine di direzione $\boldsymbol{u}_1 \wedge \boldsymbol{u}_2$. Consideriamo il piano $V = \mathrm{Span}(\boldsymbol{u}_1, \boldsymbol{u}_2)$, allora risulta $V^\perp = S^\perp$ è la retta per l'origine perpendicolare a V.

In generale, l'insieme ortogonale di un sottoinsieme soddisfa le seguenti proprietà:

PROPRIETÀ 7.17. *Fissato in $\mathbb{R}^n$ il prodotto scalare standard, sia $S \subset \mathbb{R}^n$ un sottoinsieme non vuoto.*

(1) *$S^\perp$ è un sottospazio vettoriale di $\mathbb{R}^n$,*
(2) *$S \cap S^\perp = \{\mathbf{0}_n\}$,*
(3) *se $S_1 \subset S_2$, allora $S_2^\perp \subset S_1^\perp$,*
(4) *siano $S = \{X_1, .., X_s\}$ e $V = \mathrm{Span}(S)$: allora $S^\perp = V^\perp$.*

DIMOSTRAZIONE. (1) Proviamo che $S^\perp$ è chiuso rispetto alla somma. Siano $Y_1, Y_2 \in S^\perp$, per la linearità del prodotto scalare e poiché i vettori Y_1 e Y_2 sono perpendicolari a X, per ogni $X \in S$, si ha:

$$\langle Y_1 + Y_2, X \rangle = \langle Y_1, X \rangle + \langle Y_2, X \rangle = 0 + 0 = 0, \quad \forall X \in S,$$

per cui possiamo concludere che $Y_1 + Y_2 \in S^\perp$. Proviamo ora che $S^\perp$ è chiuso rispetto al prodotto per uno scalare. Siano $Y \in S^\perp$ e $\alpha \in \mathbb{R}$, per la linearità del prodotto scalare e poiché il vettore Y è perpendicolare a X, per ogni $X \in S$, si ha:

$$\langle \alpha Y, X \rangle = \alpha \langle Y, X \rangle = \alpha.0 = 0, \quad \forall X \in S,$$

per cui possiamo concludere che $\alpha Y \in S^\perp$.
(2) Sia $Y \in S \cap S^\perp$, poiché $Y \in S^\perp$ si ha $Y \perp X$ per ogni vettore $X \in S$, in particolare $Y \perp Y$ poiché $Y \in S$. Risulta quindi $\langle Y, Y \rangle = 0$ e per la positività del prodotto scalare necessariamente $Y = \mathbf{0}_n$.
(3) Sia $S_1 \subset S_2$. Se $Y \in S_2^\perp$, risulta $Y \perp X$, $\forall X \in S_2$. Poiché $S_1 \subset S_2$, si ha anche $Y \perp X$, $\forall X \in S_1$, da cui deduciamo che $Y \in S_1^\perp$.

(4) Poiché $S \subset V$ dalla (3) si ha l'inclusione $V^\perp \subset S^\perp$. Basta provare l'inclusione $S^\perp \subset V^\perp$. Siano $Y \in S^\perp$ e $X \in V$. Poiché $V = \mathrm{Span}(S)$, risulta

$$X = a_1 X_1 + .. + a_s X_s, \quad \forall X \in V.$$

Per la linearità del prodotto sclare si ha:

$$\langle Y, X\rangle = \langle Y, a_1X_1 + .. + a_sX_s\rangle = a_1\langle Y, X_1\rangle + ... + a_s\langle Y, X_s\rangle,$$

poiché $Y \in S^\perp$, risulta $\langle Y, X_i\rangle = 0$, $i = 1, .., s$, per cui possiamo concludere che $\langle Y, X\rangle = 0$, per ogni $X \in V$. □

Il sottospazio ortogonale di un sottospazio ha la seguente proprietà fondamentale:

PROPOSIZIONE 7.18. *Fissato in $\mathbb{R}^n$ il prodotto scalare standard, sia $V \subset \mathbb{R}^n$ un sottospazio proprio di $\mathbb{R}^n$. Allora risulta:*

$$\mathbb{R}^n = V \oplus V^\perp, \tag{7.29}$$

*perciò $V^\perp$ è detto **complemento ortogonale** di V.*

DIMOSTRAZIONE. Sia $\dim V = h < n$. Dalla proprietà precedente 7.18, i sottospazi V e $V^\perp$ sono in somma diretta. Quindi abbiamo

$$\dim(V \oplus V^\perp) = \dim V + \dim V^\perp \leq n,$$

da cui ricaviamo che $\dim V^\perp \leq n-h$. Basta provare allora che $\dim V^\perp = n-h$, infatti risulta $\dim(V \oplus V^\perp) = n$ e quindi $\mathbb{R}^n = V \oplus V^\perp$.
Sia $\mathcal{B}_V = \{X_1, .., X_h\}$ una base ortogonale di V. Esiste allora una base ortogonale di $\mathbb{R}^n$ che la contiene:

$$\{X_1, .., X_h, Y_1, .., Y_{n-h}\}.$$

Per ogni $i = 1, .., n-h$, risulta:

$$Y_i \perp X_j, \quad \forall j = 1, .., h,$$

dalla proprietà precedente (3), segue che $Y_i \in V^\perp$, per ogni $i = 1, ., n-h$. Poiché i vettori $\{\ Y_1, .., Y_{n-h}\}$ sono linearmente indipendenti, questo implica che $\dim V^\perp = n-h$. □

Come conseguenza abbiamo il seguente:

COROLLARIO 7.19. *Fissato in $\mathbb{R}^n$ il prodotto scalare standard, sia $V \subset \mathbb{R}^n$ un sottospazio di $\mathbb{R}^n$ di dimensione $h < n$. Fissata una base $\mathcal{B}_V = \{Y_1, .., Y_h\}$, il complemento ortogonale di $V^\perp$ è il sottospazio di $\mathbb{R}^n$ di dimensione $n-h$ definito dalle seguenti h equazioni indipendenti:*

$$\langle X, Y_1\rangle = ... = \langle X, Y_h\rangle = 0.$$

ossia

$$\begin{cases} x_1y_{1,1} + x_2y_{1,2} + \cdots + x_ny_{1,n} & = 0 \\ x_1y_{2,1} + x_2y_{2,2} + \cdots + x_ny_{2,n} & = 0 \\ \vdots \\ x_1y_{h,1} + x_2y_{h,2} + \cdots + x_ny_{h,n} & = 0 \end{cases}$$

Esempio 7.22.
Consideriamo il seguente sottospazio di $\mathbb{R}^4$:

$$V = \left\{(x, y, z, t) \in \mathbb{R}^4 \colon x + y = z + t = 0\right\}.$$

Vogliamo determinare il sottospazio $V^\perp$. Osserviamo che i vettori

$$V_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix} \quad \text{e} \quad V_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix}$$

costituiscono una base di V. Quindi $\dim V = 2$ e dal teorema precedente

$$\dim V^\perp = \dim \mathbb{R}^4 - \dim V = 4 - 2 = 2.$$

Troviamo le equazioni di $V^\perp$:

$$Y \in V \iff \langle Y, V_1\rangle = \langle Y, V_2\rangle = 0,$$

calcolando i prodotti scalari otteniamo le equazioni cartesiane di $V^\perp$:

$$x - y = z - t = 0.$$

5. Il Teorema spettrale

Ci proponiamo ora di capire in che modo la struttura metrica introdotta in $\mathbb{R}^n$ possa influire sullo studio degli operatori lineari di $\mathbb{R}^n$. In particolare, ricordiamo che un operatore lineare L di $\mathbb{R}^n$ è diagonalizzabile se e solo se esiste una base di $\mathbb{R}^n$ formata da autovettori di L. In tale base l'operatore si rappresenta con una matrice diagonale. Poiché nei paragrafi precedenti abbiamo visto l'utilità di lavorare con una base ortonormale, è naturale chiedersi: *quando esiste una base ortonormale di $\mathbb{R}^n$ formata da autovettori di L?*

Sia $L \colon \mathbb{R}^n \to \mathbb{R}^n$ un operatore lineare che, nella base canonica $\mathcal{B} = \{\boldsymbol{e}_1, .., \boldsymbol{e}_n\}$, si rappresenta con la matrice $A \in \mathcal{M}_{\mathbb{R}}(n)$:

$$L(X) = AX, \quad \forall X \in \mathbb{R}^n.$$

Supponiamo che esista una base ortonormale $\mathcal{B}' = \left\{Y^1, .., Y^n\right\}$ di $\mathbb{R}^n$ formata da autovettori di L; l'operatore L nella base ortonormale $\mathcal{B}'$ si rappresenta con una matrice Δ diagonale:

$$L(X') = \Delta.X',$$

e l'equazione del cambiamento di coordinate $X = MX'$ è realizzato dalla matrice invertibile M le cui colonne sono le coordinate dei vettori della base $\mathcal{B}'$ nella base $\mathcal{B}$, cioè

$$M = (Y^1|..|Y^n).$$

Poiché le colonne di M sono una base ortonormale, possiamo concludere che la matrice M è ortogonale. Le matrici A e Δ sono simili e risulta:

$$\Delta = M^{-1}AM = M^TAM, \quad M \in \mathcal{O}(n).$$

Le matrici che sono diagonalizzabili attraverso matrici ortogonali hanno la seguente proprietà:

LEMMA 7.20. *Sia $A \in \mathcal{M}_{\mathbb{R}}(n)$. Se esiste una matrice ortogonale $M \in \mathcal{O}(n)$ tale che $M^T AM$ è diagonale, allora A è simmetrica, cioè $A = A^T$.*

DIMOSTRAZIONE. Dalla relazione di similitudine, ricaviamo che

$$A = M\Delta M^T, \quad M \in \mathcal{O}(n).$$

Calcolando la trasposta di entrambi i membri dell'uguaglianza e ricordando che $(AB)^T = B^T.A^T$, otteniamo la tesi :

$$A^T = \left(M\Delta M^T\right)^T = M\Delta M^T = A.$$

□

In realtà vale il vicesersa, cioè tutte le matrici reali simmetriche di ordine n sono diagonalizzabili attraverso un matrice ortogonale. Questo risultato fondamentale è provato nel Teorema spettrale e vale in generale in campo complesso per una classe più ampia di matrici.

TEOREMA 7.21. ***Teorema Spettrale***
Fissato in $\mathbb{R}^n$ il prodotto scalare standard, sia A una matrice reale simmetrica di ordine n. Valgono le seguenti proprietà:
(1) *A ammette n autovalori reali, contati con le relative molteplicità;*
(2) *due autospazi associati ad autovalori distinti di A sono perpendicolari;*
(3) *esiste una base ortogonale $\mathcal{B}$ di $\mathbb{R}^n$ formata da autovettori di A.*

Dimostriamo solo l'enunciato (2), mentre faremo alcune considerazioni ed osservazioni relativamente agli altri punti (le dimostrazioni relative vengono riportate in Appendice per alcuni corsi).

DIMOSTRAZIONE. (dell'enunciato (2)) Siano $\alpha \neq \beta$ due autovalori di A, indichiamo con V_α e V_β gli autospazi associati. Proviamo che:

$$V_\alpha \perp V_\beta.$$

Siano $X \in V_\alpha$ e $Y \in V_\beta$ due autovettori, per la linearità del prodotto scalare e poiché $A.X = \alpha X$ risulta:

$$\begin{gathered}\alpha\langle X, Y\rangle = \langle \alpha X, Y\rangle = \langle AX, Y\rangle = \\ = (AX)^T.Y = X^T.A^T.Y = X^T.AY = \langle X, \beta Y\rangle = \beta\langle X, Y\rangle,\end{gathered}$$

per la linearità del prodotto scalare e poiché $A.Y. = \beta Y$. Uguagliando la prima e l'ultima espressione abbiamo:

$$\alpha\langle X, Y\rangle = \beta\langle X, Y\rangle,$$

cioè

$$(\alpha - \beta)\langle X, Y\rangle = 0.$$

Poiché $\alpha \neq \beta$ si ha necessariamente $\langle X, Y\rangle = 0$ e quindi

$$X \perp Y, \quad \forall \;\; X \in V_\alpha, \quad \forall \;\; Y \in V_\beta.$$

□

OSSERVAZIONE 7.23 (circa i punti (1) e (3)).

(1) Sia

$$p_A(t) = \det(A - tI_n)$$

il polinomio caratteristico di A, l'equazione caratteristica di A è la seguente:

$$p_A(t) = 0. \tag{7.30}$$

Ricordiamo che $\alpha \in \mathbb{R}$ è un autovalore di A con molteplicità algebrica $\mu(\alpha) = \mu \geq 1$ se e solo se α è una soluzione dell'equazione caratteristica di A con moltiplicità μ, cioè

$$\mu(\alpha) = \mu \iff p_A(t) = (t-\alpha)^{\mu} q(t), \quad q(\alpha) \neq 0.$$

Il teorema afferma che l'equazione caratteristica 7.30 ammette h soluzioni reali, $1 \leq h \leq n$:

$$\alpha_1, .., \alpha_h \in \mathbb{R},$$

e le relative molteplicità algebriche soddisfano le relazione:

$$\mu(\alpha_1) + ... + \mu(\alpha_h) = n.$$

Ciò equivale al fatto che il polinomio caratteristico è totalmente decomponibile in fattori lineari:

$$p_A(t) = (-1)^n (t-\alpha_1)^{\mu(\alpha_1)} ... (t-\alpha_h)^{\mu(\alpha_h)}.$$

Questa proprietà è sempre soddisfatta nel campo complesso, ma in generale non è soddisfatta per i polinomio nell'anello $\mathbb{R}[t]$: ad esempio, è ben noto che il polinomio di secondo grado $p(t) = t^2 + 1$ non si può scrivere come prodotto di due fattori lineari a coefficienti reali.

Osserviamo infine che la proprietà (1) è sicuramente vera per le matrici reali simmetriche di ordine 2. Consideriamo infatti una matrice reale simmetrica di ordine 2:

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

L'equazione caratteristica di A è la seguente:

$$t^2 - (a+d)t + (ad - b^2) = 0,$$

poiché il discriminante di tale equazione risulta non negativo:

$$\Delta = (a+d)^2 - 4(ad - b^2) = (a-d)^2 + 4b^2 \geq 0,$$

l'equazione ammette sempre soluzioni reali. In particolare, le soluzioni sono reali distinte se $a \neq d$ e $b \neq 0$, le soluzioni sono reali coincidenti se $a = d$ e $b = 0$. Osserviamo che in questo caso risulta $A = aI_2$.

La dimostrazione di questa proprietà per matrici di ordine $n \geq 3$ non è banale, si tratta di provare che ogni soluzione complessa dell'equazione caratteristica di A è un numero reale. Per fare ciò è necessario studiare le proprietà di autovalori e autovettori di una matrice complessa e introdurre una struttura metrica nello spazio $\mathbb{C}^n$.

(3) Osserviamo che il risultato è vero per matrici reali simmetriche di ordine 2. Infatti, abbiamo osservato prima che l' equazione caratteristica di una matrice reale simmetrica di ordine 2 ha due soluzioni distinte se $A \neq aI_2$. In quest'ultimo caso la matrice è diagonale e ogni base ortonormale di $\mathbb{R}^2$ è formata da autovettori di A. Se $A \neq aI_2$, la matrice ha due autovalori reali semplici, α e β, quindi è diagonalizzabile. I due autospazi V_α e V_β sono rette per O e per la proprietà (2) sono ortogonali. Risulta quindi $\mathbb{R}^2 = V_\alpha \oplus V_\beta$, scegliendo un versore in ciascun autospazio si ottiene una base ortonormale di $\mathbb{R}^2$.

In generale, se A è una matrice reale simmetrica di ordine $n \geq 3$, come possiamo *determinare una base ortonormale* in $\mathbb{R}^n$ formata da autovettori di A?

Siano α_1, α_2,..,α_h gli autovalori distinti di A, $h \leq n$ e V_{α_1}, V_{α_2},.., V_{α_h} gli autospazi associati. Ci sono alcuni fatti importanti di cui tenere conto:

(1) $V_{\alpha_1} \oplus V_{\alpha_2} \oplus ... \oplus V_{\alpha_h} = \mathbb{R}^n$,

(2) $V_{\alpha_i} \perp V_{\alpha_j}$, $\forall i \neq j$.

Per il teorema di Gram-Schimdt, in ogni autospazio esiste una base ortogonale $\mathcal{B}(V_{\alpha_i})$, $i = 1, 2, ..h$. Per la proprietà (1) l'unione di tali basi

$$\mathcal{B} = \mathcal{B}(V_{\alpha_1}) \cup \mathcal{B}(V_{\alpha_2}) \cup ... \cup \mathcal{B}(V_{\alpha_h})$$

è una base di $\mathbb{R}^n$, formata da autovettori di A. Per la proprietà (2), due autospazi associati ad autovalori distinti sono ortogonali: possiamo concludere che la base $\mathcal{B}$ così ottenuta è ortogonale. Infine, normalizzando i vettori di tale base si ottiene una base ortonormale di $\mathbb{R}^n$ formata da autovettori di A.

Vale la pena di studiare un esempio.

Esempio 7.24.

Consideriamo la matrice reale simmetrica di ordine 3:

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

Vogliamo costruire una base ortogonale di $\mathbb{R}^3$ formata da autovettori di A. L' equazione caratteristica di A è la seguente:

$$p_A(\alpha) = |A - \alpha I_3| = -\alpha(4-\alpha)^2 = 0,$$

le cui radici sono:

$$\alpha = 0 \quad \mu(0) = 1 \quad \alpha = 4 \quad \mu(4) = 2.$$

Sia V_0 l'autospazio associato all'autovalore $\alpha = 0$:

$$V_0 = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \colon A \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0 \right\},$$

V_0 è una retta per l'origine di $\mathbb{R}^3$, di equazioni $x + z = y = 0$.

Sia V_4 l'autospazio associato all'autovalore $\alpha = 4$:

$$V_4 = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \colon A \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 4 \begin{pmatrix} x \\ y \\ z \end{pmatrix} \right\},$$

V_4 è un piano per l'origine di $\mathbb{R}^3$, di equazione $x - z = 0$.
Risulta $V_0 \oplus V_4 = \mathbb{R}^3$, una base ortogonale di $\mathbb{R}^3$ formata da autovettori di A si ottiene unendo una base ortogonale di V_0 ed una base ortogonale di V_4. Osserviamo che i seguenti vettori:

$$Y_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \quad Y_3 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

sono due vettori ortogonali di V_4, quindi costituiscono una base ortogonale di V_4. Una base ortogonale di V_0 si ottiene prendendo un qualsiasi vettore non nullo in V_0: sia

$$Y_1 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}.$$

La base $\{Y_1, Y_2, Y_3\}$ è una base ortogonale di $\mathbb{R}^3$ formata da autovettori di A.

Seguono immediatamente dal teorema spettrale i seguenti corollari:

Corollario 7.22. *Una matrice reale simmetrica A di ordine n è diagonalizzabile.*

Corollario 7.23. *Sia A una matrice reale simmetrica di ordine n. Esiste una matrice ortogonale $M \in O(n)$ che diagonalizza A:*

$$M^{-1}AM = M^T AM = \mathrm{diag}(\alpha_1, \alpha_2, ..., \alpha_n),$$

con $\mathrm{Spec}(A) = \{\alpha_1, \alpha_2, ..., \alpha_n\}$.

APPENDICE

6. Appendice: Matrici ortogonali di ordine 3

Anzitutto, ricordiamo che una matrice $Q \in \mathcal{O}(3)$ costituisce un cambiamento di base dalla base canonica ad una nuova base ortonormale. Alternativamente, possiamo vedere la matrice Q come quella che rappresenta un'operatore lineare che manda i 3 vettori della base canonica in un nuovo insieme di vettori ortonormali. Siccome la matrice ortogonale, e, quindi, l'applicazione linerare conservano il prodotto scalare, conservano anche le distanze: in altre parole, la matrice ortogonale descrive un operatore che realizza una trasformazione rigida sui vettori dello spazio $\mathbb{E}^3_O$, cioè una rotazione eventualmente composta con un'inversione. È facile convincersi del fatto che una rotazione deve conservare l'orientamento reciproco dei vettori della nuova base ortonormale, ossia corrisponde a matrici ortogonale con determinante $+1$, o, come abbiamo imparato a dire, matrici ortogonali speciali ($\mathcal{SO}(3)$).

Un interessante teorema si presta ad interpetazioni geometriche, ed è utilizzato in meccanica: la sua formulazione originaria risale al matematico svizzero Euler.

TEOREMA 7.24 (di Eulero). *Sia $Q \in \mathcal{SO}(3)$ una matrice ortogonale speciale di ordine 3. Allora Q ha sempre un autovalore $\lambda = +1$.*

DIMOSTRAZIONE. Sino ad ora abbiamo considerato le matrici ortogonali come sottoinsieme di $\mathcal{GL}(n, \mathbb{R})$, quindi a entrate reali. Poiché $\mathbb{R} \subset \mathbb{C}$, possiamo comunque pensare l'operatore lineare associato a Q in campo complesso. In questo caso, possiamo affermare che la matrice Q ha tutti gli autovalori in $\mathbb{C}$, grazie al teorema fondamentale dell'algebra (0.8). Siano $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{C}$ gli autovalori, non necessariamente distinti. Sicuramente, $\det(Q) = \lambda_1\lambda_2\lambda_3$, per il teorema 6.11 visto nel Capitolo sugli autovalori 6.

Inoltre, un corollario del teorema fondamentale (0.10), poiché il grado del polinomio caratteristico $p_Q(\lambda)$ è 3, quindi dispari, esiste almeno una radice del polinomio reale, e, quindi, almeno un autovalore è reale; supponiamo che questo sia λ_1.

In campo reale, allora, esiste un autovettore $\boldsymbol{v}$ con autovalore λ_1; ossia, se consideriamo l'operatore lineare L_Q associato a Q nella base canonica, abbiamo

$$L_Q(\boldsymbol{v}) = \lambda_1 \boldsymbol{v}\,; \tag{7.31}$$

In base all'Osservazione 7.20, già sappiamo che $(\lambda_1)^2 = 1$, ossia $\lambda_1 = 1$ oppure $\lambda_1 = -1$. A questo punto, se $\lambda_1 = 1$, abbiamo finito; altrimenti, se $\lambda = -1$, procediamo come segue.

Completiamo una base per $\mathbb{R}^3$ tenendo $\boldsymbol{v}_1$ come primo vettore; mediante un algoritmo di Gram-Schmidt, produciamo da essa una base ortonormale $\mathcal{B}' = \{\boldsymbol{w}'_1, \boldsymbol{w}'_2, \boldsymbol{w}'_3\}$, con $\boldsymbol{w}'_1 = \frac{\boldsymbol{v}}{||\boldsymbol{v}||}$.

Esisterà, dunque, una matrice ortogonale N che permette di passare dalla base canonica $\mathcal{B}_0$ a $\mathcal{B}'$ (Proposizione 7.14). La matrice che rappresenta l'operatore L_Q nella nuova base sarà

$$Q' = N^{-1}QN = N^{\mathrm{T}}QN = \left(\begin{array}{c|cc} \lambda_1 & a & b \\ \hline 0 & r_{11} & r_{12} \\ 0 & r_{21} & r_{22} \end{array}\right), \tag{7.32}$$

dove la prima colonna è

$$[L_{Q'}(\boldsymbol{w}'_1)]_{\mathcal{B}'} = [\lambda_1 \boldsymbol{w}'_1]_{\mathcal{B}'} = \begin{pmatrix} \lambda_1 \\ 0 \\ 0 \end{pmatrix},$$

ossia il trasformato del vettore $\boldsymbol{w}'_1$ (autovettore con autovalore λ_1), rappresentato nella base $\mathcal{B}' = \{\boldsymbol{w}'_1, \boldsymbol{w}'_2, \boldsymbol{w}'_3\}$.

Il resto della matrice ha una struttura relativamente semplice; anzitutto, chiamiamo

$$R = \begin{pmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{pmatrix} \tag{7.33}$$

la matrice 2×2 che occupa l'angolo in basso a destra.

Inoltre, osserviamo che termini a, b sulla prima riga devono essere nulli; infatti, l'inversa della matrice Q' dovrà avere la seguente struttura:

$$(Q')^{-1} = \left(\begin{array}{c|cc} \frac{1}{\lambda_1} & \alpha & \beta \\ \hline 0 & \multicolumn{2}{c}{\multirow{2}{*}{R^{-1}}} \\ 0 & & \end{array}\right), \tag{7.34}$$

con α, β da determinare (convincersene verificando il prodotto di Q' della (7.32) con la matrice $(Q')^{-1}$ della (7.34)).

Nel nostro caso, in aggiunta, la matrice Q' è ortogonale, essendo ottenuta dal prodotto di matrici tutte ortogonali (ricordiamo che $\mathcal{O}(3)$ è un gruppo: Osservazione 7.17); pertanto:

$$(Q')^{-1} = \left(\begin{array}{c|cc} \frac{1}{\lambda_1} & \alpha & \beta \\ \hline 0 & \multicolumn{2}{c}{\multirow{2}{*}{R^{-1}}} \\ 0 & & \end{array}\right) = \left(\begin{array}{c|cc} \lambda_1 & 0 & 0 \\ \hline a & \multicolumn{2}{c}{\multirow{2}{*}{R^{T}}} \\ b & & \end{array}\right) = (Q')^{\mathrm{T}} \tag{7.35}$$

che mostra subito che $a = b = \alpha = \beta = 0$, e che $R \in \mathcal{O}(2)$.

Da ultimo, se calcoliamo il determinante di Q' secondo la prima colonna otteniamo

$$\det(Q') = \lambda_1 \det(R) = \det(Q) = 1\,, \tag{7.36}$$

dove abbiamo usato il fatto che le matrici Q e Q' sono simili, quindi il determinante non cambia.

Questo mostra che, quando $\lambda_1 = -1$, deve essere $\det(R) = -1$, cioè R è una matrice ortogonale di ordine 2 con determinante -1, quindi ha due autovalori distinti ± 1 (Proposizione 7.16), e risulta diagonalizzabile. In altre parole, gli altri due autovalori della matrice Q sono $\lambda_{2,3} = \pm 1$, quindi uno fra i due è l'autovalore cercato $\lambda = +1$. □

Conseguenza immediata dei ragionamenti fatti sopra è il seguente

Corollario 7.25. *Sia $Q \in \mathcal{SO}(3)$ una matrice ortogonale speciale di ordine 3. Questa descrive una rotazione di un opportuno angolo $\vartheta \in [0, 2\pi)$ (a meno di $2k\pi$, con $k \in \mathbb{Z}$) attorno ad un asse diretto come l'autovettore associato al suo autovalore $\lambda = 1$. Inoltre, la matrice è diagonalizzabile se e solo se $\vartheta = 0$ oppure $\vartheta = \pi$.*

Dimostrazione. Si costruisca la matrice Q' come nella dimostrazione del Teorema di Eulero (7.24); se $\lambda_1 = -1$, abbiamo visto che Q è diagonalizzabile, e la sua forma diagonale è
$$Q'' = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$
ossia una rotazione di $\vartheta = \pi$ attorno all'asse diretto come il terzo vettore della base, cioè l'autovettore di autovalore 1.

Se invece $\lambda_1 = 1$, la (7.36) ci dice che $\det(R) = 1$, ossia la matrice R ortogonale di ordine 2 ha determinante $+1$, quindi è diagonalizzabile *solo se* $\vartheta = 0$ o $\vartheta = \pi$; nel primo caso, si ha la rotazione nulla ($Q = I_3$), nel secondo, ancora una rotazione di $\vartheta = \pi$ attorno alla direzione di $\boldsymbol{w}_1'$. □

7. Prodotto hermitiano

Occorre ricorrere a qualche "trucchetto" con in numeri complessi per completare la dimostrazione del teorema spettrale. Il nocciolo della questione risiede nel fatto che una matrice quadrata (reale o complessa) ha tutti gli autovalori nel campo complesso.

Occorre però modificare un po' la nozione di prodotto scalare: se i nostri spazi vettoriali sono del tipo $\mathbb{C}^n$ sul campo $\mathbb{C}$, la definizione del prodotto scalare standard non godrebbe più di una proprietà fondamentale. Infatti non è più detto che la somma di quadrati sia nulla soltanto se tutte le basi sono nulle.

Osservazione 7.25. Sia $Z = \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix} \in \mathbb{C}^n$ un vettore colonna di numeri complessi (con $n > 1$): allora $Z^{\mathrm{T}} Z = 0$ **non implica** che $z_i = 0$ per ogni $i = 1, \dots n$.

Dimostrazione. Basta considerare $Z = \begin{pmatrix} 1 \\ i \\ 0 \\ \vdots \\ 0 \end{pmatrix}$, cioè un vettore con una componente pari a 1, un'altra pari ad i e tutte le altre –se ci sono– nulle.

Si ha
$$Z^{\mathrm{T}} Z = 1^2 + i^2 + 0 + \dots 0 = 1 - 1 = 0\,.$$
□

Come dobbiamo modificare la definizione? il nostro fallimento risiede nel fatto che il quadrato di un numero complesso **non** è necessariamente reale positivo, e neppure reale, in genere. Quello che resta reale positivo (o nullo) è il **modulo** del numero complesso, o anche il suo modulo quadro. Ricordando che, se $z \in \mathbb{C}$, $|z|^2 = z\overline{z}$, ci rendiamo conto che le cose possono andare meglio se, nel definire il prodotto scalare standard in campo complesso, introduciamo il complesso coniugato del vettore cononna Z. In dettaglio, le considerazioni fatte giustificano la seguente definizione.

Definizione 7.9. Dati due vettori $X = \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix}$, e $Y = \begin{pmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{pmatrix}$ dello spazio vettoriale $\mathbb{C}^n$ sul campo complesso $\mathbb{C}$, con $n \geq 1$, chiamiamo **prodotto scalare complesso standard** o, anche **prodotto hermitiano standard** dei vettori X e Y il seguente numero reale

$$\langle X, Y\rangle_{\mathbb{C}} = x_1\overline{y_1} + x_2\overline{y_2} + \ldots + x_n\overline{y_n} = \ X^T \cdot \overline{Y}.$$

Con questa definizione, è facile mostrare che il prodotto hermitiano standard gode di proprietà molto vicine a quelle del prodotto scalare standard, definito in spazi vettoriali reali.

Proposizione 7.26. *Il prodotto scalare standard di $\mathbb{C}^n$ gode delle seguenti proprietà:*

(1) ***Proprietà di hermitianità:***

$$\langle X, Y\rangle_{\mathbb{C}} = \overline{\langle Y, X\rangle_{\mathbb{C}}} \quad \forall X, Y \in \mathbb{C}^n;$$

(2) ***Proprietà di linearità/coniugatolinearità:***

$$\langle aX_1 + bX_2, Y\rangle_{\mathbb{C}} = a\langle X_1, Y\rangle_{\mathbb{C}} + \ b\langle X_2, Y\rangle_{\mathbb{C}}, \quad \forall X_1, X_2, Y \in \mathbb{C}^n, \forall a, b \in \mathbb{C};$$

$$\langle X, aY_1 + bY_2\rangle_{\mathbb{C}} = \overline{a}\langle X, Y_1\rangle_{\mathbb{C}} + \ \overline{b}\langle X, Y_2\rangle_{\mathbb{C}}, \quad \forall X, Y_1, Y_2 \in \mathbb{C}^n, \forall a, b \in \mathbb{C};$$

(3) ***Proprietà di positività:***

$$\langle X, X\rangle_{\mathbb{C}} \geq 0 \quad \forall X \in \mathbb{C}^n,$$

$$\langle X, X\rangle_{\mathbb{C}} = 0 \iff X = \mathbf{0}_n.$$

Dimostrazione. Verifichiamo le proprietà.
(1) Siano $X, Y \in \mathbb{C}^n$ qualunque; osserviamo che risulta:

$$\begin{aligned} \overline{\langle Y, X\rangle_{\mathbb{C}}} &= \overline{Y^T \cdot \overline{X}} = \overline{y_1\overline{x_1} + y_2\overline{x_2} + \ldots + y_n\overline{x_n}} \\ &= \overline{y_1}\,\overline{\overline{x_1}} + \overline{y_2}\,\overline{\overline{x_2}} + \ldots + \overline{y_n}\,\overline{\overline{x_n}} \\ &= x_1\overline{y_1} + x_2\overline{y_2} + \ldots + x_n\overline{y_n} \\ &= X^T \cdot \overline{Y} = \langle X, Y\rangle_{\mathbb{C}} \end{aligned} \tag{7.37}$$

(2) Basta provare la prima equazione; la seconda discende da questa e dalla precedente proprietà.

Siano X_1, X_2 e Y vettori di $\mathbb{C}^n$; osserviamo che dalle proprietà del prodotto matriciale risulta:

$$\begin{aligned} \langle aX_1 + bX_2, Y\rangle_{\mathbb{C}} &= \ (aX_1 + bX_2)^T \cdot \overline{Y} = (a\ X_1^T + b\ X_2^T) \cdot \overline{Y} \\ &= a\ X_1^T \cdot \overline{Y} + b\ X_2^T \cdot \overline{Y} \\ &= a\langle X_1, Y\rangle_{\mathbb{C}} + b\langle X_2, Y\rangle_{\mathbb{C}}. \end{aligned} \tag{7.38}$$

Ora, usando questo risultato e anche la prima proprietà:

$$\begin{aligned}\langle X, aY_1 + bY_2\rangle_{\mathbb{C}} &= \overline{\langle aY_1 + bY_2, X\rangle_{\mathbb{C}}} = \overline{a\langle Y_1, X\rangle_{\mathbb{C}} + b\langle Y_2, X\rangle_{\mathbb{C}}} \\ &= \overline{a}\overline{\langle Y_1, X\rangle_{\mathbb{C}}} + \overline{b}\overline{\langle Y_2, X\rangle_{\mathbb{C}}} = \overline{a}\langle X, Y_1\rangle_{\mathbb{C}} + \overline{b}\langle X, Y_2\rangle_{\mathbb{C}}\end{aligned} \tag{7.39}$$

(3) Sia $X \in \mathbb{C}^n$, si ha:

$$\langle X, X\rangle_{\mathbb{C}} = \ X^T \cdot \overline{X} = |x_1|^2 + |x_2|^2 + ... + |x_n|^2 \geq 0,$$

inoltre $\langle X, X\rangle_{\mathbb{C}} = 0$ se e solo se $x_1 = x_2 = ... = x_n = 0$, cioè X è il vettore nullo di $\mathbb{C}^n$, poiché ogni termine $|x_i|$, per $i = 1, \dots n$ è un numero **reale**. □

Alla luce della definizione di prodotto scalare standard complesso 7.9, occorre rivedere quale legame sussiste tra matrici e prodotti scalari; nella dimostrazione del secondo punto del teorema spettrale, abbiamo visto che una matrice reale simmetrica può essere applicata ad uno dei vettori di un prodotto scalare o all'altro, producendo lo stesso risultato.

Proprietà 7.27. *Siano $X, Y \in \mathbb{R}^n$ qualunque e $A \in \mathcal{M}_{\mathbb{R}}(n)$ simmetrica; allora*

$$\langle AX, Y\rangle = \langle X, AY\rangle .$$

Dimostrazione. Infatti,

$$\langle AX, Y\rangle + (AX)^{\mathrm{T}}Y = X^{\mathrm{T}}A^{\mathrm{T}}Y = X^{\mathrm{T}}AY = \langle X, AY\rangle ,$$

sfruttando il fatto che A è simmetrica, dunque $A = A^{\mathrm{T}}$. □

Ma se lavoro in spazi complessi? non basta la simmetria: occorre che una matrice che coincida con la trasposta coniugata (detta anche *hermitiana*).

Definizione 7.10. Una matrice $A \in \mathcal{M}_{\mathbb{C}}(n)$ si dice *hermitiana* se

$$\overline{A^{\mathrm{T}}} = A .$$

Notiamo però che per le matrici reali le due cose coincidono:

Osservazione 7.26. Se una matrice reale è simmetrica, allora è anche hermitiana. Infatti, il complesso coniugato di un numero reale è il numero reale stesso (la parte immaginaria è nulla!), quindi

$$\overline{A^{\mathrm{T}}} = A^{\mathrm{T}} = A .$$

Allora la Prorietà 7.27 viene leggermente modificata nel campo complesso:

Proprietà 7.28. *Siano $X, Y \in \mathbb{C}^n$ qualunque e $A \in \mathcal{M}_{\mathbb{C}}(n)$ hermitiana; allora*

$$\langle AX, Y\rangle_{\mathbb{C}} = \langle X, AY\rangle_{\mathbb{C}} .$$

Dimostrazione. Infatti abbiamo:

$$\langle X, AY\rangle_{\mathbb{C}} = X^{\mathrm{T}}\,\overline{AY} = X^{\mathrm{T}}\,\overline{A}\,\overline{Y} = (\overline{A^{\mathrm{T}}}X)^{\mathrm{T}}\,\overline{Y} = (AX)^{\mathrm{T}}\,\overline{Y} = \langle AX, Y\rangle ,$$

sfruttando il fatto che $(\overline{A^{\mathrm{T}}})^{\mathrm{T}} = \overline{A}$, perché le operazioni di trasposizione e di coniugio si possono scambaire di ordine. □

Lasciamo allo studente il compito di verificare che valgono anche gli inversi delle Proprietà 7.27 e 7.28, ossia se per ogni scelta di vettori nello spazio i due prodotti scalari coincidono, allora la matrice nel prodotto è simmetrica o hermitiana, rispettivamente.

Siamo ora in grado di convincerci del primo punto del teorema spettrale

Dimostrazione del punto 1 del Teorema Spettrale 7.21. Siccome $\mathcal{M}_{\mathbb{R}}(n) \subset \mathcal{M}_{\mathbb{C}}(n)$, possiamo considerare la matrice A, reale e simmetrica, come una matrice di $\mathcal{M}_{\mathbb{C}}(n)$; in base all'Osservazione 7.26, possiamo dire che A è hermitiana.

La matrice A risulta quindi rappresentare un'operatore lineare di $\mathbb{C}^n$; in campo complesso la matrice ha n autovalori λ_i, on $i = 1, \ldots n$, non necessariamente distinti (la somma delle rispettive molteplicità ammonterà in tal caso a n esattamente): è quanto stabilito dal Corollario 6.15 del Capitolo 6.

Sia $\alpha = \lambda_i$, uno qualunque degli autovalori, e indichiamo con $Z \in \mathbb{C}^n$ un suo autovettore:

$$AZ = \alpha Z\,;$$

mostriamo che $\alpha \in \mathbb{R}$. Infatti, l'hermitianità di A e le proprietà 7.26 del prodotto scalare complesso standard consentono di scrivere:

$$\begin{aligned} \alpha\langle Z, Z\rangle_{\mathbb{C}} &= \langle \alpha Z, Z\rangle_{\mathbb{C}} = \langle AZ, Z\rangle_{\mathbb{C}} \\ &= \langle Z, AZ\rangle_{\mathbb{C}} = \langle Z, \alpha Z\rangle_{\mathbb{C}} \\ &= \overline{\alpha}\langle Z, Z\rangle_{\mathbb{C}}\,; \end{aligned} \tag{7.40}$$

il primo e l'ultimo termine della catena di uguaglianze permettono di scrivere

$$(\alpha - \overline{\alpha})\langle Z, Z\rangle_{\mathbb{C}} = 0\,.$$

Ricordando che un autovettore **non può essere il vettore nullo**, quindi $Z \neq \mathbf{0}_n$, e, allora, la positività del prodotto hermitiano ci dice che $\langle Z, Z\rangle_{\mathbb{C}} \neq 0$, deduciamo che

$$\alpha = \overline{\alpha}\,,$$

ossia, che $\alpha \in \mathbb{R}$.

La matrice A ha dunque n autovalori, tutti reali, contando le corrispettive molteplicità. $\square$

CAPITOLO 8

Forme quadratiche e loro applicazioni.

1. Forme quadratiche

In questo paragrafo introdurremo la nozione di forma quadratica definita sullo spazio vettoriale $\mathbb{R}^n$. Una prima applicazione fondamentale dello studio delle forme quadratiche è la classificazione delle coniche e delle quadriche di cui ci occuperemo nel prossimo paragrafo. Tuttavia gli studenti incontreranno nel loro percorso scolastico molte altre applicazioni, ad esempio in analisi nello studio dei punti di massimo e di minimo di una funzione reale di più variabili reali. Cominciamo quindi formalizzando la nozione di forma quadratica.

Definizione 8.1. Un'applicazione $Q\colon \mathbb{R}^n \to \mathbb{R}$ è detta **forma quadratica reale** su $\mathbb{R}^n$ se $Q(X)$ è un polinomio omogeneo di secondo grado a coefficienti reali nelle componenti $x_1, x_2, .., x_n$ del vettore X di $\mathbb{R}^n$.

N.B. Il termine forma reale su $\mathbb{R}^n$ significa che l'applicazione è definita su $\mathbb{R}^n$ ed ha valori in $\mathbb{R}$!!

Esempio 8.1.

(1) Una forma quadratica reale su $\mathbb{R}^2$ ha l'espressione seguente:
$$Q(X) = q_1x_1^2 + q_2x_1x_2 + q_3x_2^2,$$
con $q_1, q_2, q_3 \in \mathbb{R}$, $X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2$.

(2) Una forma quadratica reale su $\mathbb{R}^3$ ha ha l'espressione seguente:
$$Q(X) = q_1x_1^2 + q_2x_2^2 + q_3x_3^2 + q_4x_1x_2 + q_5x_1x_3 + q_6x_2x_3,$$
con $q_i \in \mathbb{R}$, $i = 1, .., 6$, $X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3$.

(3) Fissato il prodotto scalare standard in $\mathbb{R}^n$, la funzione
$$Q(X) = \|X\|^2 = \langle X, X\rangle = x_1^2 + x_2^2 + ... + x_n^2,$$
con $X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$, è una forma quadratica reale su $\mathbb{R}^n$, detta *forma quadratica definita dal prodotto scalare standard* di $\mathbb{R}^n$.

Gli esempi precedenti suggeriscono di introdurre, per ogni forma quadratica reale su $\mathbb{R}^n$, la seguente notazione:

$$Q(X) = \sum q_{ij}x_ix_j, \quad q_{ij} \in \mathbb{R}, \quad \forall\ 1 \leq i \leq j \leq n. \tag{8.1}$$

Infatti $Q(X)$ è un polinomio omogeneo di secondo grado nelle componenti $x_1, .., x_n$ del vettore $X \in \mathbb{R}^n$. Consideriamo lo spazio vettoriale reale $\mathbb{R}[x_1, ., x_n]$ dei polinomi a coefficienti reali nelle variabili $x_1, .., x_n$ ed il sottoinsieme

$$\mathcal{Q}_2 \subset \mathbb{R}[x_1, ., x_n]$$

dei polinomi omogenei di grado 2. Basta osservare che $\mathcal{Q}_2$ è un sottospazio di $\mathbb{R}[x_1, ., x_n]$ e che i monomi

$$\{x_i.x_j\}\,, \quad 1 \leq i \leq j \leq n$$

costituiscono una base di $\mathcal{Q}_2$.

Osservazione 8.2. Dalla definizione data di forma quadratica seguono immediatamente le proprietà seguenti:

(1) $Q(\mathbf{0}_{\mathbb{R}^n}) = 0$;
(2) $Q(\alpha X) = \alpha^2 Q(X)$, $\forall X \in \mathbb{R}^n$, $\forall \alpha \in \mathbb{R}$.

Verifichiamo la proprietà (2). Osseviamo che risulta, $\forall \alpha \in \mathbb{R}$:

$$Q(\alpha X) = \sum q_{ij}(\alpha x_i).(\alpha x_j) = \sum q_{ij}\alpha^2 x_i.x_j = \alpha^2 Q(X).$$

In particolare, osserviamo che una forma quadratica Q non è un' applicazione lineare!!!

Come abbiamo fatto per le applicazioni lineari di $\mathbb{R}^n$, ci chiediamo se è possibile rappresentare in forma matriciale una forma quadratica. Cominciamo ad esaminare nei dettagli le forme quadratiche di $\mathbb{R}^2$.

Esempio 8.3.
Abbiamo osservato che il prodotto scalare standard di $\mathbb{R}^2$ definisce la forma quadratica:

$$Q(X) = \|X\|^2 = \langle X, X\rangle = X^{\mathrm{T}}X = x_1^2 + x_2^2.$$

- Osserviamo che se calcoliamo il prodotto scalare dei vettori X e $A.X$, dove A è una qualsiasi matrice quadrata reale di ordine 2, otteniamo ancora una forma quadratica. Infatti, sia $A = (A^1|A^2)$ una matrice quadrata di ordine 2, risulta:

$$AX = x_1A^1 + x_2A^2,$$

per la linearità del prodotto scalare si ottiene:

$$\langle X, AX\rangle = x_1\langle X, A^1\rangle + x_2\langle X, A^2\rangle.$$

Posto $A^i = \begin{pmatrix} a_{1i} \\ a_{2i} \end{pmatrix}$, $i = 1, 2$, calcolando i prodotti scalari si ottiene:

$$\langle X, AX\rangle =$$

$$x_1(a_{11}x_1 + a_{21}x_2) + x_2(a_{12}x_1 + a_{22}x_2) = a_{11}x_1^2 + (a_{12} + a_{21})x_1x_2 + a_{22}x_2^2\,,$$

cioè un polinomo omogeneo di secondo grado nelle variabili x_1 e x_2. Possiamo quindi concludere che la matrice A definisce una forma quadratica reale su $\mathbb{R}^2$, con la seguente espressione:

$$Q_A(X) = \langle X, AX \rangle = X^{\mathrm{T}}AX.$$

- Consideriamo ora una forma quadratica non identicamente nulla su $\mathbb{R}^2$:

$$Q(X) = q_{11}x_1^2 + q_{12}x_1x_2 + q_{22}x_2^2,$$

con $q_{11}, q_{12}, q_{22} \in \mathbb{R}$, $X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2$.

Ci chiediamo se sia possibile rappresentare in forma matriciale Q, ci proponiamo quindi di determinare se esiste un'unica matrice quadrata A reale di ordine 2, per cui si possa scrivere:

$$Q(X) = X^{\mathrm{T}}AX.$$

Sia $A = (a_{ij})$, $i, j = 1, 2$ una matrice reale quadrata di ordine 2, per quanto visto prima risulta

$$Q_A(X) = \langle X, AX \rangle = X^{\mathrm{T}}AX = a_{11}x_1^2 + (a_{12} + a_{21})x_1.x_2 + a_{22}x_2^2;$$

uguagliando le due espressioni di $Q(X)$ e $Q_A(X)$ otteniamo:

$$q_{11}x_1^2 + q_{12}x_1x_2 + q_{22}x_2^2 = a_{11}x_1^2 + (a_{12} + a_{21})x_1x_2 + a_{22}x_2^2,$$

$\forall X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2$. Otteniamo quindi il seguente sistema lineare nelle incognite a_{11}, a_{12},a_{21},a_{22}:

$$\begin{cases} a_{11} = q_{11} \\ a_{12} + a_{21} = q_{12} \\ a_{22} = q_{22} \end{cases} \tag{8.2}$$

che ammette infinite soluzioni. Infatti, i valori di a_{11} e a_{22} sono determinati, ma l'equazione lineare a 2 incognite

$$a_{12} + a_{21} = q_{12}$$

ammette infinite soluzioni. Tuttavia, se aggiungiamo al sistema 8.2 l'equazione

$$a_{12} = a_{21},$$

il sistema ammette l'unica soluzione:

$$a_{11} = q_{11} \quad a_{12} = a_{21} = \frac{1}{2}q_{12} \quad a_{22} = q_{22}.$$

Osserviamo che aggiungere l'equazione $a_{12} = a_{21}$ equivale a *limitarsi a considerare le matrici reali simmetriche* di ordine 2. Possiamo quindi concludere che esiste un'unica matrice A reale simmetrica di ordine 2 tale che

$$Q(X) = Q_A(X) = X^{\mathrm{T}}AX = \langle X, AX \rangle.$$

Abbiamo verificato che ogni forma quadratica su $\mathbb{R}^2$ è definita da un'unica matrice reale simmetrica di ordine 2.

La rappresentazione matriciale di una forma quadratica con una matrice reale simmetrica è un risultato fondamentale perché permette di applicare allo studio delle forme quadratiche le tecniche fornite dal teorema spettrale. È per questo motivo che scegliamo di rappresentare ogni forma quadratica con matrici reali simmetriche. Il risultato vale in generale per forme quadratiche su $\mathbb{R}^n$, con $n \geq 2$:

Teorema 8.1. *Sia $Q\colon \mathbb{R}^n \to \mathbb{R}$ una forma quadratica reale: esiste un' unica matrice $A \in \mathcal{M}_{\mathbb{R}}(n)$, reale simmetrica di ordine n che definisce la forma quadratica Q nella base standard di $\mathbb{R}^n$:*

$$Q(X) = \langle X, AX \rangle = X^{\mathrm{T}} AX. \tag{8.3}$$

Viceversa, ogni matrice reale simmetrica A di ordine n definisce una forma quadratica su $\mathbb{R}^n$ la cui espressione è data dalla formula 8.3.

La dimostrazione del teorema è simile al caso $n = 2$. Ricordiamo solo che se

$$Q(X) = \sum q_{ij} x_i x_j, \quad q_{ij} \in \mathbb{R}, \quad \forall\, 1 \leq i \leq j \leq n,$$

allora la matrice A reale simmetrica che definisce Q viene determinata nel seguente modo:

$$a_{ii} = q_{ii} \quad \forall i = 1, .., n,$$
$$a_{ij} = a_{ji} = \frac{1}{2} q_{ij}, \quad \forall i \neq j, \quad i, j = 1, ..n.$$

Esempio 8.4.

(1) Consideriamo la seguente forma quadratica su $\mathbb{R}^2$:
$$Q(X) = 5x^2 - 4xy + y^2.$$
Ci proponiamo di determinare l'espressione matriciale di Q:
$$Q(X) = X^{\mathrm{T}} AX.$$
Si ha: $a_{11} = 5$, $a_{12} = a_{21} = \frac{-4}{2} = -2$, $a_{22} = 1$. Quindi risulta:
$$Q(X) = X^{\mathrm{T}} \begin{pmatrix} 5 & -2 \\ -2 & 1 \end{pmatrix} X.$$

(2) Consideriamo la seguente forma quadratica su $\mathbb{R}^3$:
$$Q(X) = x^2 - 4xy - z^2 + 2xz + yz.$$
Ci proponiamo di determinare l'espressione matriciale di Q:
$$Q(X) = X^{\mathrm{T}} AX.$$
Si ha: $a_{11} = 1$, $a_{12} = a_{21} = \frac{-4}{2} = -2$, $a_{13} = a_{31} = \frac{2}{2} = 1$, $a_{22} = 0$, $a_{23} = a_{32} = \frac{1}{2}$, $a_{33} = -1$. Quindi
$$A = \begin{pmatrix} 1 & -2 & 1 \\ -2 & 0 & \frac{1}{2} \\ 1 & \frac{1}{2} & -1 \end{pmatrix}.$$

Osservazione 8.5. Sia $A \in \mathcal{M}_{\mathbb{R}}(n)$ una matrice reale simmetrica di ordine n, e $Q_A\colon \mathbb{R}^n \to \mathbb{R}$ la forma quadratica definita da A.

• $\forall i = 1,..,n$ risulta: $a_{ii} = Q_A(\boldsymbol{e}_i)$, dove $\boldsymbol{e}_i$ è l'$i-$esimo vettore della base canonica di $\mathbb{R}^n$. Infatti si ha:

$$Q_A(\boldsymbol{e}_i) = \langle \boldsymbol{e}_i, A\boldsymbol{e}_i \rangle = \langle \boldsymbol{e}_i, A^i \rangle = a_{ii}.$$

• Se $A = \alpha I_n$, allora $Q_A(X) = \alpha ||X||^2$. Infatti per la linearità del prodotto scalare si ha:

$$Q_A(X) = \langle X, \alpha X \rangle = \alpha \langle X, X \rangle.$$

• Se $A = \text{diag}\ (\alpha_1, \alpha_2, ..., \alpha_n)$ è una matrice diagonale reale di ordine n, allora

$$Q_A(X) = \alpha_1 x_1^2 + \alpha_2 x_2^2 + ... + \alpha_n x_n^2.$$

Infatti risulta

$$\begin{aligned} Q_A(X) = \langle X, AX \rangle = x_1 \langle X, \alpha_1 \boldsymbol{e}^1 \rangle + ... + x_n \langle X, \alpha_n \boldsymbol{e}^n \rangle = \\ = \alpha_1 x_1^2 + \alpha_2 x_2^2 + ... + \alpha_n x_n^2. \end{aligned}$$

Una matrice diagonale definisce quindi una forma quadratica molto semplice, perché è combinazione lineare di monomi che sono quadrati. (i.e. è priva di termini rettangolari).

In molti casi è utile studiare il segno di una forma quadratica. Ad esempio, per studiare se un punto $X_0 \in \mathbb{R}^n$ è un massimo o un minimo per una funzione $f\colon \mathbb{R}^n \to \mathbb{R}$, si studia il segno della seguente forma quadratica

$$Q(X) = \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_i}(X_0) x_i x_j.$$

A tale scopo introduciamo le definizioni seguenti:

Definizione 8.2. Sia $Q\colon \mathbb{R}^n \to \mathbb{R}$ una forma quadratica reale.

(1) Q è detta **definita positiva** se

$$Q(X) > 0 \quad \forall X \neq \mathbf{0}_{\mathbb{R}^n};$$

(2) Q è detta **definita negativa** se

$$Q(X) < 0 \quad \forall X \neq \mathbf{0}_{\mathbb{R}^n};$$

(3) Q è detta **semidefinita positiva** se

$$Q(X) \geq 0 \quad \forall X \in \mathbb{R}^n, \quad \exists X \neq \mathbf{0}_{\mathbb{R}^n} \mid Q(X) = 0;$$

(4) Q è detta **semidefinita negativa** se

$$Q(X) \leq 0 \quad \forall X \in \mathbb{R}^n, \quad \exists X \neq \mathbf{0}_{\mathbb{R}^n} \mid Q(X) = 0;$$

(5) Q è detta **non definita** se non si verificano i casi 1) 2) 3) 4).

ESEMPIO 8.6. Vediamo alcuni esempi.

(1) Consideriamo la seguente forma quadratica su $\mathbb{R}^2$:
$$Q(X) = x^2 + 4xy + y^2.$$
Si ha:
$$Q\begin{pmatrix}1\\0\end{pmatrix} = a_{11} = 1 > 0, \quad Q\begin{pmatrix}1\\-1\end{pmatrix} = -2 < 0.$$
Possiamo concludere che Q è non definita.

(2) Consideriamo la seguente forma quadratica su $\mathbb{R}^3$:
$$Q(X) = x^2 - 2xy + y^2 + 4z^2.$$
Osserviamo che:
$$Q(X) = (x-y)^2 + 4z^2 \geq 0, \quad \forall X \in \mathbb{R}^3.$$
Inoltre $Q\begin{pmatrix}1\\1\\0\end{pmatrix} = 0$, quindi possiamo concludere che Q è semidefinita positiva.

(3) Sia Q la forma quadratica associata al prodotto scalare standard di $\mathbb{R}^n$
$$Q(X) = \langle X, X\rangle = x_1^2 + x_2^2 + ... + x_n^2,$$
osserviamo che $Q(X) \geq 0$, $\forall X \in \mathbb{R}^n$ e $Q(X) = 0$ se e solo se $X = \mathbf{0}_{\mathbb{R}^n}$, quindi Q è definita positiva.

(4) Sia $A = \text{diag}\,(\alpha_1, \alpha_2, ..., \alpha_n)$ una matrice diagonale reale di ordine n e Q_A la forma quadratica da essa definita:
$$Q_A(X) = \alpha_1 x_1^2 + \alpha_2 x_2^2 + ... + \alpha_n x_n^2.$$
Lo studio del segno di Q in questo caso è molto semplice. Infatti:
   (a) Q è definita positiva $\iff$ $\alpha_i > 0$, $\forall i = 1,..,n$;
   (b) Q è definita negativa $\iff$ $\alpha_i < 0$, $\forall i = 1,..,n$;
   (c) Q è semidefinita positiva $\iff$ $\alpha_i \geq 0$, $\forall i = 1,..,n$;
   (d) Q è semidefinita negativa $\iff$ $\alpha_i \leq 0$, $\forall i = 1,..,n$.

Proviamo la (a). Ovviamente se $\alpha_i > 0$, $\forall i$, allora per ogni $X \neq \mathbf{0}_{\mathbb{R}^n}$ si ha:
$$Q(X) = \alpha_1 x_1^2 + \alpha_2 x_2^2 + ... + \alpha_n x_n^2 > 0,$$
essendo somma di quantità positive. Supponiamo ora che Q sia definita positiva, $\forall i = 1,..,n$ risulta
$$\alpha_i = Q(\boldsymbol{e}_i) > 0.$$

Le altre proprietà si provano in modo analogo.

OSSERVAZIONE 8.7. Sia $Q(X) =< X, AX >= X^{\mathrm{T}}AX$ una forma quadratica reale su $\mathbb{R}^n$ definita dalla matrice A reale simetrica. Sia $M \in \mathcal{GL}(n, \mathbb{R})$ una matrice invertibile che realizza il cambio di coordinate: $X = MX'$. L'espressione di Q nelle nuove coordinate X' è la seguente:

$$Q(X') = (MX')^{\mathrm{T}}A(MX) = X'^{\mathrm{T}}(M^{\mathrm{T}}AM)X.$$

La matrice $M^{\mathrm{T}}AM$ è la matrice associata a Q nella base $\mathcal{B}'$ del sistema di coordinale X'.

Come nel caso delle applicazioni lineari, data una forma quadratica Q ci chiediamo se sia possibile cambiare il sistema di coordinate per avere un' espressione piú semplice di Q, che ci possa aiutare nello studio del suo segno. Tale risultato è un'importante applicazione del teorema spettrale ed è noto come processo di *riduzione a forma canonica di una forma quadratica.*

TEOREMA 8.2. *Sia $Q(X)$ una forma quadratica reale su $\mathbb{R}^n$ definita dalla matrice reale simmetrica A. Esiste una matrice ortogonale $O \in O(n)$, che realizza il cambio di variabili $X = OX'$, tale che nelle coordinate X' la forma quadratica ha la seguente espressione:*

$$Q(X') = \alpha_1 {x'_1}^2 + \alpha_2 {x'_2}^2 + \dots + \alpha_n {x'_n}^2,$$

dove $\{\alpha_1, \alpha_2, \dots, \alpha_n\}$ sono gli autovalori di A.
Tale espressione viene detta ***forma canonica*** *della forma quadratica Q ed il sistema di coordinate corrispondente è detto* ***riferimento canonico*** *di Q. La forma canonica è unica a meno di una permutazione degli autovalori di A.*

DIMOSTRAZIONE. Sia $Q(X) = X^{\mathrm{T}}AX$, per il corollario del teorema spettrale esiste una matrice ortogonale $O \in \mathcal{O}(n)$ che diagonalizza A:

$$O^{-1}AO = O^{\mathrm{T}}AO = \Delta = \mathrm{diag}(\alpha_1, \alpha_2, \dots, \alpha_n),$$

con $\mathrm{Spec}(A) = \{\alpha_1, \alpha_2, \dots, \alpha_n\}$.
Consideriamo il cambio di coordinate $X = OX'$, nelle coordinate X' la forma quadratica Q si scrive

$$Q(X') = X'^{\mathrm{T}}(O^{\mathrm{T}}AO)X' = X'^{\mathrm{T}}\Delta X' = \alpha_1 {x'_1}^2 + \alpha_2 {x'_2}^2 + \dots + \alpha_n {x'_n}^2.$$

□

La nuova base di $\mathbb{R}^n$ del sistema di coordinate X' è una base ortonormale di $\mathbb{R}^n$ formata da autovettori di A.

ESEMPIO 8.8.
Consideriamo la forma quadratica su $\mathbb{R}^2$:

$$Q(x, y) = x^2 - 4xy + y^2.$$

Vogliamo determinare una forma canonica di Q. Sia $A \in \mathcal{M}_{\mathbb{R}}(2)$ la matrice reale simmetrica associata a Q:

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

L'equazione caratteristica di A è la seguente:

$$p_A(\alpha) = \alpha^2 - 2\alpha - 3 = 0,$$

le cui radici sono $\alpha = -1$ e $\alpha = 3$, entrambe semplici. Una forma canonica di Q è la seguente:

$$Q(x', y') = -x'^2 + 3y'^2.$$

Osserviamo che Q non è definita.

La forma canonica è utile per lo studio del segno di una forma quadratica:

Proposizione 8.3. *Siano* $Q(X) = X^{\mathrm{T}}AX$ *una forma quadratica su* $\mathbb{R}^n$ *e* $\mathrm{Spec}(A) = \{\alpha_1, \alpha_2, .., \alpha_n\}$.

(1) *Q è definita positiva* $\iff$ $\alpha_i > 0$, $\forall i$;
(2) *Q è semidefinita positiva* $\iff$ $\alpha_i \geq 0$, $\forall i$;
(3) *Q è definita negativa* $\iff$ $\alpha_i < 0$, $\forall i$;
(4) *Q è semidefinita negativa* $\iff$ $\alpha_i \leq 0$, $\forall i$.

Per la dimostrazione vedi Esempio 8.6 (4).

Corollario 8.4. *Sia* $Q(X) = X^{\mathrm{T}}AX$ *una forma quadratica su* $\mathbb{R}^n$ *definita dalla matrice* A. *Se* Q *è definita positiva allora* $\det A > 0$ *e* $\mathrm{tr}(A) > 0$.

Dimostrazione.
Basta osservare che $\det(A)$ e $\mathrm{tr}(A)$ sono invarianti per similitudine, inoltre per una matrice diagonale $\Delta = \mathrm{diag}(\alpha_1, \alpha_2, .., \alpha_n)$ si ha:

$$\det(\Delta) = \alpha_1 \cdot \alpha_2 \cdot ....\alpha_n \quad tr(\Delta) = \alpha_1 + \alpha_2 + ...\alpha_n.$$

Essendo definita positiva, risulta $\alpha_i > 0$, $\forall i = 1, .., n$, e quindi $\det(\Delta) > 0$ e $\mathrm{tr}(\Delta) > 0$. □

Esempio 8.9.
La condizione data dal corollario è sufficiente a garantire che Q sia definita positiva solo se $n = 2$.
Consideriamo la forma quadratica su $\mathbb{R}^3$:

$$Q(X) = 3x^2 - y^2 - z^2,$$

la cui matrice associata è la matrice

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

con $\det A = 3 > 0$ e $\mathrm{tr}(A) = 1 > 0$. Tuttavia Q è non definita.

Può essere utile per lo studio del segno di una forma quadratica il seguente criterio:

Proposizione 8.5. *(Criterio dei minori incapsulati):*
Sia $Q(X) = X^{\mathrm{T}}AX$ una forma quadratica su $\mathbb{R}^n$. Per ogni $i = 1, ., n$, indichiamo con Δ_i il determinante del minore di ordine i di A costituito dalle prime i righe e i colonne di A.

Q è definita positiva se e solo se $\Delta_i > 0$, per ogni $i = 1, ., n$.

Esempio 8.10.
Consideramo la seguente matrice reale simmetrica, al variare del parametro reale k:

$$A = \begin{pmatrix} k & 1 & k \\ 1 & k & 1 \\ k & 1 & 1 \end{pmatrix}$$

Vogliamo determinare per quali valori di k la matrice A rappresenta una forma quadratica definita positiva. Siano $A_{[1]} = a_{11}$ e $A_{[2]}$ la sottomatrice di A data dalle prime due righe e colonne di A, si ha:

$$\Delta_1 = \det(A_{[1]}) = k \quad \Delta_2 = \det(A_{[2]}) = k^2 - 1 \quad \Delta_3 = \det(A) = -(k-1)^2(k+1).$$

Per il criterio dei minori incapsulati si ottiene il sistema di disequazioni:

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

che è immediato verificare non ha soluzioni reali.

2. Classificazione di coniche e quadriche

La principale applicazione del teorema spettrale consiste nella riduzione a forma canonica di coniche e quadriche e quindi nella loro classificazione.

2.1. Coniche.

Definizione 8.3. Fissato nel piano un sistema di riferimento cartesiano ortogonale $\mathcal{R}(O, \hat{\imath}, \hat{\jmath})$, una **curva algebrica** di ordine d é l'insieme $\mathcal{C}$ dei punti del piano le cui coordinate soddisfano un'equazione algebrica a coefficienti reali di ordine d:

$$\mathcal{C} = \{P = \begin{pmatrix} x \\ y \end{pmatrix} \mid \ f(x,y) = 0\}, \quad f(x,y) \in \mathbb{R}[x,y], \quad \deg f(x,y) \ = d.$$

Ció significa che $f(x,y)$ è un polinomio di grado d nelle varibili x e y con coefficienti in $\mathbb{R}$!

Osservazione 8.11.

(1) Le rette sono curve algebriche di ordine 1 nel piano.
(2) Esistono nel piano curve che non sono algebriche. Si consideri ad esempio l'insieme:
$$\mathcal{C} = \left\{ P = \begin{pmatrix} x \\ y \end{pmatrix} : y = \sin x \right\}.$$
(3) La proprietà che una curva $\mathcal{C}$ sia algebrica di ordine d non dipende dalla scelta del sistema di riferimento cartesiano fissato.

Definizione 8.4. Le curve algebriche di ordine 2 sono dette **coniche**.

Esempio 8.12.

(1) La circonferenza è una conica.
Infatti, fissato un sistema di riferimento cartesiano ortogonale nel piano $\mathcal{R}(O, \hat{\imath}, \hat{\jmath})$, siano $C = \begin{pmatrix} a \\ b \end{pmatrix}$ e $R > 0$, la circonferenza $\mathcal{C}$ di centro C e raggio R ha equazione:
$$(x-a)^2 + (y-b)^2 = R^2.$$
(2) L'ellisse è una conica.
Ricordiamo che dati due punti distinti $F \neq F'$ ed un numero reale $a > 0$, è detto **ellisse di fuochi** F **e** F' l'insieme dei punti P del piano che soddisfano la relazione
$$PF + PF' = 2a.$$
È immediato verificare che fissato un sistema di riferimento cartesiano ortogonale nel piano $\mathcal{R}(O, \hat{\imath}, \hat{\jmath})$ tale che l'asse x sia la retta FF' e l'asse y sia l'asse del segmento FF', la relazione precedente diventa:
$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1,$$

dove $c = OF$ e $b^2 = a^2 - c^2$. Tale equazione è detta **equazione canonica dell'ellisse** di fuochi F e F' ed il riferimento fissato **riferimento canonico**.

(3) L'iperbole è una conica.
Ricordiamo che dati due punti distinti $F \neq F'$ ed un numero reale $a > 0$, è detto **iperbole di fuochi** F **e** F' l'insieme dei punti P del piano che soddisfano la relazione

$$|PF - PF'| = 2a.$$

È immediato verificare che fissato un sistema di riferimento cartesiano ortogonalenel piano $\mathcal{R}(O, \hat{\imath}, \hat{\jmath})$ tale che l'asse x sia la retta FF' e l'asse y sia l'asse del segmento FF', la relazione precedente diventa:

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1,$$

dove $c = OF$ e $b^2 = c^2 - a^2$. Tale equazione è detta **equazione canonica dell'iperbole** di fuochi F e F' ed il riferimento fissato **riferimento canonico**.

(4) La parabola è una conica.
Ricordiamo che dati un punto F ed una retta d, con $F \notin d$, è detto **parabola di fuoco** F **e direttrice** d l'insieme dei punti P del piano che soddisfano la relazione

$$PF = \delta(P, d).$$

È immediato verificare che fissato un sistema di riferimento cartesiano nel piano $\mathcal{R}(O, \hat{\imath}, \hat{\jmath})$ tale che F appartenga all'asse x, l'asse y sia parallela a d ed infine $O(0, 0)$ appartenga alla parabola, la relazione scritta diventa:

$$y^2 = 2px,$$

dove $p = \delta(F, d)$. Tale equazione è detta **equazione canonica della parabola** di fuoco F e direttrice d ed il riferimento fissato **riferimento canonico**.

Un'equazione algebrica di secondo grado nelle variabili x, y non sempre rappresenta una curva nel senso intuitivo del termine, come risulta dai seguenti esempi.

Esempio 8.13.

(1) Fissato un sistema di riferimento cartesiano ortogonale $\mathcal{R}(O, \hat{\imath}, \hat{\jmath})$ nel piano, consideriamo la conica di equazione

$$x^2 - y^2 = 0.$$

Osserviamo che risulta $x^2 - y^2 = (x + y)(x - y)$, quindi la conica è l'unione delle due rette incidenti

$$l_1 \colon y = x \quad l_2 \colon y = -x.$$

(2) Fissato un sistema di riferimento cartesiano ortogonale $\mathcal{R}(O, \hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}})$ nel piano, consideriamo la conica di equazione

$$x^2 + y^2 = 0.$$

Osserviamo che l'equazione ammette l'unica soluzione reale

$$x = y = 0,$$

quindi la conica corrispondente si riduce ad un solo punto: l'origine O. Osserviamo inoltre che nell'anello $\mathbb{C}[x, y]$ risulta

$$x^2 - y^2 = (x + iy)(x - iy),$$

quindi la conica è l'unione di due rette complesse coniugate incidenti.

(3) Fissato un sistema di riferimento cartesiano ortogonale $\mathcal{R}(O, \hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}})$ nel piano, consideriamo la conica di equazione

$$x^2 - 1 = 0.$$

Poiché risulta $x^2 - 1 = (x + 1)(x - 1)$, la conica è l'unione delle due rette parallele

$$l_1 : x = 1 \quad l_2 : x = -1.$$

(4) Fissato un sistema di riferimento cartesiano ortogonale $\mathcal{R}(O, \hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}})$ nel piano, consideriamo la conica di equazione

$$x^2 + 1 = 0.$$

L'equazione non ha soluzioni reali, quindi la conica corrispondente non ha punti reali. Tuttavia nell'anello $\mathbb{C}[x, y]$ risulta

$$(x^2 + 1) = (x + i)(x - i),$$

quindi la conica è l'unione di due rette complesse coniugate parallele.

(5) Fissato un sistema di riferimento cartesiano ortogonale $\mathcal{R}(O, \hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}})$ nel piano, consideriamo la conica di equazione

$$x^2 = 0.$$

L'equazione ammette l'unica soluzione $x = 0$. I punti della conica corrispondente sono tutti e soli i punti della retta l di equazione $x = 0$. Possiamo vedere tale conica come il limite per $\epsilon \to 0$ della conica di equazione

$$x^2 - \epsilon = 0 \quad \epsilon > 0.$$

Diciamo allora che la conica è la retta l contata due volte.

Notazioni.

Fissato un sistema di riferimento cartesiano ortogonale nel piano $\mathcal{R}(O, \hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}})$, sia C la conica di equazione $f(x,y) = 0$. Poniamo:

$$f(x,y) = Q_2(x,y) + Q_1(x,y) + a_{33}, \tag{8.4}$$

dove $Q_i(x,y)$ è un polinomio omogeneo a coefficienti reali di grado i nelle variabili x, y, per $i = 1, 2$, e $a_{33} \in \mathbb{R}$. Quindi:
$Q_2(x,y)$ è una forma quadratica reale su $\mathbb{R}^2$, per cui può essere scritta in forma matriciale

$$Q_2(x,y) = (x \quad y) A_2 \begin{pmatrix} x \\ y \end{pmatrix}, \tag{8.5}$$

dove $A_2 = \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix}$ è una matrice reale simmetrica di ordine 2;
$Q_1(x,y)$ è la seguente applicazione lineare reale $\mathbb{R}^2 \to \mathbb{R}$:

$$Q_1(x,y) = 2a_{13}x + 2a_{23}y. \tag{8.6}$$

Associamo alla conica C la seguente matrice reale simmetrica di ordine 3:

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{pmatrix}; \tag{8.7}$$

infine poniamo

$$\Delta = \det(A). \tag{8.8}$$

Definizione 8.5. Fissato nel piano un sistema di riferimento cartesiano ortogonale $\mathcal{R}(O, \hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}})$, sia C la conica di equazione $f(x,y) = 0$.

(1) C è detta **non degenere** se $\Delta \neq 0$;
(2) C è detta **degenere** se $\Delta = 0$: in particolare è detta **semplicemente degenere** se $\operatorname{rg} A = 2$ e **doppiamente degenere** se $\operatorname{rg} A = 1$.

Si verifica che il rango di A non dipende dal sistema di riferimento cartesiano ortogonale scelto, di conseguenza la proprietà che una conica sia degenere non dipende dalla scelta del riferimento.

Osservazione 8.14.

Tutte le coniche descritte negli esempi 8.13 sono degeneri. Si verifica che coniche 1), 2), 3) e 4) sono semplicemente degeneri, mentre la conica 5) è doppiamente degenere.
Si può provare che C è una conica semplicemente degenere se e solo se C è l'unione di due rette reali (o immaginarie) incidenti o parallele, cioè il polinomio che defisce la conica si può scrivere come prodotto di due fattori lineari in $\mathbb{C}[x,y]$.
Infine, C è doppiamente degenere se e solo se è una retta contata due volte e quindi il polinomio che la definisce è un quadrato.

Esempio 8.15.

(1) Sia C un' ellisse di fuochi $F \neq F'$, C è una conica non degenere. Infatti la sua equazione canonica è

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1,$$

da cui ricaviamo la matrice reale simmetrica ad essa associata:

$$A = \begin{pmatrix} \frac{1}{a^2} & 0 & 0 \\ 0 & \frac{1}{b^2} & 0 \\ 0 & 0 & -1 \end{pmatrix},$$

risulta $\Delta = -\frac{1}{a^2b^2} \neq 0$. Osserviamo che nel riferimento canonico l'origine O è centro di simmetria per l'ellisse e gli assi cartesiani sono assi di simmetria ortogonali. Inoltre la matrice A_2 è diagonale.

(2) Sia C un' iperbole di fuochi $F \neq F'$, C è una conica non degenere. Infatti la sua equazione canonica è

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1,$$

da cui ricaviamo la matrice reale simmetrica ad essa associata:

$$A = \begin{pmatrix} \frac{1}{a^2} & 0 & 0 \\ 0 & -\frac{1}{b^2} & 0 \\ 0 & 0 & -1 \end{pmatrix},$$

risulta $\Delta = \frac{1}{a^2b^2} \neq 0$. Osserviamo che nel riferimento canonico l'origine O è centro di simmetria per l'iperbole e gli assi cartesiani sono assi di simmetria ortogonali. Inoltre la matrice A_2 è diagonale.

(3) Sia C una parabola di fuoco F e direttrice d, $F \notin d$, C è una conica non degenere. Infatti la sua equazione canonica è

$$y^2 = 2px,$$

da cui ricaviamo la matrice reale simmetrica ad essa associata:

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

risulta $\Delta = -p^2$. Osserviamo che la conica non ammette un centro di simmetria, e l'asse x è l'unico asse di simmetria ortogonale. Infine, la matrice A_2 è diagonale.

Osserviamo che nelle equazioni canoniche di ellisse, iperbole e parabola la forma quadratica Q_2 è in forma canonica (cioè la matrice associata A_2 è diagonale). Il primo passo nello studio di una conica è quindi quello di ridurre a forma canonica la forma quadratica Q_2, attraverso una trasformazione ortogonale delle coordinate. Il secondo passo è quello di semplificare la parte lineare attraverso una traslazione. Questo procedimento consente di classificare le coniche non degeneri e di ridurle a forma canonica. Il risultato che ora enunciamo fornisce la classificazione completa delle coniche non degeneri.

TEOREMA 8.6. *Sia C una conica non degenere nel piano euclideo, esiste un sistema di riferimento cartesiano ortogonale $\mathcal{R}(O, \hat{\imath}, \hat{\jmath})$ in cui C ha una delle seguente equazioni:*

(1) $\frac{X^2}{a^2} + \frac{Y^2}{b^2} = 1$, *la conica è un' ellisse reale;*
(2) $\frac{X^2}{a^2} + \frac{Y^2}{b^2} = -1$, *la conica è un' ellisse immaginaria;*
(3) $\frac{X^2}{a^2} - \frac{Y^2}{b^2} = 1$, *la conica è un'iperbole;*
(4) $Y^2 - 2pX = 0$, *la conica è una parabola.*

Tale riferimento è detto ***riferimento canonico*** *per la conica.*

OSSERVAZIONE 8.16. Nel caso dell'iperbole e della parabola, considereremo, equivalentemente, come riferimento canonico anche quello con gli assi X ed Y scambiati fra di loro, ossia se la conica C ha equazione $\frac{X^2}{a^2} - \frac{Y^2}{b^2} = -1$ se è un iperbole, o $X^2 - 2pY = 0$, se è una parabola.

Otteniamo quindi il seguente criterio per riconoscere una conica:

COROLLARIO 8.7. *Fissato nel piano un sistema di riferimento cartesiano ortogonale $\mathcal{R}(O, \hat{\imath}, \hat{\jmath})$, sia C una conica non degenere a punti reali di equazione*

$$Q_2(x, y) + Q_1(x, y) + q_0 = 0.$$

Sia A_2 la matrice reale simmetrica di ordine 2 associata alla forma quadratica $Q_2(x, y)$.

C è un'ellisse se e solo se $|A_2| > 0$;
C è un'iperbole se e solo se $|A_2| < 0$;
C è una parabola se e solo se $|A_2| = 0$.

ESEMPIO 8.17.

(1) Fissato nel piano un sistema di riferimento cartesiano ortogonale $\mathcal{R}(O, \hat{\imath}, \hat{\jmath})$, si consideri la conica C di equazione:

$$2x^2 - 4x - y^2 - 6y - 8 = 0.$$

Le matrici A ed A_2 associate alla conica sono le seguenti:

$$A = \begin{pmatrix} 2 & 0 & -2 \\ 0 & -1 & -3 \\ -2 & -3 & -8 \end{pmatrix} \qquad A_2 = \begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix}.$$

Risulta $\Delta \neq 0$ e $|A_2| = -2$, per cui possiamo concludere che la conica è un' iperbole non degenere.

Per ridurla ad equazione canonica, osserviamo che la forma quadratica è in forma canonica, poiché manca il termine in xy, è sufficiente quindi determinare la traslazione per portare gli assi nel centro della conica. A tal fine completiamo i quadrati:

$$2(x^2 - 2x + 1) - (y^2 + 6y + 9) = 8 + 2 - 9 = 1,$$

con la traslazione di equazione

$$\begin{cases} X = x - 1 \\ Y = y + 3 \end{cases}$$

si ottiene l'equazione canonica della conica:

$$2X^2 - Y^2 = 1.$$

(2) Fissato nel piano un sistema di riferimento cartesiano ortogonale $\mathcal{R}(O, \hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}})$, si consideri la conica C di equazione:

$$2x^2 + 2xy + 2y^2 - 1 = 0.$$

Le matrici A ed A_2 associate alla conica sono le seguenti:

$$A = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & -1 \end{pmatrix} \qquad A_2 = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.$$

Risulta $\Delta \neq 0$ e $|A_2| = 3$, per cui possiamo concludere che la conica è un' ellisse non degenere.

Osserviamo che nell'equazione mancano i termini lineari in x e y, per cui la conica ammette l'origine come centro di simmetria. È sufficiente quindi ridurre $Q_2(x, y)$ a forma canonica. A tal fine calcoliamo gli autovalori di A_2:

$$\det(A_2 - \lambda I_2) = 0 \quad \to \lambda_1 = 1 \quad \lambda_2 = 3.$$

La forma canonica della forma quadratica è la seguente

$$Q_2(X, Y) = X^2 + 3Y^2,$$

e quindi l'equazione canonica della conica è la seguente

$$X^2 + 3Y^2 = 1.$$

Le equazioni del cambio di coordinate sono le seguenti:

$$\begin{pmatrix} x \\ y \end{pmatrix} = M. \begin{pmatrix} X \\ Y \end{pmatrix},$$

dove M è una matrice ortogonale che diagonalizza A_2. Ricordiamo che gli autospazi associati agli autovalori $\lambda_1 = 1$ e $\lambda_2 = 3$ sono le seguenti rette ortogonali:

$$x + y = 0 \quad x - y = 0.$$

Una base ortonormale di autovettori di A_2 è la seguente:

$$\boldsymbol{v}_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \quad \boldsymbol{v}_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

per cui la matrice M è la seguente:

$$M = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}.$$

(3) Fissato nel piano un sistema di riferimento cartesiano ortogonale $\mathcal{R}(O, \hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}})$, si consideri la conica C di equazione:

$$x^2 + 2xy + y^2 - 4x + 1 = 0.$$

Le matrici A ed A_2 associate alla conica sono le seguenti:

$$A = \begin{pmatrix} 1 & 1 & -2 \\ 1 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix} \qquad A_2 = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

Risulta $\Delta \neq 0$ e $|A_2| = 0$, per cui possiamo concludere che la conica è una parabola non degenere.

Per ridurla ad equazione canonica, dobbiamo innanzittutto ridurre a forma canonica la forma quadratica $Q_2(x, y)$. A tal fine calcoliamo gli autovalori di A_2:

$$\det(A_2 - \lambda I_2) = 0 \quad \rightarrow \lambda_1 = 0 \quad \lambda_2 = 2.$$

La forma canonica della forma quadratica è la seguente

$$Q_2(x', y') = 2{y'}^2,$$

con il cambiamento di coordinate di equazione

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \cdot \begin{pmatrix} x' \\ y' \end{pmatrix}.$$

Nelle coordinate (x', y') l'equazione di C è la seguente:

$$2{y'}^2 - 2\sqrt{2}(x' + y') + 1 = 0.$$

Completiamo ora i quadrati

$$2({y'}^2 - \sqrt{2}y' + \frac{1}{2}) - 2\sqrt{2}x' = 1 - 1 = 0,$$

con la traslazione di equazione

$$\begin{cases} X = x' \\ Y = y' - \frac{1}{\sqrt{2}} \end{cases}$$

si ottiene l'equazione canonica della conica:

$$Y^2 = \sqrt{2}X.$$

2.2. Superfici quadriche.

Definizione 8.6. Fissato nello spazio un sistema di riferimento cartesiano ortogonale $\mathcal{R}(O, \hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}, \hat{\boldsymbol{k}})$, una **superficie algebrica** di ordine d é l'insieme $\mathcal{S}$ dei punti dello spazio le cui coordinate soddisfano un'equazione algebrica a coefficienti reali di ordine d:

$$\mathcal{S} = \{P = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mid f(x,y,z) = 0\ \}, \quad f(x,y,z) \in \mathbb{R}[x,y,z], \quad \deg f(x,y,z) = d.$$

Una superficie algebrica di ordine 2 è detta **superficie quadrica**.

Esempio 8.18.

(1) I piani sono superfici algebriche di ordine 1.
(2) La sfera è una superficie quadrica.
Fissato un sistema di riferimento cartesiano ortogonale $\mathcal{R}(O, \hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}, \hat{\boldsymbol{k}})$, sia $\mathcal{S}$ la sfera di centro $C = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$ e raggio R. L'equazione di $\mathcal{S}$ è la seguente
$$(x-a)^2 + (y-b)^2 + (z-c)^2 = R^2,$$
per cui $\mathcal{S}$ è una quadrica.
(3) La superficie ottenuta dalla rotazione di una conica non degenere attorno ad un suo asse di simmetria è una quadrica, detta **quadrica di rotazione**.
Sia C una parabola contenuta in un piano π, fissiamo il sistema di riferimento in modo tale che π sia il piano x, z e la parabola abbia equazione canonica:
$$x^2 = 2pz,$$
infine l'unico asse di simmetria di C è l'asse z. Vogliamo determinare la superficie $\mathcal{S}$ ottenuta dalla rotazione di C attorno all'asse z. Consideriamo un generico punto della parabola:
$$P = \begin{pmatrix} t \\ 0 \\ \frac{t^2}{2p} \end{pmatrix} \in C \qquad t \in \mathbb{R}.$$
Durante la rotazione P descrive una circonferenza, chiamata **parallelo** per P, contenuta nel piano per P ortogonale all'asse z, di equazione:
$$\begin{cases} x^2 + y^2 + (z - \frac{t^2}{2p})^2 = t^2 \\ z = \frac{t^2}{2p} \end{cases}$$
Poiché la superficie $\mathcal{S}$ è l'unione di tutti i paralleli, al variare di $P \in C$, per ottenere l'equazione di $\mathcal{Q}$ eliminiamo il parametro t:
$$2pz = x^2 + y^2.$$

Possiamo quindi concludere che $\mathcal{Q}$ è una quadrica, detta **paraboloide di rotazione**.
Sia C un' ellisse contenuta nel piano x, z di equazione

$$\frac{x^2}{a^2} + \frac{z^2}{b^2} = 1.$$

Si verifica che le superfici ottenute dalla rotazione di C rispettivamente attorno all'asse z e all'asse x sono le quadriche di equazione rispettivamente

$$\frac{x^2 + y^2}{a^2} + \frac{z^2}{b^2} = 1 \qquad \frac{x^2}{a^2} + \frac{y^2 + z^2}{b^2} = 1,$$

dette **ellissoidi di rotazione**.
Sia Sia C un' iperbole contenuta nel piano x, z di equazione

$$\frac{x^2}{a^2} - \frac{z^2}{b^2} = 1.$$

Si verifica che le superfici ottenute dalla rotazione di C rispettivamente attorno all'asse z e all'asse x sono le quadriche di equazione rispettivamente

$$\frac{x^2 + y^2}{a^2} - \frac{z^2}{b^2} = 1 \qquad \frac{x^2}{a^2} - \frac{y^2 + z^2}{b^2} = 1,$$

dette rispettivamente **iperboloide di rotazione a una falda** e **iperboloide di rotazione a due falde**. Osserviamo che la quadrica ottenuta dalla rotazione atorno all'asse trasverso di C ha due componenti connesse distinte.

Anche nel caso delle quadriche esistono equazioni che non rappresentano superfici nel senso intuitivo del termine.

Esempio 8.19.

(1) Fissato un sistema di riferimento cartesiano ortogonale $\mathcal{R}(O, \hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}, \hat{\boldsymbol{k}})$ sia $\mathcal{S}$ la quadrica di equazione:

$$x^2 + y^2 + z^2 = -1.$$

Poiché l'equazione non ha soluzioni reali, la quadrica non ha punti reali.

(2) Fissato un sistema di riferimento cartesiano ortogonale $\mathcal{R}(O, x, y, z)$ sia $\mathcal{S}$ la quadrica di equazione:

$$x^2 + y^2 - z^2 = 0.$$

La quadrica $\mathcal{S}$ è un cono circolare retto con vertice nell'origine O. Osserviamo che intersecando la quadrica con piani paralleli al piano $z = 0$ si ottengono le seguenti circonferenze:

$$\begin{cases} z = k \\ x^2 + y^2 = k^2 \end{cases}$$

Infine, l'equazione di $\mathcal{S}$ è omogenea di secondo grado in x, y, z, ciò significa che se $P \in \mathcal{S}$, allora la retta OP è interamente contenuta in $\mathcal{S}$.

(3) Fissato un sistema di riferimento cartesiano ortogonale $\mathcal{R}(O, \hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}, \hat{\boldsymbol{k}})$ sia $\mathcal{S}$ la quadrica di equazione:
$$x^2 + y^2 = 1.$$
La quadrica $\mathcal{S}$ è un cilindro circolare retto con generatrici parallele all'asse z. Osserviamo che intersecando la quadrica con piani paralleli al piano $z = 0$ si ottengono le seguenti circonferenze:
$$\begin{cases} z = k \\ x^2 + y^2 = 1 \end{cases}$$
Infine, l'equazione di $\mathcal{S}$ è priva della variabile z, ciò implica che se $P \in \mathcal{S}$, allora la retta per P parallela all'asse z è interamente contenuta in $\mathcal{S}$.

(4) Fissato un sistema di riferimento cartesiano ortogonale $\mathcal{R}(O, \hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}, \hat{\boldsymbol{k}})$ sia $\mathcal{S}$ la quadrica di equazione:
$$xy - xz = 0.$$
Osserviamo che risulta $xy - xz = x(y - z)$, quindi la quadrica è l'unione dei due piani incidenti di equazioni $x = 0$ e $y - z = 0$.

(5) Fissato un sistema di riferimento cartesiano ortogonale $\mathcal{R}(O, x, y, z)$ sia $\mathcal{S}$ la quadrica di equazione:
$$z^2 = 0.$$
L'equazione ammette l'unica soluzione $z = 0$. Come per le coniche, la quadrica è il piano $z = 0$ contato due volte.

Come nel caso delle coniche si puó associare ad una superficie quadrica una matrice reale simmetrica A di ordine 4.

Definizione 8.7. Una superficie quadrica $\mathcal{S}$ é **non singolare** se $detA \neq 0$, é **singolare** se $detA = 0$.

Come nel caso delle coniche, riducendo a forma canonica la forma quadratica $Q_2(x, y, z)$ che compare nell'equazione della superificie quadrica, eseguendo poi opportune traslazioni, è possibile ridurre a forma canonica le quadriche ed ottenere il seguente teorema di classificazione delle superficie quadriche.

Teorema 8.8. *Sia $\mathcal{S}$ una superficie quadrica non singolare nello spazio euclideo, a punti reali, esiste un sistema di riferimento cartesiano ortogonale $\mathcal{R}(O, \hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}, \hat{\boldsymbol{k}})$ in cui $\mathcal{S}$ ha una delle seguente equazioni:*

(1) $\frac{X^2}{a^2} + \frac{Y^2}{b^2} + \frac{Z^2}{c^2} = 1$, *la superficie quadrica é un ellissoide;*
(2) $\frac{X^2}{a^2} + \frac{Y^2}{b^2} - \frac{Z^2}{c^2} = 1$, *la superficie quadrica é un iperboloide a una falda;*
(3) $\frac{X^2}{a^2} - \frac{Y^2}{b^2} - \frac{Z^2}{c^2} = 1$, *la superficie quadrica è un iperboloide a due falde;*

(4) $\frac{X^2}{a^2}+\frac{Y^2}{b^2}=Z$, *la superficie quadrica é un paraboloide ellittico;*
(5) $\frac{X^2}{a^2}-\frac{Y^2}{b^2}=Z$, *la superficie quadrica é un paraboloide iperbolico.*

Tale riferimento è detto ***riferimento canonico*** *per la quadrica.*

Esempio 8.20.

(1) Fissato un sistema di riferimento cartesiano ortogonale $\mathcal{R}(O,x,y,z)$ sia $\mathcal{S}$ la quadrica di equazione:
$$z-2xy=0.$$
Per ridurre a forma canonica la forma quadratica
$$Q_2(x,y,z)=-2xy,$$
la cui matrice associata è la seguente
$$A_2=\begin{pmatrix}0&-1&0\\-1&0&0\\0&0&0\end{pmatrix}.$$
Gli autovalori di A_2 sono $\alpha_1=1$, $\alpha_2=-1$, $\alpha_3=0$, la forma canonica di Q_2 è la seguente
$$Q_2(X,Y,Z)=X^2-Y^2.$$
Gli autospazi associati sono i seguenti:
$$V_{\alpha_1}\colon x+y=z=0$$
$$V_{\alpha_2}\colon x-y=z=0$$
$$V_{\alpha_3}\colon x=y=0,$$
una base ortonormale formata da autovettori è la seguente:
$$v_1=\frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\\0\end{pmatrix}\quad v_2=\frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\\0\end{pmatrix}\quad v_3=\begin{pmatrix}0\\0\\1\end{pmatrix}.$$
Le equazioni del cambio di coordinate sono
$$\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{2}}&0\\-\frac{1}{\sqrt{2}}&\frac{1}{\sqrt{2}}&0\\0&0&1\end{pmatrix}\cdot\begin{pmatrix}X\\Y\\Z\end{pmatrix},$$
per cui l'equazione di $\mathcal{S}$ nelle coordinate (X,Y,Z) è la seguente
$$Z=Y^2-X^2.$$
Abbiamo trovato l'equazione canonica di $\mathcal{S}$, possiamo concludere che la quadrica è non singolare ed è un paraboloide iperbolico.

(2) Fissato un sistema di riferimento cartesiano ortogonale $\mathcal{R}(O, x, y, z)$ sia $\mathcal{S}$ la quadrica di equazione:

$$3x^2 + y^2 + 6z^2 + 4xz - 1 = 0.$$

Innanzitutto riduciamo a forma canonica la forma quadratica

$$Q_2(x, y, z) = 3x^2 + y^2 + 6z^2 + 4xz,$$

la cui matrice associata è la seguente

$$A_2 = \begin{pmatrix} 3 & 0 & 2 \\ 0 & 1 & 0 \\ 2 & 0 & 6 \end{pmatrix}.$$

Gli autovalori di A_2 sono $\alpha_1 = 1$, $\alpha_2 = 2$, $\alpha_3 = 7$, la forma canonica di Q_2 è la seguente

$$Q_2(X, Y, Z) = X^2 + 2Y^2 + 7Z^2.$$

Gli autospazi associati sono i seguenti:

$$V_{\alpha_1} : x = z = 0$$

$$V_{\alpha_2} : x + 2z = y = 0$$

$$V_{\alpha_3} : 2x - z = y = 0,$$

una base ortonormale formata da autovettori è la seguente:

$$v_1 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad v_2 = \frac{1}{\sqrt{5}} \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix} \quad v_3 = \frac{1}{\sqrt{5}} \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}.$$

Le equazioni del cambio di coordinate sono

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 & -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{pmatrix} \cdot \begin{pmatrix} X \\ Y \\ Z \end{pmatrix},$$

per cui l'equazione di $\mathcal{S}$ nelle coordinate (X, Y, Z) è la seguente

$$X^2 + 2Y^2 + 7Z^2 = 1.$$

Tale equazione è canonica, possiamo concludere che la quadrica è non singolare ed un ellissoide a punti reali.

(3) Fissato un sistema di riferimento cartesiano ortogonale $\mathcal{R}(O, x, y, z)$ sia $\mathcal{S}$ la quadrica di equazione:

$$4x^2 + y^2 + 2y - 6z^2 + 12z - 7 = 0.$$

Osserviamo che la forma quadratica

$$Q_2(x, y, z) = 3x^2 + y^2 - 6z^2,$$

è in forma canonica. Quindi per ridurre la quadrica a forma canonica basta completare i quadrati:

$$3x^2 + (y^2 + 2y + 1) - 6(z^2 - 2z + 1) = 7 + 1 - 6 = 2,$$

cioè

$$3x^2 + (y + 1)^2 - 6(z - 1)^2 = 2.$$

Con la traslazione di equazione

$$\begin{cases} X = x \\ Y = y + 1 \\ Z = z - 1 \end{cases}$$

si ottiene l'equazione canonica della quadrica:

$$\frac{X^2}{\frac{1}{2}} + \frac{Y^2}{2} - \frac{Z^2}{\frac{1}{3}} = 1.$$

Possiamo concludere che la quadrica è non singolare ed è un iperboloide a una falda.

CAPITOLO 9

Esercizi

Proponiamo alcuni esercizi sui vari argomenti presentati, e per alcuni di essi forniamo la soluzione o la risposta.

In generale, le risposte riportate in colore rosso sono univoche; quelle in colore ciano sono soggette ad un certo grado di arbitrarietà. Se una risposta è riportata in colore magenta è univocamente determinata, ma dipende dalle scelte fatte in precedenza.

1. Introduzione

Gli esercizi proposti in questa sezione si riferiscono ai contenuti del Capitolo 0, e vengono divisi in più sottosezioni per comodità di organizzazione degli argomenti.

2. Numeri complessi

Riportiamo alcuni esercizi proposti di recente nelle prove d'esame. Conveniamo *per le risposte* di indicare con $\mathrm{Arg}(Z)$ l'argomento principale del numero complesso z, ossia l'angolo ricondotto all'intervallo principale $(-\pi, \pi]$. Ciò eliminerà alcune arbitrarietà nelle risposte, per motivi di semplicità.

Per alcuni esercizi viene riportata la soluzione dettagliata; per altri, le sole risposte corrette.

Esercizio 2.1. (27 novembre 2008, prova in itinere)
Dato il numero complesso $w = 1+i\sqrt{3}$, calcolare: (a) $|w^4| =$: (b) $\mathrm{Arg}(w^4) =$
(c) $\Re(w^{-1}) =$ (d) $\Im(w^{-1}) =$

Risoluzione. Calcoliamo prima modulo e argomento di w:

$$|w| = \sqrt{1^2 + (\sqrt{3})^2} = \sqrt{4} = 2\,, \qquad \mathrm{Arg}(w) = \arctan(\sqrt{3}) = \frac{\pi}{3}\,. \tag{9.1}$$

Con la formula di De Moivre, otteniamo immediatamente

$$|w^4| = |w|^4 = 2^4 = 16 \qquad \mathrm{Arg}(w^4) = 4\,\mathrm{Arg}(w) = \frac{4}{3}\pi\,. \tag{9.2}$$

Inoltre, $w^{-1} = \overline{w}/|w|^2$; abbiamo, pertanto:

$$\Re(w^{-1}) = \Re(\frac{\overline{w}}{|w|^2}) = \Re(\frac{1 - i\sqrt{3}}{|w|^2}) = \frac{1}{4}, \tag{9.3}$$

e

$$\Im(w^{-1}) = \Im(\frac{\overline{w}}{|w|^2}) = \Im(\frac{1 - i\sqrt{3}}{|w|^2}) = -\frac{1}{4}\sqrt{3}. \tag{9.4}$$

Esercizio 2.2. (27 novembre 2008, prova in itinere)
Determinare i numeri complessi z che soddisfano le seguenti condizioni:

$$\begin{cases} |z| = 1 \\ \mathrm{Arg}(z^2) = \pi + 2k\pi \end{cases}$$

Risoluzione. La prima equazione dà direttamente il modulo dei numeri z cercati. La seconda equazione ci dice che

$$\mathrm{Arg}(z) = \frac{1}{2}\mathrm{Arg}(z^2) = \frac{\pi}{2} + k\pi \qquad k \in \mathbb{Z}. \tag{9.5}$$

Al variare di $k \in \mathbb{Z}$ otteniamo diversi argomenti, ma tutti sono equivalenti ai due argomenti principali

$$\vartheta_1 = \frac{\pi}{2} \qquad \vartheta_1 = \frac{3}{2}\pi; \tag{9.6}$$

I numeri $z_{1,2}$ cercati perciò sono:

$$z_1 = 1\,e^{i\frac{\pi}{2}} = i \qquad z_2 = 1\,e^{i\frac{3\pi}{2}} = -i \tag{9.7}$$

Esercizio 2.3. (4 febbraio 2009, appello)
Si considerino i numeri complessi $z = 1 + \sqrt{3}i$ e $w = 1 + i$. Si calcolino:
(a) $|z| = 2$ (b) $\mathrm{Arg}(w) = \frac{\pi}{4}$ (c) $\mathrm{Arg}(z^3 w) = \frac{5}{4}\pi$ (d) $|\frac{w^3}{z^2}| = \frac{\sqrt{2}}{2}$

Esercizio 2.4. (19 febbraio 2009, appello)
Si consideri il numero complesso $z = \frac{1-\sqrt{3}i}{\sqrt{3}+3i}$. Si calcolino:
(a) $|z| =$ (b) $\mathrm{Arg}(z) =$ (c) $z^{-1} =$ (d) $z^2 =$

Esercizio 2.5. (6 luglio 2009, appello)
Si consideri il numero complesso $z = (1-i)^{-1}$. Si calcolino:
(a) $|z| = \frac{\sqrt{2}}{2}$ (b) $\mathrm{Arg}(z) = \frac{\pi}{4}$ (c) $\Re(z^{-1}) = 1$ (d) $z^2 = \frac{i}{2}$

Esercizio 2.6. (16 settembre 2009, appello) Si considerino i numeri complessi $z = 1 + \sqrt{3}i$ e $w = 1 - i$. Si calcolino: (a) $z^3 = -8$ (b) $\mathrm{Arg}(zw) = \frac{\pi}{12}$ (c) $|zw^2| = 4$ (d) $w^{-1} = \frac{1}{2}(1+i)$

ESERCIZIO 2.7. (23 novembre 2009, appello straord.)
Determinare il numero $z \in \mathbb{C}$ tale che

$$\Re(z+i) = 2 \qquad \Re(1-iz) = -1 \qquad z = 2-2i.$$

Noto z, si calcolino:
(a) $\mathrm{Arg}(z^3) = -\frac{3}{4}\pi$ (b) $|2z^{-1}| = \frac{1}{\sqrt{2}}$ (c) $\Im(z^2) = -8$.

ESERCIZIO 2.8. (14 settembre 2010)
Considerare il numero complesso $z = (2-2i)^{-3} \in \mathbb{C}$. Determinare
(a) $\mathrm{Arg}(z) = \frac{3}{4}\pi$ (b) $|z| = \frac{\sqrt{2}}{32}$ (c) $\frac{1}{z} = -16(1+i)$ (d) $z^{-2} = 512i$.

ESERCIZIO 2.9. (11 febbraio 2010)
Considerare il numero complesso $z = \sqrt{2} - \sqrt{6}i$.
Sia $w \in \mathbb{C}$ tale che $|w| = \frac{|z|}{\sqrt{2}}$, e $\mathrm{Arg}(w) = \frac{\pi}{4}$. Determinare
(a) $w = \sqrt{2}(1+i)$ (b) $\mathrm{Arg}(z) = -\frac{\pi}{3}$ (c) $z^3w^2 = -64\sqrt{2}i$ (d) $\Re(w^3) = -4\sqrt{2}$

---

3. Vettori applicati in $\mathbb{E}^3_O$ e geometria nello spazio

Gli esercizi proposti in questa sezione si riferiscono ai contenuti del Capitolo 1.

ESERCIZIO 3.1. (27 novembre 2008, prova in itinere)
Fissata la base standard $\left\{\hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}, \hat{\boldsymbol{k}}\right\}$ di $\mathbb{E}^3_O$, sono dati i vettori: $\boldsymbol{u} = \hat{\boldsymbol{\imath}} - \hat{\boldsymbol{\jmath}}$ e $\boldsymbol{v} = \hat{\boldsymbol{k}} - \hat{\boldsymbol{\jmath}}$.
Determinare:
(1) Un versore $\hat{\boldsymbol{w}} \in \mathrm{Span}(\boldsymbol{u} - \boldsymbol{v})$:
(2) L'equazione cartesiana di $\mathrm{Span}(\boldsymbol{u}, \boldsymbol{v})$:
(3) Un vettore $\boldsymbol{n}$ ortogonale ai vettori $\boldsymbol{u}$ e $\boldsymbol{v}$:
(4) Un vettore $\boldsymbol{x} \in \mathrm{Span}(\boldsymbol{u}, \boldsymbol{v})$ ortogonale al vettore $\boldsymbol{v}$:
(5) Le coordinate del vettore $\boldsymbol{x}$ rispetto alla base $\{\boldsymbol{v}, \boldsymbol{n}, \boldsymbol{u}\}$:

RISOLUZIONE. Cominciamo con lo scrivere il vettore $\boldsymbol{w} = \boldsymbol{u} - \boldsymbol{v} = \hat{\boldsymbol{\imath}} - \hat{\boldsymbol{k}}$, per risolvere il punto (1); sicuramente, $\boldsymbol{w} \in \mathrm{Span}((\boldsymbol{u}, \boldsymbol{v})$. Per avere un versore, basta dividere $\boldsymbol{w}$ per il suo modulo:

$$\hat{\boldsymbol{w}} = \frac{\boldsymbol{w}}{|\boldsymbol{w}|} = \frac{\hat{\boldsymbol{\imath}} - \hat{\boldsymbol{k}}}{\sqrt{1+1}}, \tag{9.8}$$

ossia,

$$\hat{\boldsymbol{w}} = \frac{1}{\sqrt{2}}(\hat{\boldsymbol{\imath}} - \hat{\boldsymbol{k}}\,.)$$

Per il punto (2), scriviamo ora una generica combinazione lineare di $\boldsymbol{u}, \boldsymbol{v}$ secondo i coefficienti reali λ, μ; questo darà un vettore generico $\boldsymbol{x} \in \mathrm{Span}(\boldsymbol{u}, \boldsymbol{v})$ (osserviamo che questo insieme è il piano passante per l'origine che contiene i due vettori):

$$\boldsymbol{x} = \lambda \boldsymbol{u} + \mu \boldsymbol{v}\,; \tag{9.9}$$

riscriviamo la (9.9) secondo le componenti $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ di $\boldsymbol{x}$, e otteniamo le equazioni parametriche:

$$\begin{cases} x &= \lambda \\ y &= -\lambda - \mu \\ z &= \mu \end{cases} \tag{9.10}$$

L'eliminazione dei parametri dall'Eq. (9.10) per avere l'equazione cartesiana del piano generato da $\boldsymbol{u}$ e $\boldsymbol{v}$: otteniamo $y = -x - z$, cioè

$$x + y + z = 0\,.$$

L'osservazione fatta sopra, ossia che lo span dei due vettori è un piano, permette di risolvere immediatamente il punto (3): infatti, un vettore $\boldsymbol{n}$ ortogonale allo span deve essere ortogonale al piano, ossia diretto come la normale al piano, quindi le sue componenti devono essere proporzionali ai coefficienti delle variabili nell'equazione del piano. Pertanto,

$$\boldsymbol{n} = \hat{\boldsymbol{\imath}} + \hat{\boldsymbol{\jmath}} + \hat{\boldsymbol{k}}\,.$$

Per il punto (4), scriviamo ora la condizione che il vettore dell'Eq. (9.10) sia ortogonale a $\boldsymbol{v}$: per fare questo, basta scrivere che il prodotto scalare fra i due vettori è nullo. Usando l'espressione per il prodotto scalare espressa tramite le componenti, otteniamo

$$\langle \boldsymbol{v}, \boldsymbol{x} \rangle = v_x\, x + v_y\, y + v_z\, z = 0 + (-1)(-\lambda - \mu) + 1\mu = \lambda + 2\mu = 0. \tag{9.11}$$

Questo esprime il legame che deve sussistere tra i due parametri; scegliendo $\mu = -1$ abbiamo:

$$\begin{cases} x &= \lambda = -2\mu = 2 \\ y &= -\lambda - \mu = -(2) - (-1) = -1 \\ z &= \mu = -1 \end{cases}, \tag{9.12}$$

che consente di scrivere

$$\boldsymbol{x} = 2\hat{\boldsymbol{\imath}} - \hat{\boldsymbol{\jmath}} - \hat{\boldsymbol{k}}\,.$$

Resta ora da risolvere il punto (5); scriviamo il vettore $\boldsymbol{x}$ come combinazione lineare di vettori della base:

$$\boldsymbol{x} = \alpha \boldsymbol{v} + \beta \boldsymbol{n} + \gamma \boldsymbol{u}\,. \tag{9.13}$$

Osserviamo, anzitutto, che $\boldsymbol{x} \in \mathrm{Span}(\boldsymbol{u}, \boldsymbol{v})$ e $\boldsymbol{n} \perp \boldsymbol{u}, \boldsymbol{v}$; nell'Eq. (9.13) è immediato, quindi, ottenere che $\beta = 0$, poiché $\boldsymbol{x}$ non può avere componente lungo $\boldsymbol{n}$. Per trovare α e γ possiamo moltiplicare scalarmente l'equazione vettoriale (9.13) per due vettori della base standard, e ottenere due equazioni scalari. Se scegliamo di proiettare lungo $\hat{\boldsymbol{\imath}}$ e lungo $\hat{\boldsymbol{k}}$ le equazioni si risolvono immediatamente, poiché i due vettori della nuova base sono rispettivamente ortogonali a queste direzioni:

$$\langle \boldsymbol{x}, \hat{\boldsymbol{\imath}} \rangle = \alpha \langle \boldsymbol{v}, \hat{\boldsymbol{\imath}} \rangle + \gamma \langle \boldsymbol{u}, \hat{\boldsymbol{\imath}} \rangle = 0 + \beta = 2 \tag{9.14a}$$

$$\langle \boldsymbol{x}, \hat{\boldsymbol{k}} \rangle = \alpha \langle \boldsymbol{v}, \hat{\boldsymbol{k}} \rangle + \gamma \langle \boldsymbol{u}, \hat{\boldsymbol{k}} \rangle = \alpha + 0 = -1\,. \tag{9.14b}$$

Pertanto, le coordinate richieste sono:

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

Esercizio 3.2. (27 novembre 2008)

Fissato nello spazio un sistema di riferimento cartesiano ortogonale $\mathcal{R}(O, \hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}, \hat{\boldsymbol{k}})$, si considerino i punti $P_1 = (0, 1, -1)$, $P_2 = (1, 0, -1)$, e $P_3 = (0, 2, 0)$.

(1) Determinare l'equazione cartesiana del piano π passante per i punti P_1, P_2 e P_3:

(2) Dire quale/i fra i punti $Q_1 = (1, -1, 0)$, $Q_2 = (-1, 1, 0)$, $Q_3 = (1, 1, 0)$ e $Q_4 = (0, 0, -2)$ appartengono al piano π.

Risoluzione. Il piano richiesto è parallelo al piano passante per l'origine determinato da $\mathrm{Span}(\overrightarrow{P_1P_2}, \overrightarrow{P_1P_3})$. Cominciamo a determinare questo piano, scrivendo esplicitamente le combinazioni lineari dei due vettori

$$\overrightarrow{P_1P_2} = P_2 - P_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} = \hat{\boldsymbol{\imath}} - \hat{\boldsymbol{\jmath}} \quad \text{e} \quad \overrightarrow{P_1P_3} = P_3 - P_1 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = \hat{\boldsymbol{\jmath}} + \hat{\boldsymbol{k}}\,.$$

Otteniamo:

$$\begin{cases} x &= \lambda \\ y &= -\lambda + \mu \\ z &= \mu \end{cases} \tag{9.15}$$

da cui, eliminando i parametri $y = -x + z$. L'equazione del piano richiesto avrà, pertanto, la forma $x + y - z = \delta$. Per calcolare il valore di δ basta imporre il passaggio per uno qualunque dei tre punti: scegliendo P_3 abbiamo immediatamente per il piano π l'equazione:

$$x + y - z = 2\,. \tag{9.16}$$

Per il secondo quesito, basta semplicemente verificare se le coordinate dei punti soddisfano o meno l'equazione (9.16) del piano trovata. È facile convincersi che

$$Q_3, Q_4 \in \pi,$$

mentre gli altri due punti sono esterni al piano.

Esercizio 3.3. (27 novembre 2008)
Fissato nello spazio un sistema di riferimento cartesiano ortogonale $\mathcal{R}(O, \hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}, \hat{\boldsymbol{k}})$, si considerino il punto $P = (0, 2, 0)$ e la retta r di equazioni $y - z = y - x = 0$. Determinare:

(1) La direzione della retta r:
(2) L' equazione cartesiana del piano π_1 contenente la retta r ed il punto P:
(3) L'equazione cartesiana del piano π_2 ortogonale ad r e passante per P:
(4) L'angolo fra i piani π_1 e π_2:
(5) La distanza del punto P dalla retta r:

Risoluzione. Le equazioni della retta r possono essere scritte come

$$x = y \,, \tag{9.17a}$$

$$z = y \,, \tag{9.17b}$$

che permettono subito di scegliere y come parametro cartesiano; in altre parole, uguagliando y ad un parametro $\lambda \in \mathbb{R}$ scriveremo:

$$x = \lambda \,, \tag{9.18a}$$

$$y = \lambda \,, \tag{9.18b}$$

$$z = \lambda \,. \tag{9.18c}$$

Dalle Eq. (9.18) otteniamo immediatamente che la retta passa per l'origine, e un vettore che individua la sua direzione sarà

$$\boldsymbol{d} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} . \tag{9.19}$$

L'equazione del piano π_1 avrà la forma

$$\alpha\, x + \beta\, y + \gamma\, z = \delta \,; \tag{9.20}$$

osserviamo che il suo vettore normale $\boldsymbol{N} = \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix}$ deve essere ortogonale a $\boldsymbol{d}$, quindi il loro prodotto scalare deve essere nullo:

$$\langle \boldsymbol{N}, \boldsymbol{d} \rangle = \alpha + \beta + \gamma = 0 \,, \tag{9.21}$$

che ci fornisce una prima condizione per i valori delle componenti di $\boldsymbol{N}$.

Imponendo che il piano passi per il punto P otteniamo:

$$\alpha\, 0 + \beta\, 2 + \gamma\, 0 = 2\beta = \delta \,; \tag{9.22}$$

osservando che il piano deve contenere la retta r, quindi anche l'origine, imponiamo $\delta = 0$ e, quindi, dall'Eq. (9.22), $\beta = 0$. Nell'Eq. (9.21) possiamo prendere, quindi, $\alpha = 1$, da cui $\gamma = -1$. L'equazione del piano risulta essere, quindi:

$$x - z = 0. \tag{9.23}$$

La scrittura dell'equazione del piano π_2 è quasi immediata: imponendo il passaggio per il punto P e che il vettore $\boldsymbol{d}$ rappresenta un suo vettore normale, data la condizione imposta dal testo, abbiamo

$$1(x-0)+1(y-2)+1(z-0)=0\,, \tag{9.24}$$

ossia

$$x+y+z=2.$$

L'angolo fra i piani π_1 e π_2 è immediatamente determinato, senza bisogno di calcoli: poiché uno contiene la retta r e l'altro è ortogonale a questa, i due piani sono ortogonali, pertanto il loro angolo è

$$\vartheta=\frac{\pi}{2}.$$

Resta, infine, da determinare la distanza di P da r. Determiniamo il punto Q intersezione fra la retta r ed il piano π_2 ortogonale ad essa passante per P, di cui abbiamo determinato l'equazione (9.24); tale punto risolve il sistema

$$\begin{cases} y-z=0 \\ y-x=0 \\ x+y+z=2\,. \end{cases} \tag{9.25}$$

La soluzione del sistema è semplice, e fornisce $Q=(\frac{2}{3},\frac{2}{3},\frac{2}{3})$; la distanza richiesta coincide con la distanza fra P e Q:

$$d(P,Q)=\sqrt{(\frac{2}{3}-)^2+(\frac{2}{3}-2)^2+(\frac{2}{3}-0)^2}\,; \tag{9.26}$$

sviluppando i calcoli otteniamo

$$d(P,Q)=d(P,r)=2\frac{\sqrt{6}}{3}.$$

Esercizio 3.4. (4 febbraio 2008, appello)

Fissato nello spazio un sistema di riferimento cartesiano ortogonale $\mathcal{R}(O,\hat{\boldsymbol{\imath}},\hat{\boldsymbol{\jmath}},\hat{\boldsymbol{k}})$, si considerino il punto $P=(1,1,1)$ ed il piano π di equazione $x+2y+z=1$.

(1) Scrivere le eq. cartesiane della retta r ortogonale a π e passante per P:
(2) Scrivere l'eq. del piano τ contenente l'origine e la retta r:
(3) Scrivere l'eq. del piano σ passante per P parallelo a π:

Esercizio 3.5. (19 febbraio 2009, appello)

Fissato nello spazio un sistema di riferimento cartesiano ortogonale $\mathcal{R}(O,\hat{\boldsymbol{\imath}},\hat{\boldsymbol{\jmath}},\hat{\boldsymbol{k}})$, si considerino il punto $P=(1,-1,-1)$ ed i vettori $\boldsymbol{u}=\begin{pmatrix}2\\1\\0\end{pmatrix}$ e $\boldsymbol{v}=\begin{pmatrix}0\\1\\2\end{pmatrix}$.

(1) Scrivere l'equazione del piano $\pi=\operatorname{Span}(\boldsymbol{u},\boldsymbol{v})$.
(2) Calcolare la distanza δ di P da π.

(3) Scrivere le equazioni cartesiane della retta r passante per l'origine del riferimento ed il piede della perpendicolare a π passante per P.
(4) Scrivere l'equazione del piano che contiene P ed r.

ESERCIZIO 3.6. (6 luglio 2009, appello)
Fissato nello spazio un sistema di riferimento cartesiano ortogonale $\mathcal{R}(O, \hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}, \hat{\boldsymbol{k}})$, si considerino la retta r di equazioni $x = y$ e $z - 2y = 0$ ed il punto $P = (1, 1, h)$.
(1) Scrivere l' equazione del piano π passante per l'origine e ortogonale a r: $x + y + 2z = 0$.
(2) Determinare i valori del parametro h per i quali $P \in \pi$: $h = -1$.
(3) Trovare i punti sulla retta r distanti $\sqrt{6}$ dal piano π:
$P_1 = (1, 1, 2)$, $P_2 = (-1, -1, -2)$.

ESERCIZIO 3.7. (16 settembre 2009, appello)
Fissato nello spazio un sistema di riferimento cartesiano ortogonale $\mathcal{R}(O, \hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}, \hat{\boldsymbol{k}})$, si considerino i punti $A = (1, 1, 0)$, $B = (1, 0, 1)$ e $C = (1, h, 2)$.
(1) Scrivere le equazioni cartesiane della retta AB:
$$x = 1 \quad y + z = 1$$
(2) Stabilire per quale/i valori del parametro reale h i punti A, B e C sono allineati: $h = -1$
(3) Posto $h = 2$, scrivere l'equazione cartesiana del piano per i tre punti:
$$x = 1$$

ESERCIZIO 3.8. (16 settembre 2009, appello)
Fissato nello spazio un sistema di riferimento cartesiano ortogonale $\mathcal{R}(O, \hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}, \hat{\boldsymbol{k}})$, si considerino la sfera Σ di equazione
$$x^2 + y^2 + z^2 - 2x + 4y + z + 5 = 0$$
ed il piano π di equazione $x + z = 0$. Determinare:
(1) Le coordinate del centro C ed il raggio R della sfera:
$$C = (1, -2, -\frac{1}{2}) \quad R = \frac{1}{2}$$
(2) La distanza del centro C dal piano π:
$$\delta = \frac{\sqrt{2}}{4}$$
(3) Il raggio r della circonferenza che π taglia su Σ:
$$r = \frac{\sqrt{2}}{4}$$

ESERCIZIO 3.9. (23 novembre 2009, appello straord.)
Fissato nello spazio un sistema di riferimento cartesiano ortogonale $\mathcal{R}(O, \hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}, \hat{\boldsymbol{k}})$, si considerino i punti $A = (1, 1, 1)$ e $B = (2, -1, 3)$.

(1) Scrivere l'equazione del piano π passante per O e perpendicolare alla retta $r = AB$.
$$x - 2y + 2z = 0\,.$$
(2) Calcolare le distanze fra A e B, fra A e π, fra B e π:
$$d_{AB} = 3, \quad d_{A\pi} = \frac{1}{3}, \quad d_{B\pi} = \frac{10}{3}\,.$$
(3) Calcolare la distanza fra O e la retta AB:
$$d_{O,AB} = \frac{\sqrt{26}}{3}\,.$$

ESERCIZIO 3.10. (14 febbraio 2010)
Fissato nello spazio un sistema di riferimento cartesiano ortogonale $\mathcal{R}(O, \hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}, \hat{\boldsymbol{k}})$, si considerino i punti $A = \begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix}$, $B = \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix}$, e $P = \begin{pmatrix} -1 \\ 0 \\ -1 \end{pmatrix}$.

(1) Scrivere le equazioni cartesiane della retta r passante per A e B:
$$\begin{cases} x - y = 1 \\ y - z = 2 \end{cases}$$
(2) Scrivere le equazioni cartesiane della retta s passante per P e avente direzione $\boldsymbol{d} = \hat{\boldsymbol{\imath}} - \hat{\boldsymbol{k}}$: $\begin{cases} x + z = -2 \\ y = 0 \end{cases}$
(3) Precisare la posizione reciproca delle rette r ed s: rette SGHEMBE.
(4) Scrivere l'equazione del piano π che contiene A, B, e P: $\pi\colon x - 3y + 2z + 3 = 0$
(5) Calcolare la distanza dell'origine O del riferimento da π: $\delta = \frac{3}{\sqrt{14}}$.

ESERCIZIO 3.11. (14 settembre 2010)
Fissato nello spazio un sistema di riferimento cartesiano ortogonale $\mathcal{R}(O, \hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}, \hat{\boldsymbol{k}})$, si considerino i punti $A = (0, 2, 1)$, $B = (2, 0, -1)$.

(1) Determinare le equazioni parametriche della retta r passante per A e B:
$$\begin{cases} x = 1 + t \\ y = 1 - t \\ z = -t \end{cases} \qquad t \in \mathbb{R}$$
Determinare in forma cartesiana il piano π passante per l'origine $O = (0, 0, 0)$ ed ortogonale a r: $x - y - z = 0$

(2) Determinare il punto H di intersezione fra r e π: $H = (1, 1, 0)$
(3) Determinare in forma cartesiana le equazioni della retta OH:

$$x - y = z = 0$$

---

4. Spazi vettoriali

Gli esercizi proposti in questa sezione si riferiscono ai contenuti del Capitolo 2.

Esercizio 4.1. Considerati i seguenti sottospazi di $\mathbb{R}^3$

$$U = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \,|\, x - y = z = 0 \right\}$$

e

$$V = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \,|\, x + y = 0 \right\}$$

determinare:

(1) la dimensione ed una base di U e V;
(2) il sottospazio $U \cap V$;
(3) la sommma $U + V$. U e V sono in somma diretta?

Risoluzione. Le equazioni che definiscono gli elementi di U sono 2, e sono indipendenti. Scriviamole in forma parametrica; per fare questo, osserviamo che la coordinata z è fissata: $z = 0$. Scrivendo la prima equazione cartesiana nella forma $x = y$ risulta naturale scegliere come parametro una qualunque delle due coordinate:

$$x = \lambda\,, \tag{9.27a}$$

$$y = \lambda\,, \tag{9.27b}$$

$$z = 0\,, \tag{9.27c}$$

con $\lambda \in \mathbb{R}$. Essendoci un solo parametro libero, $\dim(U) = 1$. Per avere una base $\mathcal{B}_U$ di U, basta prendere $\lambda = 1$: il vettore ottenuto genera U.

$$\mathcal{B}_U = \left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \right\}.$$

Per V, z è libera; inoltre, $y = -x$: le equazioni parametriche sono:

$$x = \lambda\,, \tag{9.28a}$$

$$y = -\lambda\,, \tag{9.28b}$$

$$z = \mu\,, \tag{9.28c}$$

con $\lambda, \mu \in \mathbb{R}$. Con due parametri liberi, abbiamo $\dim(V) = 2$. Per avere una base $\mathcal{B}_V$ di V, basta prendere $\lambda = 1$ e $\mu = 0$, e poi $\lambda = 0$ e $\mu = 1$: i vettori ottenuti sono indipendenti, e quindi generano V

$$\mathcal{B}_V = \left\{ \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}.$$

Per il secondo punto, dobbiamo trovare le coordinate che soddisfano il sistema

$$\begin{cases} x - y & = 0 \\ z & = 0 \\ x + y & = 0 \end{cases}. \tag{9.29}$$

Sommando e sottraendo membro a membro la prima e l'ultima otteniamo

$$\begin{cases} 2x & = 0 \\ z & = 0 \\ 2y & = 0 \end{cases}, \tag{9.30}$$

cioè $x = y = z = 0$. Solo il vettore nullo si trova nell'intersezione.

La formula di Grassmann garantisce che

$$\dim(U) + \dim(V) = \dim(U + V) + \dim(U \cap V);$$

siccome $\dim(U \cap V) = 0$, $\dim(U + V) = 1 + 2 = 3$.

Ossia, $U + V$ è un sottospazio di $\mathbb{R}^3$ di dimensione 3; ma $\dim(\mathbb{R}^3) = 3$, quindi $U + V = \mathbb{R}^3$. La somma è diretta: $U + V = U \oplus V$.

Esercizio 4.2. Dati i sottospazi di $\mathbb{R}^4$:

$$U = \left\{ \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \in \mathbb{R}^4 \,|\, x - y = 0 \right\}$$

e

$$V = \left\{ \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \in \mathbb{R}^4 \,|\, x - t = z = 0 \right\}$$

determinare:

(1) la dimensione ed una base di U e V;
(2) la dimensione ed una base di $U + V$;
(3) la sommma $U \cap V$. U e V sono in somma diretta?

Risoluzione. Una sola equazione definisce gli elementi di U. Scriviamola in forma parametrica; per fare questo, osserviamo che le coordinate z e t sono libere; scrivendo l'equazione cartesiana nella forma $x = y$ risulta naturale scegliere come parametro una qualunque delle due coordinate:

$$x = \lambda\,, \tag{9.31a}$$

$$y = \lambda\,, \tag{9.31b}$$

$$z = \mu\,, \tag{9.31c}$$

$$t = \nu\,, \tag{9.31d}$$

con $\lambda, \mu, \nu \in \mathbb{R}$. Essendoci tre parametri liberi, $\dim(U) = 3$. Per avere una base $\mathcal{B}_U$ di U, basta prendere un parametro di valore unitario, e gli altri due nulli: i 3 vettori ottenuti generano U.

$$\mathcal{B}_U = \left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\}.$$

Come conseguenza, abbiamo, chiaramente, $\dim(U) = 3$.

Per V, y è libera; inoltre, $x = t$, e $z = 0$. Le equazioni parametriche che definiscono l'insieme sono:

$$x = \lambda\,, \tag{9.32a}$$

$$y = \mu\,, \tag{9.32b}$$

$$z = 0\,, \tag{9.32c}$$

$$t = \lambda\,, \tag{9.32d}$$

con $\lambda, \mu \in \mathbb{R}$. Con due parametri liberi, abbiamo $\dim(V) = 2$. Per avere una base $\mathcal{B}_V$ di V, basta prendere $\lambda = 1$ e $\mu = 0$, e poi $\lambda = 0$ e $\mu = 1$: i vettori ottenuti sono indipendenti, e quindi generano V

$$\mathcal{B}_V = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \right\}.$$

Per il secondo punto, dobbiamo trovare le coordinate che soddisfano il sistema

$$\begin{cases} x - y & = 0 \\ z & = 0 \\ x - t & = 0 \end{cases}. \tag{9.33}$$

Le tre equazioni sono, banalmente, indipendenti; prendendo come parametro libero la variabile x, abbiamo le seguenti equazioni parametriche, dipendenti da un solo parametro:

$$x = \lambda\,, \tag{9.34a}$$

$$y = \lambda\,, \tag{9.34b}$$

$$z = 0\,, \tag{9.34c}$$

$$t = \lambda\,; \tag{9.34d}$$

si ha, quindi, $\dim(U \cap V) = 1$, e una sua base è costituita dal vettore

$$\begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix} .$$

La formula di Grassmann garantisce che

$$\dim(U) + \dim(V) = \dim(U + V) + \dim(U \cap V) ;$$

siccome $\dim(U \cap V) = 1$, $\dim(U + V) = 3 + 2 - 1 = 4$; ossia, $U + V$ è un sottospazio di $\mathbb{R}^4$ di dimensione 4; ma $\dim(\mathbb{R}^4) = 4$, quindi $U + V = \mathbb{R}^4$. Per convincersi di ciò, osserviamo che $U + V$ è generato da

$$\mathcal{B}_U \cup \mathcal{B}_V = \left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} , \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} , \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} , \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} , \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \right\} ;$$

gli ultimi 4 vettori sno linearmente indipendenti, e permettono di scrivere ogni vettore $\begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix}$ di $\mathbb{R}^4$ come loro combinazione lineare. Come si può verificare facilmente, si ha:

$$\begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = x \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} + y \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} + z \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} + (t - x) \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} .$$

La somma non è diretta, perché $\dim(U \cap V) = 1 \neq 0$.

Esercizio 4.3. (27 novembre 2008, prova in itinere)
Si considerino i seguenti sottospazi di $\mathbb{R}^4$:

$$U = \left\{ \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \in \mathbb{R}^4 \colon \quad x + y = t = 0 \right\}$$

$$V = \operatorname{Span} \left(\begin{pmatrix} 0 \\ 0 \\ -1 \\ 1 \end{pmatrix} , \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \end{pmatrix} \right) .$$

(1) Determinare una base per U:
(2) Determinare un sistema di generatori per $U + V$:
(3) Calcolare: $\dim(U + V) =$ $\qquad\qquad$ $\dim(U \cap V) =$

RISOLUZIONE. Scriviamo in forma parametrica le equazioni che definiscono il sottospazio U; per fare questo, osserviamo che la coordinata z è completamente libera. Scrivendo la prima equazione cartesiana nella forma $y = -x$ risulta naturale scegliere come parametro una qualunque delle due coordinate:

$$x = \lambda\,, \tag{9.35a}$$

$$y = -\lambda\,, \tag{9.35b}$$

$$z = \mu\,, \tag{9.35c}$$

$$t = 0\,, \tag{9.35d}$$

dove $\lambda, \mu \in \mathbb{R}$. La dimensione di U è dunque 2, come il numero di parametri liberi. Scegliendo $\lambda = 1, \mu = 0$ e $\lambda = 0, \mu = 1$ otteniamo due vettori linearmente indipendenti che possono essere usati come base. Perciò una base di U è:

$$\left\{ \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\}. \tag{9.36}$$

Per avere un sistema di generatori di $U + V$ basta unire i vettori della base di U con i generatori di V, che sono noti per la definizione di V come span di un insieme di vettori. Possiamo scrivere per questo sistema

$$\left\{ \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ -1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \end{pmatrix} \right\}. \tag{9.37}$$

Notando che il primo e l'ultimo vettore sono banalmente linearmente dipendenti (uno l'opposto dell'altro), ne terremo uno solo. Quindi il sistema

$$\left\{ \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ -1 \\ 1 \end{pmatrix}, \right\} \tag{9.38}$$

non solo genera $U + V$ come il precedente, ma è anche una sua base, visto che i tre vettori rimasti sono linearmente indipendenti, come è facile verificare. Allora abbiamo in modo immediato che

$$\dim(U + V) = 3\,;$$

la formula di Grassmann ci dà

$$\dim(U + V) = \dim(U) + \dim(V) - \dim(U \cap V)\,,$$

e, poiché $\dim(U) = \dim(V) = 2$, abbiamo

$$\dim(U \cap V) = 1\,.$$

5. Matrici, determinante e rango

Gli esercizi proposti in questa sezione si riferiscono ai contenuti del Capitolo 3.

Esercizio 5.1. (27 novembre 2008, prova in itinere)
Si consideri la seguente matrice reale quadrata di ordine 4, al variare del parametro reale h:

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

(1) Calcolare, al variare di h, $|A| =$:
(2) Determinare per quali valori di h la matrice A è singolare (non invertibile):
(3) Determinare per quali valori di h le righe di A sono vettori linearmente indipendenti:
(4) Determinare per quali valori di h la matrice ha rango 2:
(5) Posto $h = -1$, calcolare: $|A^5| =$ $\quad\quad$ $|-4A| =$

Risoluzione. La matrice è triangolare; il suo determinante è il prodotto degli elementi sulla diagonale; in funzione di h abbiamo

$$\det(A) = h^2(1-h). \tag{9.39}$$

La matrice è singolare se e solo se il suo determinante è nullo; dall'Eq. (9.39), uguagliando il determinante a 0, otteniamo che la condizione di singolarità è

$$h = 1 \lor h = 0.$$

I vettori riga di A sono linearmente indipendenti se il rango è massimo, ossia se il determinante non è nullo; quindi, nella situazione complementare rispetto al punto precedente:

$$h \neq 1 \land h \neq 0.$$

Per avere rango 2, la matrice non deve avere rango massimo, ossia, $\det(A) = 0$. Gli unici valori da esplorare sono, quindi $h = 1 \lor h = 0$.

Per $h = 0$, la matrice diventa

$$A = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \end{pmatrix}; \tag{9.40}$$

nella matrice (9.40), osserviamo che la terza e la quarta riga coincidono, la prima è nulla, e la seconda è sicuramente indipendente dalla terza. Possiamo, dunque, dire che il rango in questo caso è 2.

Invece, per $h = 1$, la matrice diventa

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \end{pmatrix}; \tag{9.41}$$

nella matrice (9.41), stavolta la seconda e la terza riga coincidono, ma la prima, seconda e quarta righe sono linearmente indipendenti. Il rango in questo caso è 3. Pertanto, il rango è 2 per

$$h = 0.$$

Posto $h = -1$, dall'Eq. (9.39) abbiamo

$$\det(A) = 2\,; \tag{9.42}$$

quindi, usando le regole per il calcolo dei determinanti:

$$|A^5| = 2^5 = 32\,,$$

e

$$|-4A| = (-4)^4 \times 2 = 512\,.$$

Esercizio 5.2. Siano A e B due matrici reali quadrate di ordine $n \geq 2$. Dimostrare che

$$\mathrm{rg}(A+B) \leq \mathrm{rg}(A) + \mathrm{rg}(B).$$

Risoluzione. Consideriamo gli operatori lineari $L_A\colon \mathbb{R}^n \to \mathbb{R}^n$ e $L_B\colon \mathbb{R}^n \to \mathbb{R}^n$ definiti rispettivamente da

$$L_A(X) = AX, \quad L_B(X) = BX \quad \forall X \in \mathbb{R}^n.$$

Si ha $\mathrm{rg}(A) = \dim \mathrm{Im}\, L_A$ e $r(B) = \dim \mathrm{Im}\, L_B$. Consideriamo inoltre l'operatore somma degli operatori L_A e L_B, definito nel seguente modo:

$$(L_A + L_B)(X) = L_A(X) + L_B(X) = AX + BX = (A+B)X, \quad \forall X \in \mathbb{R}^n.$$

Osserviamo che la matrice associata a $(L_A + L_B)$, nella base standard di $\mathbb{R}^n$, è la matrice $A+B$. Si ha quindi:

$$\mathrm{rg}(A+B) = \dim \mathrm{Im}(L_A + L_B).$$

Osserviamo che si ha la seguente inclusione fra sottospazi:

$$\mathrm{Im}(L_A + L_B) \subset \mathrm{Im}\, L_A + \mathrm{Im}\, L_B\,.$$

Infatti, siano $X \in \mathbb{R}^n$ e $Y = (L_A + L_b)(X)$ si ha:

$$Y = (A+B)X = AX + BX \in \mathrm{Im}\, L_A + \mathrm{Im}\, L_B.$$

Abbiamo quindi

$$\mathrm{rg}(A+B) = \dim \mathrm{Im}(L_A + L_B) \leq \dim(ImL_A + \mathrm{Im}\, L_B),$$

dal teorema di Grassmann si ha

$$\dim(ImL_A + \mathrm{Im}\, L_B) \leq \dim \mathrm{Im}\, L_A + \dim \mathrm{Im}\, L_B = \mathrm{rg}(A) + \mathrm{rg}(B),$$

possiamo pertanto concludere che

$$r(A+B) \leq \mathrm{rg}(A) + \mathrm{rg}(B).$$

---

6. Applicazioni lineari

Gli esercizi proposti in questa sezione si riferiscono ai contenuti del Capitolo 4.

Esercizio 6.1. (4 febbraio 2009, prova in itinere)
Si consideri l'applicazione lineare $f\colon \mathbb{R}^3 \to \mathbb{R}^4$ definita da

$$f(x, y, z) = (x + y + 2z, x + z, y + z, x + y + 2z).$$

Determinare:

(1) La matrice A che rappresenta f rispetto alle basi canoniche:
(2) $\dim \operatorname{Im} f =$ $\qquad\qquad$ $\dim \operatorname{Ker} f =$
(3) Le eq. cartesiane per $\operatorname{Ker} f$ e $\operatorname{Im} f$:
(4) Una matrice $B \in M_{\mathbb{R}}(3)$, non nulla, per cui si abbia $AB = 0$:

Esercizio 6.2. (4 febbraio 2009, appello)
Si consideri la seguente applicazione lineare $f\colon \mathbb{R}^2 \to \mathbb{R}^3$:

$$f(x, y) = (x + y, -x, -y).$$

(1) Determinare la matrice A che rappresenta f rispetto alle basi standard:

(2) Determinare: $\dim \operatorname{Im} f =$ $\qquad\qquad$ $\dim \operatorname{Ker} f =$
(3) f è iniettiva? Giustificare la risposta.

(4) Determinare le equazioni cartesiane per $\operatorname{Im} f$:
(5) Sia $\boldsymbol{w} = (1, a, 0)$. Determinare per quali valori di $a \in \mathbb{R}$ si ha $\boldsymbol{w} \in \operatorname{Im} f$:

Esercizio 6.3. (19 febbraio 2009, appello)
Si consideri la seguente applicazione lineare $f\colon \mathbb{R}^3 \to \mathbb{R}^4$:

$$f(x, y, z) = (x, x - y + z, z - 2x - y, x + z).$$

(1) Determinare la matrice A che rappresenta f rispetto alle basi standard:
(2) Determinare: $\dim \operatorname{Im} f =$ $\qquad\qquad$ $\dim \operatorname{Ker} f =$
(3) Determinare una base di $\operatorname{Im} f$
(4) Completare la base di $\operatorname{Im} f$ per avere una base di $\mathbb{R}^4$

(5) (*non assegnato*) Determinare la matrice B che rappresenta f rispetto alla base standard in $\mathbb{R}^3$ ed alla seguente base di $\mathbb{R}^4$:

$$\mathcal{B}_4 = \left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 2 \end{pmatrix} \right\}.$$

(6) Determinare la matrice M tale che $AM = B$

Esercizio 6.4. (6 luglio 2009, appello)
Si consideri la seguente applicazione lineare $f\colon \mathbb{R}^3 \to \mathbb{R}^2$:

$$f\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2x - 2y - z \\ y + z \end{pmatrix}.$$

(1) Determinare la matrice A che rappresenta f rispetto alle basi standard:

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

(2) Determinare: $\dim \operatorname{Im} f = 2$ $\qquad \dim \operatorname{Ker} f = 1$

(3) Il vettore $\boldsymbol{v} = \begin{pmatrix} 3 \\ 3 \end{pmatrix}$ appartiene a $\operatorname{Im} f$? Giustificare la risposta.
Sì: f è surgettiva ($\dim \operatorname{Im} f = \operatorname{rg} A = 2$).

(4) f è iniettiva? Giustificare la risposta.
No: $\dim \operatorname{Ker} f = 1$.

(5) Sia $U \subset \mathbb{R}^3$ il sottospazio di equazione $2x + z = 0$.
Calcolare $\dim f(U) = 1$.

Esercizio 6.5. (16 settembre 2009, appello)
Fissata la base standard $\mathcal{B} = \{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ di $\mathbb{R}^3$, si consideri l'applicazione lineare $\boldsymbol{f}\colon \mathbb{R}^3 \to \mathbb{R}^3$ tale che:

$$\boldsymbol{f}(\boldsymbol{e}_1) = \boldsymbol{e}_2 + \boldsymbol{e}_3 \quad \boldsymbol{f}(\boldsymbol{e}_2) = \boldsymbol{e}_1 + \boldsymbol{e}_2 + \boldsymbol{e}_3 \quad \boldsymbol{f}(\boldsymbol{e}_3) = 2\boldsymbol{e}_1 + \boldsymbol{e}_2 + \boldsymbol{e}_3.$$

(1) Determinare la matrice A che rappresenta $\boldsymbol{f}$ rispetto alla base standard:

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

(2) Calcolare: $\dim \operatorname{Im} \boldsymbol{f} = 2$, $\dim \operatorname{Ker} \boldsymbol{f} = 1$

(3) Determinare le equazioni di $\operatorname{Im} \boldsymbol{f}$ e $\operatorname{Ker} \boldsymbol{f}$:

$$\operatorname{Im} \boldsymbol{f} = \left\{(x, y, z) \in \mathbb{R}^3 \mid y - z = 0\right\},$$
$$\operatorname{Ker} \boldsymbol{f} = \left\{(x, y, z) \in \mathbb{R}^3 \mid y + 2z = 0,\ x + y + z = 0\right\}$$

ossia, $\operatorname{Ker} \boldsymbol{f}$ è la retta per l'origine diretta come $\boldsymbol{d} = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$.

(4) Determinare un piano per l'origine $U \subset \mathbb{R}^3$ la cui immagine sia una retta per l'origine di $\mathbb{R}^3$: basta prendere un piano che contenga $\operatorname{Ker} \boldsymbol{f}$, ossia che sia generato da $\boldsymbol{d}$ e da un altro vettore indipendente, per esempio, $\boldsymbol{u} = (0, 1, 0)$. Con questa scelta il piano ha equazione $x = z$.

Esercizio 6.6. (23 novembre 2009, appello straord.)
Fissata le basi standard $\mathcal{B}_3 = \{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ di $\mathbb{R}^3$, e $\mathcal{B}' = \{\boldsymbol{e}'_1, \boldsymbol{e}'_2, \boldsymbol{e}'_3 \boldsymbol{e}'_4\}$ di $\mathbb{R}^4$ si consideri l' applicazione lineare $L : \mathbb{R}^4 \to \mathbb{R}^3$ tale che:

$L(\boldsymbol{e}'_1) = \boldsymbol{e}_1 - \boldsymbol{e}_2 - \boldsymbol{e}_3 \quad L(\boldsymbol{e}'_2) = \boldsymbol{e}_2 - \boldsymbol{e}_3 \quad L(\boldsymbol{e}'_3) = -\boldsymbol{e}_1 + 2\boldsymbol{e}_3 \quad L(\boldsymbol{e}'_4) = 2\boldsymbol{e}_1 - \boldsymbol{e}_2 - 3\boldsymbol{e}_3.$

(1) Determinare la matrice A che rappresenta L rispetto alle basi standard:

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

(2) Calcolare: $\dim(\operatorname{Im} L) = 2$, $\qquad \dim(\operatorname{Ker} L) = 2$

(3) Precisare se L è iniettiva e/o surgettiva, motivando la risposta.
Non è iniettiva ($\dim(\operatorname{Ker} L) \neq 0$);
non è surgettiva ($\dim(\operatorname{Im} L) = 2 \neq \dim \mathbb{R}^3$).

(4) Determinare l'equazione cartesiana di $\operatorname{Im} L$:

$$2x + y + z = 0$$

(5) Determinare una base di $\operatorname{Ker} L$:

$$\mathcal{B}_{Ker} = \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -2 \\ -1 \\ 0 \\ 1 \end{pmatrix} \right\}$$

Esercizio 6.7. (11 febbraio 2010)
Si consideri l'applicazione lineare $L\colon \mathbb{R}^2 \to \mathbb{R}^3$ definita da

$$L \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x + hy \\ hx + y \\ (h+1)y \end{pmatrix},$$

con $h \in \mathbb{R}$ e siano $\{\boldsymbol{e}_1, \boldsymbol{e}_2\}$ i vettori della base standard di $\mathbb{R}^2$.

(1) Calcolare: $L(\boldsymbol{e}_1) = \begin{pmatrix} 1 \\ h \\ 0 \end{pmatrix} \qquad L(\boldsymbol{e}_2) = \begin{pmatrix} h \\ 1 \\ 1+h \end{pmatrix}$

(2) Scrivere la matrice A_L rappresentativa di L nelle basi standard di $\mathbb{R}^2$ e $\mathbb{R}^3$:

$$A_L = \begin{pmatrix} 1 & h \\ h & 1 \\ 0 & 1+h \end{pmatrix}$$

(3) Dire per quale/i valori di h l'applicazione è iniettiva: $h \neq -1$

(4) Posto $h = 1$, determinare la/le equazioni cartesiane di $\operatorname{Im} L$: $x - y = 0$

(5) Posto $h = -1$, determinare una base di $\operatorname{Ker} L$: $\mathcal{B}_{\mathrm{Ker}} = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}$

Esercizio 6.8. (14 settembre 2010)
Fissata la base standard $\mathcal{B} = \{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ di $\mathbb{R}^3$, si consideri l' applicazione lineare $L : \mathbb{R}^3 \to \mathbb{R}^3$ tale che:

$$L(\boldsymbol{e}_1) = \boldsymbol{e}_1 + \boldsymbol{e}_2; \quad L(\boldsymbol{e}_2) = \boldsymbol{e}_2 - \boldsymbol{e}_3; \quad L(\boldsymbol{e}_3) = \boldsymbol{e}_1 + \boldsymbol{e}_3.$$

(1) Calcolare: $\dim(\operatorname{Im} L) = 2$, $\dim(\operatorname{Ker} L) = 1$
(2) Determinare l'equazione cartesiana di $\operatorname{Im} L$: $x - y - z = 0$

(3) Determinare una base di $\operatorname{Ker} L$:

$$\mathcal{B} = \left\{ \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix} \right\}$$

(4) Sia $\mathcal{B}'$ la base composta dai seguenti vettori:

$$\boldsymbol{u}_1 = \boldsymbol{e}_1 + \boldsymbol{e}_2, \qquad \boldsymbol{u}_2 = \boldsymbol{e}_2 + \boldsymbol{e}_3, \qquad \boldsymbol{u}_3 = \boldsymbol{e}_1 + \boldsymbol{e}_3.$$

Determinare la matrice B rappresentativa di L nella base $\mathcal{B}'$

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

Esercizio 6.9. Stabilire quali tra le seguenti applicazioni tra spazi vettoriali sono lineari:
(1) $f\colon \mathbb{R}^2 \to \mathbb{R} \quad f(x, y) = x.y$;
(2) $f\colon \mathbb{R}^2 \to \mathbb{R}^2 \quad f(x, y) = (x, x - y)$;
(3) $f\colon \mathbb{R} \to \mathbb{R}^2 \quad f(x) = (ax, x + b), \quad a, b \in \mathbb{R}$.

(1) Ricordiamo che $f\colon \mathbb{R}^2 \to \mathbb{R}$ è un'applicazione lineare se e solo se sono verificate le seguenti condizioni:

$$f(\boldsymbol{v}_1 + \boldsymbol{v}_2) = f(\boldsymbol{v}_1) + f(\boldsymbol{v}_2) \quad \forall \boldsymbol{v}_1, \boldsymbol{v}_2 \in \mathbb{R}^2,$$
$$f(\lambda \boldsymbol{v}) = \lambda f(\boldsymbol{v}) \quad \forall \boldsymbol{v} \in \mathbb{R}^2, \forall \lambda \in \mathbb{R}.$$

Sia $\boldsymbol{v} = (x, y) \in \mathbb{R}^2$, osserviamo che risulta

$$f(\lambda \boldsymbol{v}) = f(\lambda x, \lambda y) = \lambda x.\lambda y = \lambda^2 x.y,$$
$$\lambda f(\boldsymbol{v}) = \lambda x.y;$$

poiché risulta $f(\lambda \boldsymbol{v}) \neq \lambda f(\boldsymbol{v})$, possiamo concludere che f non è lineare.
(2) Sia $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ definita da $f(x, y) = \begin{pmatrix} x \\ x - y \end{pmatrix}$. Osserviamo che l'applicazione f può essere scritta in forma matriciale

$$f(x, y) = A \cdot \begin{pmatrix} x \\ y \end{pmatrix}, \quad A = \begin{pmatrix} 1 & 0 \\ 1 & -1 \end{pmatrix}.$$

La linearità di f segue dalle proprietà di linearità del prodotto matriciale, infatti si ha:

$$A \cdot (X + X') = A \cdot X + A \cdot X', \quad \forall X, X' \in M_{\mathbb{R}}(2, 1),$$

$$A \cdot (\lambda X) = \lambda(A \cdot X), \quad \forall \lambda \in \mathbb{R}, \ \ \forall X \in M_{\mathbb{R}}(2,1).$$

(3) Sia $f\colon \mathbb{R} \to \mathbb{R}^2$ definita da $f(x) = \begin{pmatrix} ax \\ x+b \end{pmatrix}$, $a, b \in \mathbb{R}$. Condizione necessaria affinché f sia lineare è che risulti $f(O_{\mathbb{R}}) = O_{\mathbb{R}^2}$, quindi

$$f(0) = (0, b) = (0, 0) \Rightarrow b = 0.$$

Verifichiamo ora che $f(x) = (ax, x)$ è lineare $\forall a \in \mathbb{R}$. A tal fine possiamo scrivere f in forma matriciale

$$f(x) = A \cdot x, \quad A = \begin{pmatrix} a \\ 1 \end{pmatrix},$$

da cui possiamo concludere che f è lineare (cf. p.to (b)).

Esercizio 6.10. Si consideri l'endomorfismo $f\colon \mathbb{R}^3 \to \mathbb{R}^3$ definito da:

$$f \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2x - y + z \\ y + z \\ x + z \end{pmatrix}.$$

(1) Scrivere la matrice associata a f rispetto alla base standard $\{e_1, e_2, e_3\}$ di $\mathbb{R}^3$;
(2) calcolare la dimensione degli spazi $\operatorname{Ker} f$ e $\operatorname{Im} f$;
(3) scrivere le equazioni dei sottospazi $\operatorname{Ker} f$ e $\operatorname{Im} f$;
(4) determinare una base per i sottospazi $\operatorname{Ker} f$ e $\operatorname{Im} f$.

(1) Fissata la base standard $\mathcal{B} = \{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ di $\mathbb{R}^3$, calcoliamo le immagini dei vettori della base

$$f(\boldsymbol{e}_1) = f \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} \quad f(\boldsymbol{e}_2) = f \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \quad f(\boldsymbol{e}_3) = f \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

La matrice $A \in M_{\mathbb{R}}(3)$ associata ad f nella base $\mathcal{B}$ ha come colonne, rispettivamente, le coordinate dei vettori $f(\boldsymbol{e}_1), f(\boldsymbol{e}_2), f(\boldsymbol{e}_3)$ nella base $\mathcal{B}$:

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

(2) Ricordiamo come sono definiti i sottospazi $\operatorname{Ker} f$ e $\operatorname{Im} f$:

$$\operatorname{Ker} f = \left\{ \boldsymbol{v} \in \mathbb{R}^3 | \ \ f(\boldsymbol{v}) = 0_{\mathbb{R}^3} \right\},$$

$$\operatorname{Im} f = \left\{ \boldsymbol{w} \in \mathbb{R}^3 | \ \ \exists \boldsymbol{v} \ \in \mathbb{R}^3\colon \ f(\boldsymbol{v}) = \boldsymbol{w} \right\}.$$

Osserviamo che i vettori $f(\boldsymbol{e}_1)$, $f(\boldsymbol{e}_2), f(\boldsymbol{e}_3)$ sono un sistema di generatori del sottospazio immagine $\operatorname{Im} f$ e che risulta

$$\dim(\operatorname{Im} f) = \operatorname{rg} A.$$

Poiché $|A| = -2 + 2 = 0$, e le prime due colonne di A sono linearmente indipendenti si ha $r(A) = 2$. Possiamo concludere che $\dim \operatorname{Im} f = 2$ e quindi $\operatorname{Im} f$

è un piano in $\mathbb{R}^3$ passante per l'origine, generato dai vettori $f(\boldsymbol{e}_1)$ e $f(\boldsymbol{e}_2)$. Dal teorema delle dimensoni si ha che:

$$\dim(\mathbb{R}^3) = \dim(\operatorname{Ker} f) + \dim(\operatorname{Im} f),$$

da cui ricaviamo che $\dim \operatorname{Ker} f = 1$ e quindi $\operatorname{Ker} f$ è una retta in $\mathbb{R}^3$ passante per l'origine.

(3) Oserviamo che risulta:

$$\operatorname{Im} f = \operatorname{Span}(f(\boldsymbol{e}_1), f(\boldsymbol{e}_2)) = \left\{ \boldsymbol{w} \in \mathbb{R}^3 | \ \ \boldsymbol{w} = \alpha f(\boldsymbol{e}_1) + \beta f(\boldsymbol{e}_2), \ \alpha, \beta \in \mathbb{R} \right\},$$

le equazioni parametriche di $\operatorname{Im} f$ sono le seguenti:

$$\operatorname{Im} f = \left\{ (x, y, z) \in \mathbb{R}^3 \colon x = 2\alpha - \beta, \quad y = \beta, \quad z = \alpha, \quad \alpha, \beta \in \mathbb{R} \right\},$$

eliminando i parametri otteniamo l'equazione cartesiana di $\operatorname{Im} f$:

$$x + y - 2z = 0.$$

Il sottospazio $\operatorname{Ker} f$ è lo spazio delle soluzioni del seguente sistema lineare omogeneo

$$\operatorname{Ker} f = \left\{ (x, y, z) \in \mathbb{R}^3 | \ \ A \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\},$$

poiché la matrice A ha rango 2 e la seconda e la terza riga di A sono linearmente indipendenti, le equazioni di $\operatorname{Ker} f$ sono le seguenti:

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

(4) Osserviamo che il sottospazio $\operatorname{Ker} f$ può essere descritto con equazioni parametriche

$$\operatorname{Ker} f = \left\{ (x, y, z) \in \mathbb{R}^3 | \ \ x = t, \ y = t, \ z = -t, \quad t \in \mathbb{R} \right\},$$

quindi $\operatorname{Ker} f = \operatorname{Span}((1, 1, -1))$, per cui possiamo scegliere come base di $\operatorname{Ker} f$ il vettore $\boldsymbol{v} = (1, 1, -1)$.

Infine, per trovare una base del sottospazio $\operatorname{Im} f$ basta osservare che $\operatorname{Im} f = \operatorname{Span}(f(\boldsymbol{e}_1), f(\boldsymbol{e}_2))$, e i vettori $f(\boldsymbol{e}_1)$ e $f(\boldsymbol{e}_2)$ sono linearmente indipendenti, una base di $\operatorname{Im} f$ è la seguente : $\boldsymbol{w}_1 = (2, 0, 1)$ e $\boldsymbol{w}_2 = (-1, 1, 0)$.

Esercizio 6.11. Si consideri l'applicazione $L \colon \mathcal{M}_{\mathbb{R}}(2) \to \mathcal{M}_{\mathbb{R}}(2)$ definita da

$$L \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & \frac{b+c}{2} \\ \frac{b+c}{2} & d \end{pmatrix}.$$

(1) Verificare che L è lineare.
(2) Scrivere la matrice associata a L nella base standard di $\mathcal{M}_{\mathbb{R}}(2)$.
(3) Determinare una base per i sottospazi $\operatorname{Ker} f$ e $\operatorname{Im} f$.
(4) Scrivere l'applicazione L^2.

(1) Ricordiamo che L è un'applicazione lineare se sono verificate le seguenti condizioni:

$$L(A+B) = L(A) + L(B) \quad \forall A, B \in \mathcal{M}_{\mathbb{R}}(2)$$

$$L(\lambda A) = \lambda L(A) \quad \forall A \in \mathcal{M}_{\mathbb{R}}(2), \quad \forall \lambda \in \mathbb{R}.$$

Siano $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ e $B = \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix}$, si ha:

$$L(A+B) = L\begin{pmatrix} a+a' & b+b' \\ c+c' & d+d' \end{pmatrix} = \begin{pmatrix} a+a' & \frac{(b+b'+c+c')}{2} \\ \frac{(b+b'+c+c')}{2} & d+d' \end{pmatrix} = L(A) + L(B).$$

Siano $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ e $\lambda \in \mathbb{R}$, si ha:

$$L(\lambda A) = L\begin{pmatrix} \lambda a & \lambda b \\ \lambda c & \lambda d \end{pmatrix} = \begin{pmatrix} \lambda a & \lambda\frac{b+c}{2} \\ \lambda\frac{b+c}{2} & \lambda d \end{pmatrix} = \lambda L(A).$$

(2) Ricordiamo che la base standard dello spazio vettoriale $\mathcal{M}_{\mathbb{R}}(2)$ è costituita dalle matrici $\{E_{11}, E_{12}, E_{21}, E_{22}\}$, dove

$$E_{11} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad E_{12} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad E_{21} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \quad E_{22} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Sia $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{M}_{\mathbb{R}}(2)$, A si scrive come combinazione lineare delle matrici della base standard

$$A = aE_{11} + bE_{12} + cE_{21} + dE_{22},$$

il vettore $X \in \mathbb{R}^4$ delle coordinate di A rispetto alla base fissata è il seguente:

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

La matrice associata ad l nella base fissata è la matrice $M \in \mathcal{M}_{\mathbb{R}}(4)$ le cui colonne sono rispettivamente le coordinate di $f(E_{11})$, $f(E_{12})$, $f(E_{21})$, $f(E_{22})$ rispetto alla base fissata. Si ha:

$$L(E_{11}) = L\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = 1E_{11},$$

$$L(E_{12}) = L\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{pmatrix} = \frac{1}{2}E_{12} + \frac{1}{2}E_{21},$$

$$L(E_{21}) = L\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{pmatrix} = \frac{1}{2}E_{12} + \frac{1}{2}E_{21},$$

$$L(E_{22}) = L\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = 1E_{22},$$

concludendo la matrice M è la seguente:

$$M = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

(3) Il sottospazio $\operatorname{Im} L$ è per definizione:

$$\operatorname{Im} L = \{B \in \mathcal{M}_{\mathbb{R}}(2) | \ \ \exists A \in \mathcal{M}_{\mathbb{R}}(2) : \ L(A) = B\},$$

poiché si ha $\operatorname{rg}(M) = 3$, $\dim(\operatorname{Im} L) = 3$ e una base per $\operatorname{Im} L$ è data dalle matrici

$$L(E_{11}) = E_{11}, \quad L(E_{12}) = \frac{1}{2}(E_{12} + E_{21}), \quad L(E_{22}) = E_{22}.$$

Posiamo concludere che lo spazio immagine è il seguente sottospazio di $\mathcal{M}_{\mathbb{R}}(2)$:

$$\operatorname{Im} L = \left\{ B = \begin{pmatrix} a & b \\ c & d \end{pmatrix} : \quad a, b, c, d \in \mathbb{R}, \quad b = c \right\},$$

cioè il sottospazio delle matrici quadrate di ordine 2 reali e simmetriche.

Il sottospazio $\operatorname{Ker} L$ è per definizione:

$$\operatorname{Ker} L = \left\{ A \in \mathcal{M}_{\mathbb{R}}(2) | \ \ L(A) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \right\}.$$

Per il teorema delle dimensioni (4.8) si ha:

$$\dim \mathcal{M}_{\mathbb{R}}(2) = \dim(\operatorname{Ker} L) + \dim(\operatorname{Im} L),$$

da cui si ricava che $\dim(\operatorname{Ker} L) = 1$. Osserviamo che $L \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ se e solo se si ha

$$a = d = 0, \quad b + c = 0.$$

Otteniamo quindi che

$$\operatorname{Ker} L = \operatorname{Span}(\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}),$$

per cui come base per il sottospazio possiamo scegliere la matrice $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$.

(4) L'applicazione L^2 è la composizione di L con se stessa nel seguente modo: $L^2(A) = L(L(A))$. Sia $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{M}_{\mathbb{R}}(2)$, si ha

$$L^2(A) = L \begin{pmatrix} a & \frac{b+c}{2} \\ \frac{b+c}{2} & d \end{pmatrix} = \begin{pmatrix} a & \frac{b+c}{2} \\ \frac{b+c}{2} & d \end{pmatrix} = L(A) \quad \forall A \in \mathcal{M}_{\mathbb{R}}(2).$$

Possiamo concludere che $L^2 = L$; analogamente vale la stessa relazione per le matrici associate:

$$M^2 = M.$$

---

7. Sistemi di equazioni lineari

Gli esercizi proposti in questa sezione si riferiscono ai contenuti del Capitolo 5.

Esercizio 7.1. (4 febbraio 2009, prova in itinere)
Si consideri il seguente sistema lineare dipendente dal parametro reale h:

$$\begin{cases} (h+1)x + (h-1)y + z = 0 \\ (2h-1)x + y + (h-1)z = h \end{cases}$$

(1) Determinare per quali valori di h il sistema ammette soluzioni:
(2) Determinare per quali valori di h l'insieme delle soluzioni è una retta in $\mathbb{R}^3$:
(3) Risolvere il sistema per $h = -1$.

Risoluzione. (1) Osserviamo che il sistema dato è un sistema di 2 equazioni e 3 incognite. La matrice dei coefficienti e la colonna dei termini noti sono le seguenti:

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

Prima di tutto calcoliamo il rango di A al variare di $h \in \mathbb{R}$; $\forall h \in \mathbb{R}$ risulta $1 \leq \mathrm{rg}(A) \leq 2$. Scegliamo un minore di ordine 2 in A, ad esempio quello determinato dalle colonne A^2 e A^3:

$$\delta = \begin{vmatrix} h-1 & 1 \\ 1 & h-1 \end{vmatrix} = (h-1)^2 - 1 = h^2 - 2h = h(h-2).$$

Osserviamo che risulta:

$$\delta = 0 \iff h = 0 \ \vee h = 2.$$

Possiamo quindi affermare che:
$\forall h \in \mathbb{R}$, con $h \neq 0$ e $h \neq 2$, $\mathrm{rg}(A) = 2$. Consideriamo ora la matrice completa del sistema $\tilde{A} = (A|B)$; questa è una matrice di ordine 2×4, pertanto si ha $\mathrm{rg}(\tilde{A}) \leq 2$, per ogni $h \in \mathbb{R}$. Possiamo concludere che, $\forall h \in \mathbb{R}$, con $h \neq 0$ e $h \neq 2$:

$$\mathrm{rg}(\tilde{A}) = \mathrm{rg}(A) = 2.$$

Per il Teorema di Rouché-Capelli, il sistema ammette soluzioni e l'insieme delle soluzioni $\mathcal{V}_h$ è una varietà lineare di dimensione $3 - 2 = 1$ in $\mathbb{R}^3$, quindi è una retta in $\mathbb{R}^3$, che non passa per l'origine.

Esaminiamo ora, caso per caso, i valori del parametro rimasti.
Sia $h = 0$: sostituiamo tale valore nel sistema, abbiamo

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

Osserviamo che il sistema è omogeneo, quindi ammette sempre soluzioni. Poiché risulta $A_2 = -A_1$, possiamo concludere che $\mathrm{rg}(A) = 1$. L'insieme delle soluzioni è quindi un sottospazio V_0 di $\mathbb{R}^3$ di dimensione $3 - 1 = 2$, è quindi un piano di $\mathbb{R}^3$ passante per l'origine.

Sia ora $h = 2$: sostituiamo tale valore nel sistema, abbiamo

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

Si osserva che le righe della matrice A sono uguali: $\mathrm{rg}(A) = 1$. Indichiamo con $\tilde{A} = (A, B)$; si vede subito che ogni sottomatrice che contiene la colonna B ha determinante non nullo: risulta $\mathrm{rg}(\tilde{A}) = 2$. Quindi abbiamo:

$$\mathrm{rg}(A) \neq \mathrm{rg}(\tilde{A}).$$

Pertanto, per il teorema di Rouché-Capelli, il sistema non ammette soluzioni per $h = 2$. Riassumendo il sistema ammette soluzioni per $h \neq 2$..

(2) La discussione riportata al punto precedente permette di affermare che il sistema ha per soluzione una retta di $\mathbb{R}^3$ per $h \neq 2$ e $h \neq 0$.

(3) Con $h = -1$ il sistema diventa

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

Scambiando le due righe, possiamo individuare in z la variabile libera; dall'equazione $-2y + z = 0$ ricaviamo $y = -z/2$. Sostituendo nell'altra, otteniamo

$$-3x - z/2 - 2z = -3x - (5/2)z = -1$$

da cui $x = -(5/6)z + 1/3$.

La soluzione in forma parametrica ha la forma

$$\begin{cases} x = -\frac{5}{6}\lambda + \frac{1}{3} \\ y = -\frac{\lambda}{2} \\ z = \lambda \end{cases} \qquad \lambda \in \mathbb{R}$$

Esercizio 7.2. (4 febbraio 2009, appello)

Si consideri il seguente sistema lineare dipendente dal parametro reale h:

$$\begin{cases} hx + y = 1 \\ x + hy = h \\ (1-h)x + y + hz = 0 \\ 2x + (2+h)y + hz = 1 + h \end{cases}$$

(1) Determinare per quali valori di h il sistema ammette soluzioni:
(2) Determinare per quali valori di h l'insieme delle soluzioni è una retta in $\mathbb{R}^3$:
(3) Determinare per quali valori di h il sistema ha un'unica soluzione:

Esercizio 7.3. (19 febbraio 2009, appello)

Si consideri il seguente sistema lineare dipendente dal parametro reale k:

$$\begin{cases} x + (k+1)y + z = 0 \\ -4x + y + kz = 0 \\ (k+4)x - y = k + 1 \end{cases}$$

(1) Determinare per quali valori di k il sistema ammette un'unica soluzione:
(2) Determinare per quali valori di k l'insieme delle soluzioni è una retta in $\mathbb{R}^3$:
(3) Risolvere il sistema per $k = 1$:

Esercizio 7.4. (6 luglio 2009, appello)
Si consideri la matrice dipendente dal parametro reale h

$$A = \begin{pmatrix} h & 1 & 0 \\ 0 & 1 & h \\ h & h+1 & 1 \\ h & h & 0 \end{pmatrix}$$

(1) Si determini il rango di A al variare di h:
    Per $h = 0, 1$ $\operatorname{rg} A = 2$; per $h \neq 0, 1$ $\operatorname{rg} A = 3$.
(2) Si consideri il sistema dipendente da h

$$\begin{cases} hx + y = 0 \\ y + hz = 2 \\ hx + (h+1)y + z = 1 \\ hx + hy = 0 \end{cases}$$

   (a) Si determini per quali valori di h il sistema ammette soluzioni: $h = 2$
   (b) Si determini per quali valori di h la soluzione esiste ed è unica: $h = 2$

Esercizio 7.5. (16 settembre 2009, appello)
Si consideri il seguente sistema lineare dipendente dal parametro reale k:

$$\begin{cases} kx + y + (k+2)z = 2 \\ x + ky + (2k+1)z = k+1 \end{cases}$$

(1) Determinare per quali valori di k **non** ammette soluzione: $k = -1$
(2) Determinare per quali valori di k l'insieme delle soluzioni è una retta in $\mathbb{R}^3$: $k \neq 1, -1$
(3) Determinare per quali valori di k l'insieme delle soluzioni è un piano in $\mathbb{R}^3$: $k = +1$

Esercizio 7.6. (23 settembre 2009, appello straord.)
Si consideri il seguente sistema lineare dipendente dal parametro reale k:

$$\begin{cases} kx + z = 2 \\ x - 2y - kz = -1 \\ x + kz = -2 \end{cases}$$

(1) Determinare per quali valori di k il sistema ammette **un'unica** soluzione:
$k \neq \pm 1$
(2) Determinare per quali valori di k il sistema **non ammette** soluzioni:
$k = +1$
(3) Determinare per quali valori di k il sistema ammette **infinite** soluzioni:
$k = -1$
(4) Posto $k = 2$, risolvere il sistema.

$$x = 2, \quad y = \frac{7}{2}, \quad z = -2\,.$$

ESERCIZIO 7.7. (14 settembre 2010)
Si consideri il seguente sistema lineare dipendente dal parametro reale h:

$$\begin{cases} x + hy + z + (h^2 - 1)t = h + 1 \\ y + z + t = 0 \\ x + hy + z = 0 \end{cases}$$

(1) Determinare per quali valori di h il sistema ammette soluzioni: $h \neq 1$
(2) Determinare per quali valori di h l'insieme delle soluzioni ha dimensione 2: $h = -1$
(3) Posto $h = 2$, risolvere il sistema. $\begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} 2 \\ -1 \\ 0 \\ 1 \end{pmatrix} + z \begin{pmatrix} 1 \\ -1 \\ 1 \\ 0 \end{pmatrix}$, z libero

---

ESERCIZIO 7.8. (14 settembre 2010)
Si consideri il seguente sistema lineare dipendente dal parametro reale k:

$$\begin{cases} x + y - kz = k \\ x + y + z = 2 + 3k \\ 2x - ky + z = 2 \end{cases}$$

(1) Si determini per quali valori di k il sistema ammette soluzioni: $k \neq -2\,.$
(2) Si determini per quale valore di k la soluzione non è unica: $k = -1\,.$
(3) In corrispondenza del valore di k del punto precedente, si determini la soluzione generale del sistema:

$$\begin{cases} x = 3 \\ y = -4 - t \\ z = t \end{cases} \qquad t \in \mathbb{R}$$

ESERCIZIO 7.9. Dato il sistema lineare $\begin{cases} hx + y = 0 \\ x + hy - z = 1 \\ -y + hz = 0 \end{cases}$, stabilire per quali valori del parametro reale h

(1) il sistema ha soluzioni reali;
(2) il sistema ammette un'unica soluzione;
(3) l'insieme delle soluzioni è una retta in $\mathbb{R}^3$;
(4) l'insieme delle soluzioni è un piano in $\mathbb{R}^3$.

RISOLUZIONE. (1) Il sistema lineare è non omogeneo, ha 3 equazioni in 3 incognite. In forma matriciale il sistema si scrive $AX = B$, dove la matrice A dei coefficienti, la colonna B dei termini noti e il vettore X delle incognite sono i seguenti:

$$A = \begin{pmatrix} h & 1 & 0 \\ 1 & h & -1 \\ 0 & -1 & h \end{pmatrix} \quad B = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

Indicata con $\tilde{A}$ la matrice completa, ottenuta aggiungendo ad A la colonna B, per il teorema di Rouché Capelli il sistema ammette soluzioni se e solo si ha:

$$\mathrm{rg}(A) = \mathrm{rg}(\tilde{A}).$$

Per determinare il rango della matrice A, cerchiamo i valori del parametro h che annullano il determinante di A; risulta:

$$|A| = 0 \quad \Leftrightarrow h(h^2 - 2) = 0 \quad \Leftrightarrow h = 0,\ h = \pm\sqrt{2}.$$

Poiché $r(\tilde{A}) \leq 3$, possiamo concludere che se $h \neq 0$ e $h \neq \pm\sqrt{2}$ risulta:

$$\mathrm{rg}(A) = 3 = \mathrm{rg}(\tilde{A}).$$

Sia,ora, $h = 0$, abbiamo $A = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & -1 & 0 \end{pmatrix}$: osserviamo che i vettori colonna A^1 e $A^2 2$ sono linearmente indipendenti, inoltre si ha $A^3 = -A^1$ e $B = A^1$. Quindi risulta:

$$\mathrm{rg}(A) = \mathrm{rg}(\tilde{A}) = 2.$$

Sia, infine, $h = \pm\sqrt{2}$ abbiamo $A = \begin{pmatrix} \pm\sqrt{2} & 1 & 0 \\ 1 & \pm\sqrt{2} & -1 \\ 0 & -1 & \pm\sqrt{2} \end{pmatrix}$: osserviamo che i vettori A^1, A^2 e B sono linearmente indipendenti; infatti, si ha:

$$\det \begin{pmatrix} \pm\sqrt{2} & 1 & 0 \\ 1 & \pm\sqrt{2} & 1 \\ 0 & -1 & 0 \end{pmatrix} \neq 0.$$

Risulta quindi:

$$\mathrm{rg}(A) = 2 \quad r(\tilde{A}) = 3.$$

Possiamo quindi concludere che il sistema ammette soluzioni reali se e solo se $h \neq \pm\sqrt{2}$, rispondendo al quesito (1).
2 Ricordiamo che, per i valori del parametro h per cui il sistema è compatibile, l'insieme delle soluzioni del sistema è la varietà lineare

$$\mathcal{V} = \mathrm{Ker}\, L_A + X_0,$$

dove $\operatorname{Ker} L_A$ è il nucleo dell'operatore $f_A\colon \mathbb{R}^3 \to \mathbb{R}^3$ associato alla matrice A, ossia definito da $L_A(X) = AX\ \forall X \in \mathbb{R}^3$, $X_0 \in \mathbb{R}^3$ è una soluzione del sistema, i.e.

$$AX_0 = B.$$

La dimensione della varietà $\mathcal{V}$ è la dimensione del sottospazio $Kerf_A$:

$$\dim \mathcal{V} = \dim Kerf_A = 3 - \operatorname{rg}(A).$$

Il sistema pertanto ammette un'unica soluzione se e solo se $\mathcal{V} = \{X_0\}$, se e solo se $\dim \mathcal{V} = 0$, se e solo se $\operatorname{rg}(A) = 3$. Quindi $h \neq 0$ e $h \neq \pm\sqrt{2}$.
3 La varietà $\mathcal{V}$ è una retta in $\mathbb{R}^3$ se e solo se $\dim \mathcal{V} = 1$, se e solo se $\operatorname{rg}(A) = 2$, se e solo se $h = 0$.
4 Osserviamo che poiché risulta $\forall h\ \ \operatorname{rg}(A) \geq 2$, allora $\dim \mathcal{V} \leq 1$. Concludendo non esistono valori reali di h per cui $\mathcal{V}$ sia un piano in $\mathbb{R}^3$.

Esercizio 7.10. Dato il sistema lineare $\begin{cases} (2h-1)x - hy + 2z = 2 - h \\ hx - y + (h+1)z = 2h - h^2 \end{cases}$,
stabilire per quali valori del parametro reale h

(1) il sistema ha soluzioni reali;
(2) l'insieme delle soluzioni è una retta in $\mathbb{R}^3$;
(3) l'insieme delle soluzioni è un piano in $\mathbb{R}^3$.

Risoluzione. Il sistema lineare è non omogeneo, ha 2 equazioni in 3 incognite. In forma matriciale il sistema si scrive $AX = B$, dove la matrice A dei coefficienti, la colonna B dei termini noti e il vettore X delle incognite sono i seguenti:

$$A = \begin{pmatrix} 2h-1 & -h & 2 \\ h & -1 & h+1 \end{pmatrix} \quad B = \begin{pmatrix} 2-h \\ 2h-h^2 \end{pmatrix} \quad X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

Idicata con $\tilde{A}$ la matrice completa, ottenuta aggiungendo ad A la colonna B, per il teorema di Rouché Capelli il sistema ammette soluzioni se e solo si ha:

$$\operatorname{rg}(A) = \operatorname{rg}(\tilde{A}).$$

Osserviamo che $\operatorname{rg}(A) \leq 2$ e $\operatorname{rg}(\tilde{A}) \leq 2$. Consideriamo il seguente minore di ordine 2 di A:

$$\Delta = det \begin{pmatrix} 2h-1 & -h \\ h & -1 \end{pmatrix} = (h-1)^2,$$

poiché $\Delta = 0$ se e solo se $h = 1$, possiamo concludere che se $h \neq 1$ allora risulta:

$$\operatorname{rg}(A) = 2 = \operatorname{rg}(\tilde{A}),$$

e il sistema ammette soluzioni reali.

Se $h = 1$, le matrici A ed $\tilde{A}$ sono le seguenti:

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

risulta, pertanto

$$\operatorname{rg}(A) = 1 = \operatorname{rg}(\tilde{A}),$$

il sistema ammette quindi soluzioni reali. La risposta al quesito (1) è la seguente: il sistema ammette soluzioni reali per ogni $h \in \mathbb{R}$.
(2) Ricordiamo che l' insieme delle soluzioni del sistema è la varietà lineare

$$\mathcal{V} = \operatorname{Ker} f_A + X_0,$$

dove $\operatorname{Ker} L_A$ è il nucleo dell'applicazione lineare $L_A \colon \mathbb{R}^3 \to \mathbb{R}^2$ data da $L_A(X) = AX$, $\forall X \in \mathbb{R}^3$, $X_0 \in \mathbb{R}^3$ è una soluzione del sistema, i.e.

$$AX_0 = B.$$

La dimensione della varietà $\mathcal{V}$ è la dimensione del sottospazio $\operatorname{Ker} L_A$:

$$\dim \mathcal{V} = \dim KerL_A = 3 - \operatorname{rg}(A).$$

Osserviamo che $\mathcal{V}$ è una retta in $\mathbb{R}^3$ se e solo se $\dim \mathcal{V} = 1$, se e solo se $r(A) = 2$: ci ò avviene per ogni $h \neq 1$.
Osserviamo che se $h = 1$, poiché $\operatorname{rg}(A) = 1$, $\dim \mathcal{V} = 2$, quindi $\mathcal{V}$ è il piano in $\mathbb{R}^3$ di equazione

$$x - y + 2z = 1.$$

Esercizio 7.11. Dato il sistema lineare $\begin{cases} h\,x + y + z = 1 \\ (1+h)x + y + k\,z = h^2 - 2 \\ -x + h\,y + (h+k)z = k \end{cases}$, discuterne la solubilità al variare dei parametri $h, k \in \mathbb{R}$.

Risoluzione. La matrice dei coefficienti del sistema

$$A = \begin{pmatrix} h & 1 & 1 \\ 1+h & 1 & k \\ -1 & h & h+k \end{pmatrix}$$

è quadrata; calcoliamo direttamente il suo determinante:

$$\det(A) = \begin{vmatrix} h & 1 & 1 \\ 1+h & 1 & k \\ -1 & h & h+k \end{vmatrix} = h^2 - kh^2 - 2k + 1 = h^2 + 1 - k(h^2 + 2)$$

Essendo $h^2 + 2 > 0$, per $k \neq \frac{h^2+1}{h^2+2}$ $|A| \neq 0$, quindi $\operatorname{rg}(A) = 3$, e $\operatorname{rg}(\tilde{A}) = \operatorname{rg}(A|B) = 3$, poiché la matrice completa è di ordine 3×4, quindi non può avere rango superiore a 3. In questo caso il sistema ammette soluzione, anzi, un'unica soluzione, perché il sistema è quadrato di ordine $3 = \operatorname{rg}(A)$.

Per $k = \frac{h^2+1}{h^2+2}$, la matrice A diventa

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

considerando la sottomatrice 2×2 che contiene le prime due righe e le prime due colonne, si vede subito che $\operatorname{rg}(A) = 2$: infatti

$$\begin{vmatrix} h & 1 \\ 1+h & 1 \end{vmatrix} = h - (1+h) = -1 \neq 0\,.$$

Determiniamo il rango della matrice completa con $k = \frac{h^2+1}{h^2+2}$

$$(A|B) = \begin{pmatrix} h & 1 & 1 & 0 \\ 1+h & 1 & \frac{h^2+1}{h^2+2} & h^2-2 \\ -1 & h & h+\frac{h^2+1}{h^2+2} & \frac{h^2+1}{h^2+2} \end{pmatrix} \tag{9.43}$$

Studiamo il determinante della sottomatrice formata dalla prima, seconda e ultima colonna:

$$\begin{vmatrix} h & 1 & 0 \\ 1+h & 1 & h^2-2 \\ -1 & h & \frac{h^2+1}{h^2+2} \end{vmatrix} = -\frac{(h^2+1)(h^4-3)}{h^2+2}\,; \tag{9.44}$$

questo si annulla per $h = \pm\sqrt[4]{3}$.

Si verifica facilmente che per questi valori del parametro h il rango della matrice completa resta $\mathrm{rg}(A|B) = 2$, quindi il sistema ammette soluzioni. Altrimenti, $\mathrm{rg}(A|B) = 3$ ed il sistema non ammette soluzioni.

---

8. Autovalori ed autovettori di un operatore; diagonalizzazione

Gli esercizi proposti in questa sezione si riferiscono ai contenuti del Capitolo 6.

Esercizio 8.1. (4 febbraio 2009, prova in itinere)
Si consideri la seguente matrice reale di ordine 3:

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

(1) Calcolare il rango di A:
(2) Verificare che il vettore $\boldsymbol{v} = (1, -1, 0)$ è autovettore di A, calcolarne l'autovalore relativo:
(3) Scrivere le equazioni degli autospazi di A:

(4) Determinare una base di $\mathbb{R}^3$ formata da autovettori di A:

(5) Scrivere una matrice B con lo stesso polinomio caratteristico di A che non sia simile ad A:

Esercizio 8.2. (4 febbraio 2009, appello)
Si considerino la matrice A e la matrice B dipendente dal parametro reale k:

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \qquad B = \begin{pmatrix} 0 & k & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 3 \end{pmatrix}$$

(1) Il vettore $\boldsymbol{v} = (1,1,1)$ è autovettore di A, determinarne l'autovalore relativo:
(2) Determinare una base per l'autospazio associato all'autovalore $\alpha = 0$:

(3) Determinare al variare di k gli autovalori di B e le relative molteplicità algebriche e geometriche:
(4) Stabilire per quali valori di k le matrici A e B sono simili:

Esercizio 8.3. (19 febbraio 2009, appello)
Si considerino la matrice A dipendente dal parametro reale k:

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

(1) Determinare gli autovalori della matrice A:
(2) Determinare le molteplicità algebriche e geometriche degli autovalori, al variare di k:
(3) Determinare per quali valori di k la matrice A è diagonalizzabile:
(4) Posto $k = 1$, determinare una base per ciascuno degli autospazi

---

Esercizio 8.4. (19 febbraio 2009, appello)
Si consideri la matrice

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

(1) Determinare gli autovalori di A e le loro molteplicità algebriche:
$\lambda_1 = -1$, $\rho(\lambda_1) = 1$; $\lambda_2 = 3$, $\rho(\lambda_2) = 2$.
(2) Scrivere per ogni autovalore le equazioni del corrispondente autospazio:
$V_{\lambda_1} = \{(x,y,z) \in \mathbb{R}^3 \,|\, x + y = 0, x + 2y + 4z = 0\}$.
$V_{\lambda_2} = \{(x,y,z) \in \mathbb{R}^3 \,|\, x = 0, y = 0\}$.
(3) Gli autovalori di A sono regolari? Giustificare la risposta.
λ_1 sì: $\rho(\lambda_1) = \gamma(\lambda_1) = 1$ (molteplicità algebrica = molteplicità geometrica).
λ_2 no: $\rho(\lambda_2) = 2 \neq \gamma(\lambda_2) = 1$ (moltepl. alg. $\neq$ moltepl. geom.).
(4) Scrivere una matrice B avente lo stesso polinomio caratteristico di A, che non sia simile ad A:

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

Esercizio 8.5. (16 settembre 2009, appello)

Si consideri la seguente matrice:

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

(1) Calcolare $|A|^5 = 0$
(2) Determinare gli autovalori di A e le loro molteplicità algebriche: $\lambda_1 = 0$, $\lambda_3 = \sqrt{3}$, $\lambda_2 = -\sqrt{3}$. Tutti con molteplicità algebrica e geometrica unitaria.
(3) Il vettore $\boldsymbol{v} = (1, 1, -1)$ è autovettore di A? Giustificare la risposta. Sì: basta moltiplicare la matrice A per il vettore $\boldsymbol{v}$. Il risultato è il vettore nullo, quindi tale vettore è autovettore con autovalore 0.
(4) Determinare una base di $\mathbb{R}^3$ formata da autovettori di A:

$$\mathcal{B} = \left\{ \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} \frac{1}{2-\sqrt{3}} \\ \frac{2}{1-\sqrt{3}} \\ 1 \end{pmatrix}, \begin{pmatrix} \frac{1}{2+\sqrt{3}} \\ \frac{2}{1+\sqrt{3}} \\ 1 \end{pmatrix} \right\}$$

Esercizio 8.6. (23 novembre 2009, appello straord.)
Si consideri la seguente matrice:

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

(1) Calcolare $\det A^7 = 0$
(2) Determinare gli autovalori λ di A e le loro molteplicità algebriche μ:

$$\lambda_1 = 0\, \mu_1 = 2\,, \lambda_2 = 1\, \mu_2 = 12\,.$$

(3) Determinare le equazioni per l'autospazio di ciascun autovalore trovato:

$$V_{\lambda_1} = V_0 : \quad x - y = z = 0\,,$$
$$V\lambda_2 = V_1 \quad x = y = 0\,.$$

(4) Precisare se la matrice è diagonalizzabile, motivando la risposta.
Non è diagonalizzabile, perché $\dim V_0 + \dim V_1 = 2 \neq 3 = \dim \mathbb{R}^3$.

Esercizio 8.7. (11 febbraio 2010)
Si consideri la seguente matrice reale quadrata di ordine 4:

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

(1) Determinare gli autovalori della matrice A e le loro molteplicità algebriche:

$$\lambda_1 = 0,\ \mu_1 = 1 \qquad \lambda_2 = 2,\ \mu_2 = 3$$

(2) Determinare dimensione ed equazioni cartesiane di ciascun autospazio di A: dimensioni (molteplicità geometriche): $\nu_1 = 1, \quad \nu_2 = 2$

$$V_0 = \left\{ \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \mid x + y = z = t = 0 \right\}, V_2 = \left\{ \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \mid x = y = 0 \right\}$$

(3) Determinare una base per ciascun autospazio di A:

$$\mathcal{B}_{V_0} = \left\{ \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix} \right\}, \mathcal{B}_{V_2} = \left\{ \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\},$$

(4) A è diagonalizzabile? Giustificare la risposta. NO: $\nu_2 = 2 \neq \mu_2 = 3$

Esercizio 8.8. (14 settembre 2010)
Si consideri la seguente matrice reale quadrata di ordine 4 dipendente dal parametro $h \in \mathbb{R}$:

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

(1) Determinare gli autovalori della matrice A:

$$\lambda_1 = -1, \quad \lambda_2 = 1, \quad \lambda_3 = 3.$$

(2) Determinare per quale/i valori di h la matrice A è diagonalizzabile:

$$h = 0$$

(3) Posto $h = 0$, determinare dimensione, equazioni cartesiane ed una base per ciascun autospazio di A:

$$\dim(V_{-1}) = 1,\ V_{-1} = \{(x, y, z, t) \in \mathbb{R}^4 \mid x = y = z + t = 0\},\ \mathcal{B}_{V_{-1}} = \left\{ \begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix} \right\}$$

$$\dim(V_1) = 1,\ V_1 = \{(x, y, z, t) \in \mathbb{R}^4 \mid x + y = z = t = 0\},\ \mathcal{B}_{V_1} = \left\{ \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix} \right\}$$

$\dim(V_3) = 2$, $V_3 = \{(x, y, z, t) \in \mathbb{R}^4 \,|\, x - y = z - t = 0\}$, $\mathcal{B}_{V_3} = \left\{ \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} \right\}$

(4) Determinare per quale valore di h esiste una base ortogonale di $\mathbb{R}^4$ formata da autovettori di A: $h = 0$.

Esercizio 8.9. Sia $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ la base standard di $\mathbb{R}^3$, si consideri l'operatore lineare $L\colon \mathbb{R}^3 \to \mathbb{R}^3$ definito da:

$$L(\boldsymbol{e}_1) = \boldsymbol{e}_1 - 2\boldsymbol{e}_3 \quad L(\boldsymbol{e}_2) = \boldsymbol{e}_1 + \boldsymbol{e}_2 - 2\boldsymbol{e}_3 \quad L(\boldsymbol{e}_3) = 3\boldsymbol{e}_3.$$

(1) Determinare gli autovalori di L e le loro molteplicità algebriche;
(2) scrivere le equazioni degli autospazi;
(3) determinare una base per ciascun autospazio;
(4) è possibile trovare una base di $\mathbb{R}^3$ formata da autovettori di L?

Risoluzione. L'operatore $f\colon \mathbb{R}^3 \to \mathbb{R}^3$ è univocamente determinato dalle immagini dei vettori della base; infatti:

$$\begin{aligned} &L(x\boldsymbol{e}_1 + y\boldsymbol{e}_2 + z\boldsymbol{e}_3) \\ &= xL(\boldsymbol{e}_1) + yL(\boldsymbol{e}_2) + zL(\boldsymbol{e}_3) \\ &= x(\boldsymbol{e}_1 - 2\boldsymbol{e}_3) + y(\boldsymbol{e}_1 + \boldsymbol{e}_2 - 2\boldsymbol{e}_3) + z(3\boldsymbol{e}_3) \\ &= (x + y)\boldsymbol{e}_1 + ye_2 + (3z - 2x)\boldsymbol{e}_3, \end{aligned} \tag{9.45}$$

da cui ricaviamo

$$L(x, y, z) = (x + y, y, 3z - 2x) = A \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \quad A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ -2 & -2 & 3 \end{pmatrix}.$$

(1) Ricordiamo che gli autovalori di L sono le soluzioni dell'equazione caratteristica di L, che si ottiene uguagliando a zero il polinomio caratteristico di f. Il polinomio caratteristico di f è:

$$p_L(t) = p_A(t) = |A - tI_3| = \det \begin{pmatrix} 1-t & 1 & 0 \\ 0 & 1-t & 0 \\ -2 & -2 & 3-t \end{pmatrix} = (1 - t)^2(3 - t),$$

osserviamo che $p_L(t)$ è totalmente decomponibile in $\mathbb{R}$. L'equazione caratteristica di L è quindi la seguente:

$$(1 - t)^2(3 - t) = 0\,,$$

le cui soluzioni sono $t = 1$ con molteplicità algebrica $\mu(1) = 2$ e $t = 3$ con moplteplicità algebrica $\mu(3) = 1$.

(2), (3) Sia $t = 1$; l'autospazio V_1 è per definizione il seguente sottospazio di $\mathbb{R}^3$:

$$V_1 = \left\{\boldsymbol{v} \in \mathbb{R}^3 | \;\; f(\boldsymbol{v}) = \boldsymbol{v}\right\} = \left\{ (x, y, z) \in \mathbb{R}^3 | \;\; (A - I_3) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0_{\mathbb{R}^3} \right\},$$

con $\dim V_1 = 3 - \mathrm{rg}(A - I_3)$. Risulta:

$$A - I_3 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ -2 & -2 & 2 \end{pmatrix},$$

quindi $\mathrm{rg}(A - I_3) = 2$ e di conseguenza $\dim V_1 = 1$. Ricordiamo che la molteplicità geometrica di un autovalore α è $m(\alpha) = \dim V_\alpha$ e che l'autovalore è regolare se e solo se risulta $m(\alpha) = \mu(\alpha)$. Poiché risulta $m(1) < \mu(1)$ l'autovalore $t = 1$ non è regolare.

L'autospazio V_1 è la retta per l'origine di equazioni

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

pertanto $V_1 = \mathrm{Span}(v)$, dove $\boldsymbol{v} = \big(1, 0,\big)$. Una base per V_1 è data dal vettore non nullo $\boldsymbol{v}$.

Sia $t = 3$, l'autospazio V_3 è per definizione il seguente sottospazio di $\mathbb{R}^3$:

$$V_3 = \{v \in \mathbb{R}^3 |\ f(v) = 3v\} = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 |\ (A - 3I_3) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \mathbf{0}_{\mathbb{R}^3} \right\},$$

con $\dim V_3 = 3 - \mathrm{rg}(A - 3I_3)$. Risulta:

$$A - 3I_3 = \begin{pmatrix} -2 & 1 & 0 \\ 0 & -2 & 0 \\ -2 & -2 & 0 \end{pmatrix},$$

quindi $\mathrm{rg}(A - 3I_3) = 2$ e di conseguenza $\dim V_3 = 1$. L'autovalore $t = 3$ è regolare in quanto autovalore semplice (i.e. $m(3) = 1 = \mu(3)$). L'autospazio V_3 è la retta per l'origine di equazioni

$$\begin{cases} y = 0 \\ x = 0 \end{cases},$$

pertanto $V_3 = \mathrm{Span}(\boldsymbol{e}_3)$. Una base per V_3 è data dal vettore $\boldsymbol{e}_3$.

(4) Ricordiamo che esiste una base di $\mathbb{R}^3$ formata da autovettori di L se e solo se L è diagonalizzabile. Condizione necessaria e sufficiente affinché L sia diagonalizzabile è che il polinomio caratteristico di L sia totalmente decomponibile e che gli autovalori di L siano regolari.

Abbiamo osservato che la prima condizione è verifcata, tuttavia l'autovalore $t = 1$ non è regolare, quindi L non è diagonalizzabile.

Esercizio 8.10. Considerare le seguenti matrici quadrate reali di ordine 3:

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

(1) Verificare che A e B hanno lo stesso polinomio caratteristico.
(2) Le matrici A e B sono simili?

Risoluzione. (1) Calcoliamo rispettivamente il polinomio caratteristico delle matrici A e B:

$$p_A(t) = |A - tI_3| = \det \begin{pmatrix} 1-t & 0 & 1 \\ 0 & 1-t & 1 \\ 0 & 0 & 2-t \end{pmatrix} = (1-t)^2(2-t),$$

$$p_B(t) = |B - tI_3| = \det \begin{pmatrix} 1-t & 1 & 0 \\ 0 & 2-t & 0 \\ 1 & 0 & 1-t \end{pmatrix} = (1-t)^2(2-t),$$

osserviamo che $p_A(t) = p_B(t)$ nell'anello dei polinomi $\mathbb{R}[t]$.
(2) Ricordiamo che la condizione appena verificata è una condizione necesssaria, ma non sufficiente per la relazione di similitudine tra matrici. Cerchiamo di avere maggiori informazioni sulle matrici date, ad esempio se siano diagonalizzabili. Osserviamo che entrambe le matrici hanno polinomio caratteristico totalmente decomponibile in $\mathbb{R}$, quindi rimane da verificare se gli autovalori siano regolari. Gli autovalori delle matrici sono $t = 1$ con $\mu(1) = 2$ e $t = 2$ con $\mu(2) = 1$. L'autovalore $t = 2$é semplice, quindi sicuramente regolare, $m(2) = \mu(2) = 1$.

Analizziamo l'autovalore $t = 1$. Ricordiamo che l'autospazio $V_{1,A}$ è il seguente sottospazio di $\mathbb{R}^3$:

$$V_{1,A} = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 |\ (A - I_3) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \mathbf{0}_{\mathbb{R}^3} \right\},$$

con $\dim V_{1,A} = 3 - \mathrm{rg}(A - I_3)$. Poiché risulta:

$$A - I_3 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

abbiamo $r(A - I_3) = 1$ e quindi $m(1) = \dim V_{1,A} = 2$. Poiché risulta $m(1) = 2 = \mu(1)$, l'autovalore $t = 1$ è regolare per la matrice A. La matrice A è diagonalizzabile e quindi simile alla matrice $\Delta = diag(1, 1, 2)$.

Analogamente l'autospazio $V_{1,B}$ è il seguente sottospazio:

$$V_{1,B} = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \colon (B - I_3) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \mathbf{0}_{\mathbb{R}^3} \right\},$$

con $\dim V_{1,B} = 3 - \mathrm{rg}(B - I_3)$. Poiché risulta:

$$B - I_3 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

abbiamo $r(B - I_3) = 2$ e quindi $m(1) = \dim V_{1,B} = 1$. Poiché risulta $m(1) = 1 < \mu(1)$, l'autovalore $t = 1$ non è regolare per la matrice B. La matrice B non è pertanto diagonalizzabile.

Possiamo concludere che le matrici A e B non sono simili: infatti se B fosse simile ad A, per la proprietà transitiva della similitudine B sarebbe simile anche alla matrice diagonale Δ e quindi B sarebbe diagonalizzabile.

Esercizio 8.11. Sia $L\colon \mathbb{R}^3 \to \mathbb{R}^3$ un operatore lineare con la seguente proprietà: esiste un numero reale non nullo α tale che ogni vettore non nullo $\boldsymbol{v} \in \mathbb{R}^3$ è un autovettore associato all'autovalore α.

(1) Descrivere l'operatore L;
(2) scrivere la matrice associata ad L rispetto ad una qualsiasi base $\mathcal{B}$ di $\mathbb{R}^3$.

Risoluzione. (1) Sia $\boldsymbol{v} \in \mathbb{R}^3$ un vettore non nullo, poiché $\boldsymbol{v}$ è un autovettore di L associato all'autovalore α, si ha

$$L(\boldsymbol{v}) = \alpha\boldsymbol{v},$$

se $\boldsymbol{v} = \mathbf{0}_{\mathbb{R}^3}$, poiché L è un endomorfismo, si ha $L(\mathbf{0}_{\mathbb{R}^3}) = \mathbf{0}_{\mathbb{R}^3}$. Posto $\boldsymbol{v} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ otteniamo che L è definito da

$$L\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \alpha\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \alpha x \\ \alpha y \\ \alpha z \end{pmatrix}.$$

Ricordiamo che il nucleo di L per definizione è il seguente sottospazio di $\mathbb{R}^3$

$$\operatorname{Ker} f = \left\{ \boldsymbol{v} \in \mathbb{R}^3 \mid f(\boldsymbol{v}) = \mathbf{0}_{\mathbb{R}^3} \right\},$$

osserviamo che $\operatorname{Ker} L = \{\mathbf{0}_{\mathbb{R}^3}\}$, infatti

$$\alpha\boldsymbol{v} = \mathbf{0}_{\mathbb{R}^3} \Longrightarrow \boldsymbol{v} = 0_{\mathbb{R}^3}$$

essendo $\alpha \neq 0$ per ipotesi. Ciò implica che L è un endomorfismo iniettivo e quindi un automorfismo di $\mathbb{R}^3$. Infine ricordiamo che

$$\operatorname{Im} L = \left\{ \boldsymbol{w} \in \mathbb{R}^3 \mid \exists \boldsymbol{v} \in \mathbb{R}^3 \colon f(\boldsymbol{v}) = \boldsymbol{w} \right\},$$

possiamo quindi concludere che $\operatorname{Im} L = \mathbb{R}^3$.
(2) Sia $\mathcal{B} = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$ una base di $\mathbb{R}^3$, calcoliamo le immagini dei vettori della base:

$$L(\boldsymbol{v}_1) = \alpha\boldsymbol{v}_1 \quad L(\boldsymbol{v}_2) = \alpha\boldsymbol{v}_2 \quad L(\boldsymbol{v}_3) = \alpha\boldsymbol{v}_3.$$

La matrice A associata ad L rispetto alla base $\mathcal{B}$ ha come colonne, rispettivamente, le coordinate delle immagini dei vettori della base:

$$A = \begin{pmatrix} \alpha & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & \alpha \end{pmatrix},$$

ciè $A = \alpha I(3)$, la matrice identità di ordine 3.

Esercizio 8.12. Considerare la seguente matrice reale quadrata di ordine 4:

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

(1) Scrivere il polinomio caratteristico di A;
(2) la matrice A è diagonalizzabile?
(3) scrivere la matrice A^{-1} utilizando il Teorema di Cayley Hamilton.

RISOLUZIONE. (1) Il polinomio caratteristico di A è :

$$p_A(t) = |A - tI(4)| = \det \begin{pmatrix} 1-t & 1 & 1 & 1 \\ 0 & -1-t & 1 & 1 \\ 0 & 0 & -1-t & 1 \\ 0 & 0 & 0 & 1-t \end{pmatrix} = (1-t)^2(-1-t)^2,$$

osserviamo che è totalmente decomponibile in $\mathbb{R}$.
(2) Gli autovalori di A sono le soluzioni dell'equazione caratteristica di A:

$$(1-t)^2(-1-t)^2 = 0$$

le cui soluzioni sono $t = 1$ con $\mu(1) = 2$ e $t = -1$ con $\mu(-1) = 2$. La matrice A è diagonalizzabile se e solo se il polinomio caratteristico di A è totalmente decomponibile e gli autovalori di A sono regolari. La prima condizione è verificata, rimane da verificare la regolarità degli autovalori.

Sia $t = 1$, l'autospazio associato è il seguente sottospazio di $\mathbb{R}^4$:

$$V_1 = \left\{ X \in \mathbb{R}^4 |\ AX = X \right\} = \left\{ X \in \mathbb{R}^4 |\ (A - I_4)X = \vec{0}_{\mathbb{R}^4} \right\}$$

con $\dim V_1 = 4 - \mathrm{rg}(A - I(4))$. Risulta:

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

abbiamo quindi $r(A - I_4) = 3$, per cui $m(1) = \dim V_1 = 1$. Poiché risulta $m(1) < \mu(1)$, l'autovalore $t = 1$ non è regolare, quindi A non è diagonalizzabile.
(3) Il teorema di Cayley Hamilton afferma che ogni matrice reale quadrata A di ordine n soddisfa il proprio polinomio caratteristico. Ricordiamo che il polinomio caratteristico di A è il seguente polinomio:

$$p_A(t) = t^4 - 2t^2 + 1,$$

allora la matrice A è una radice del seguente polinomio matriciale

$$p_A(X) = X^4 - 2X^2 + I(4),$$

cioé risulta

$$p_A(A) = A^4 - 2A^2 + I(4) = 0_n.$$

Possiamo quindi scrivere la matrice $I(4)$ come combinazione lineare di potenze di A:

$$I(4) = 2A^2 - A^4,$$

moltiplicando entrambi i membri per la matrice A^{-1} otteniamo:

$$A^{-1} = 2A - A^3.$$

Calcoliamo la matrice A^3:

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

Infine otteniamo la matrice A^{-1}:

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

Esercizio 8.13. Siano A e B matrici reali quadrate di ordine n, verificare che:

(1) se $X \in \mathbb{R}^n$ è autovettore di A associato all'autovalore α ed è autovettore di B associato all'autovalore β, allora X è autovettore di $A+B$ associato all'autovalore $\alpha+\beta$;
(2) se $X \in \mathbb{R}^n$ è autovettore di A associato all'autovalore α, allora per ogni $m \geq 2$, X è autovettore di A^m associato all'autovalore α^m;
(3) sia $|A| \neq 0$: se $X \in \mathbb{R}^n$ è autovettore di A associato all'autovalore α, allora X è autovettore di A^{-1} associato all'autovalore α^{-1}.

Risoluzione. (1) Poiché X è autovettore di A associato all'autovalore α si ha:

$$AX = \alpha X,$$

poiché X è autovettore di B associato all'autovalore β si ha:

$$BX = \beta X.$$

Abbiamo quindi:

$$(A+B)X = AX + BX = \alpha X + \beta X = (\alpha+\beta)X,$$

da cui deduciamo che X è autovettore di $A+B$ associato all' autovalore $\alpha+\beta$.
(2) Dimostriamo la proprietà per induzione su m. Sia $m=2$, si ha

$$A^2X = A(AX) = A(\alpha X) = \alpha(AX) = \alpha(\alpha X) = (\alpha)^2 X,$$

la proprietà è vera.

Supponiamo che la proprietà sia verificata per $m-1$: $A^{m-1}X = \alpha^{m-1}X$, proviamo che vale per m. Infatti si ha:

$$A^mX = A(A^{m-1}X) = A(\alpha^{m-1}X) = \alpha^{m-1}(AX) = \alpha^{m-1}(\alpha X) = \alpha^m X,$$

che prova la tesi.
(3) Poiché $|A| \neq 0$ la matrice A è invertibile; indicata con A^{-1} l 'inversa di A, si ha

$$I(n) = A^{-1} \cdot A.$$

Sia $X \in \mathbb{R}^n$ un autovettore di A associato all'autovalore α, risulta:

$$I_nX = (A^{-1} \cdot A)X = A^{-1}(AX) = A^{-1}(\alpha X) = \alpha A^{-1}X.$$

Osserviamo che essendo $|A| \neq 0$, si ha $\alpha \neq 0$, si ottiene allora:

$$A^{-1}X = \alpha^{-1}X,$$

quindi X è autovettore di A^{-1} associato all'autovalore α^{-1}.

Esercizio 8.14. Si consideri la seguente matrice reale quadrata di ordine 3, con k parametro reale:

$$B = \begin{pmatrix} 1 & k & 1 \\ 0 & k+1 & 1 \\ 0 & 0 & -1 \end{pmatrix},$$

determinare per quali valori di k la matrice B è diagonalizzabile.

Risoluzione. Calcoliamo il polinomio caratteristico di B:

$$p_B(t) = |B - tI_3| = \det \begin{pmatrix} 1-t & k & 1 \\ 0 & k+1-t & 1 \\ 0 & 0 & -1-t \end{pmatrix} = (1-t)(-1-t)(k+1-t).$$

Osserviamo che il polinomio caratteristico è totalmente decomponibile in $\mathbb{R}$. L'equazione caratteristica di B è:

$$(1-t)(-1-t)(k+1-t) = 0,$$

gli autovalori sono quindi $t = 1$, $t = -1$ e $t = k+1$.

Osserviamo che se $k \neq 0$ e $k \neq -2$, allora la matrice B ha 3 autovalori distinti (semplici) e quindi è diagonalizzabile (condizione sufficiente per la diagonalizzazione).

Sia $k = 0$: abbiamo $B = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & -1 \end{pmatrix}$, gli autovalori di B sono $t = 1$ con $\mu(1) = 2$ e $t = -1$ con $\mu(-1) = 1$. Condizione necessaria e sufficiente affiché B sia diagonalizzabile è che il polinomio caratteristico di B sia totalmente decomponibile in $\mathbb{R}$ e che gli autovalori siano regolari. Poiché $t = -1$ è semplice, quindi regolare, rimane da verificare che $t = 1$ sia regolare. L'autospazio V_1 è il seguente sottospazio di $\mathbb{R}^3$:

$$V_1 = \left\{ X \in \mathbb{R}^3 | \ BX = X \right\} = \left\{ X \in \mathbb{R}^3 | \ (B - I(3))X = 0_{\mathbb{R}^3} \ \right\}$$

con $\dim V_1 = 3 - \mathrm{rg}(B - I(3))$. Risulta:

$$B - I(3) = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & -2 \end{pmatrix},$$

si ha quindi $r(B - I_3) = 1$, da cui $m(1) = \dim V_1 = 2 = \mu(1)$: l'autovalore $t = 1$ è regolare e quindi la matrice B è diagonalizzabile.

Sia $k = -2$: abbiamo $B = \begin{pmatrix} 1 & -2 & 1 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{pmatrix}$, gli autovalori di B sono $t = 1$ con $\mu(1) = 1$ e $t = -1$ con $\mu(-1) = 2$. Condizione necessaria e sufficiente affiché B sia diagonalizzabile è che il polinomio caratteristico di B sia totalmente decomponibile in $\mathbb{R}$ e che gli autovalori siano regolari. Poiché $t = 1$ è semplice, quindi regolare, rimane da verificare che $t = -1$ sia regolare. L'autospazio V_{-1} è il seguente sottospazio di $\mathbb{R}^3$:

$$V_{-1} = \left\{X \in \mathbb{R}^3 | \ BX = -X\right\} = \left\{X \in \mathbb{R}^3 | \ (B + I_3)X = 0_{\mathbb{R}^3} \ \right\}$$

con $\dim V_1 = 3 - \text{rg}(B + I(3))$. Risulta:

$$B + I(3) = \begin{pmatrix} 2 & -2 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix},$$

si ha quindi $\text{rg}(B + I_3) = 2$, da cui $m(-1) = \dim V_{-1} = 1 < \mu(1)$: l'autovalore $t = -1$ non è regolare e quindi la matrice B non è diagonalizzabile.

---

9. Prodotto scalare in $\mathbb{R}^3$; ortogonalità

Gli esercizi proposti in questa sezione si riferiscono ai contenuti del Capitolo 7.

Esercizio 9.1. (4 febbraio 2009, prova in itinere)
Fissato in $\mathbb{R}^4$ il prodotto scalare standard, si consideri il sottospazio di $\mathbb{R}^4$

$$U = \left\{ \boldsymbol{u} = \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \in \mathbb{R}^4 | \ x + y + 2z = y + 3z + t = 0 \right\}.$$

Determinare:
(1) $\dim U =$ $\qquad \dim U^\perp =$
(2) una base di $U^\perp$:
(3) una base ortonormale di U:

Esercizio 9.2. (4 febbraio 2009, appello)

Fissati in $\mathbb{R}^4$ il prodotto scalare standard, si considerino i vettori $\boldsymbol{u}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}$ e $\boldsymbol{u}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}$ e i seguenti sottospazi:

$$V = \left\{ \boldsymbol{v} = \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \in \mathbb{R}^4 | \ \ x + y + t = 0 \right\}, \quad U = \text{Span}(\boldsymbol{u}_1, \boldsymbol{u}_2),$$

determinare:

(1) la dimensione di V e una base di V:

(2) $\dim(U \cap V) =$ $\dim(U + V) =$
(3) la proiezione ortogonale di $\boldsymbol{u}_2$ lungo $\mathrm{Span}(\boldsymbol{u}_1)$:
(4) una base ortonormale per U:

Esercizio 9.3. (19 febbraio 2009, appello)
Fissati in $\mathbb{R}^4$ il prodotto scalare standard, si consideri il seguente sottospazio:

$$V = \left\{ \boldsymbol{v} = \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \in \mathbb{R}^4 | \quad x - y + z + 3t = 0 \right\}.$$

Determinare:

(1) la dimensione di V:
(2) Una base ortogonale di V:

(3) Il complemento ortogonale $V^{\perp}$ di V:

(4) una base ortonormale per $V^{\perp}$:

Esercizio 9.4. (6 luglio 2009, appello)
Fissato in $\mathbb{R}^4$ il prodotto scalare standard, si consideri il seguente sottospazio:

$$U = \left\{ \boldsymbol{u} = \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \in \mathbb{R}^4 \colon x + 2y - 3z = 0, t - z = 0 \right\},$$

ed i vettori $v_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ h \end{pmatrix}$ e $v_2 = \begin{pmatrix} 2 \\ -1 \\ 0 \\ 0 \end{pmatrix}$. Determinare:

(1) $\dim U = 2$:
(2) Una base ortogonale di U: $\left\{ \begin{pmatrix} 2 \\ -1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 6 \\ 5 \\ 5 \end{pmatrix} \right\}$.
(3) Le equazioni di $U^{\perp}$: $x + y + z + t = 0; 2x - y = 0$.
(4) Per quali valori di h, $\mathrm{Span}(\boldsymbol{v}_1, \boldsymbol{v}_2)$ è un sottospazio di U: $h = 1$.

Esercizio 9.5. (16 settembre 2009)
Fissati in $\mathbb{R}^4$ il prodotto scalare standard, si considerino i vettori

$$\boldsymbol{v}_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 1 \end{pmatrix}, \text{ e } \boldsymbol{v}_2 = \begin{pmatrix} 2 \\ 1 \\ 1 \\ 0 \end{pmatrix};$$

indichiamo con $V = \mathrm{Span}(\boldsymbol{v}_1, \boldsymbol{v}_2)$ il sottospazio generato dai due vettori. Sia inoltre U il sottospazio definito mediante equazioni cartesiane:

$$U = \left\{ \boldsymbol{u} = \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \in \mathbb{R}^4 \;\middle|\; x - t - z = 0\,,\, z = y + t \right\},$$

Determinare:

(1) la dimensione di $V^\perp$: $\dim(V^\perp) = 2$
(2) le equazioni cartesiane di V:

$$V = \left\{ \boldsymbol{v} = \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \in \mathbb{R}^4 \;\middle|\; x - t - 2z = 0\,,\, z = y + t \right\},$$

(3) una base di U:

$$\mathcal{B}(U) = \left\{ \begin{pmatrix} 1 \\ -1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 1 \\ 1 \end{pmatrix} \right\}$$

(4) Sia $\boldsymbol{w} = \begin{pmatrix} h \\ -1 \\ 1 \\ 2 \end{pmatrix}$; determinare per quale valore di h si ha $\boldsymbol{w} \in V$:

$$h = 4$$

(5) $\dim(U + V) = \dim(U) + \dim(V) - \dim(U \cap V) = 2 + 2 - 1 = 3$

Esercizio 9.6. (23 novembre 2009, appello straord.)

Fissato in $\mathbb{R}^4$ il prodotto scalare standard, si considerino i vettori $\boldsymbol{u}_1 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}$, e $\boldsymbol{u}_2 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$; siano $U = \operatorname{Span}(\boldsymbol{u}_1, \boldsymbol{u}_2)$ e $V = \left\{ \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \in \mathbb{R}^4 \mid x + y = z = 0 \right\}$ sottospazi di $\mathbb{R}^4$. Determinare:

(1) le equazioni cartesiane di U

$$z - y = t = 0\,.$$

(2) una base di $V^\perp$:

$$\left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \right\}.$$

(3) $\dim(U \cap V) = 0$, $\qquad \dim(U + V) = 4$

(4) una base di $U + V$.

$$\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\}.$$

Esercizio 9.7. (14 settembre 2010)
Fissato in $\mathbb{R}^4$ il prodotto scalare standard, si consideri il seguente sottospazio:

$$U = \{\boldsymbol{u} = \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \in \mathbb{R}^4 \colon x - 2y + 3z = z - t = 0\}.$$

Determinare:

(1) $\dim U =$: 2

(2) Una base ortogonale di U:

$$\mathcal{B}_U = \left\{ \begin{pmatrix} -3 \\ 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 4 \\ 11 \\ 6 \\ 6 \end{pmatrix} \right\}$$

(3) Le equazioni di $U^\perp$:

$$U^\perp = \{(x, y, z, t) \in \mathbb{R}^4 \,|\, 3x - z - t = 2x + y = 0\}$$

(4) Una base di di $U^{\perp}$:

$$\mathcal{B}_{U^{\perp}} = \left\{ \begin{pmatrix} 1 \\ -2 \\ 3 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ -2 \\ 0 \\ 3 \end{pmatrix} \right\}$$

---

10. Forme quadratiche, coniche e complementi

Gli esercizi proposti in questa sezione si riferiscono ai contenuti del Capitolo 8.

Esercizio 10.1. (4 febbraio 2009, prova in itinere)
Si consideri la seguente forma quadratica reale $Q\colon \mathbb{R}^3 \to \mathbb{R}$, in funzione del parametro reale k:

$$Q(x, y, z) = x^2 - 2kxy + y^2 + kz^2.$$

(1) Scrivere Q in forma matriciale:

(2) Determinare per quali valori di k la forma Q è definita positiva:

Posto $k = 1$:

(1) scrivere l'equazione canonica di Q:
(2) scrivere la trasformazione ortogonale $X = MX'$ che riduce Q alla forma canonica trovata:

Esercizio 10.2. (4 febbraio 2009, prova in itinere)
Fissato nel piano un sistema di riferimento cartesiano ortogonale $\mathcal{R}(O, \hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}, \hat{\boldsymbol{k}})$, si consideri l'applicazione lineare $\rho\colon \mathbb{R}^2 \to \mathbb{R}^2$, dove $\rho(\boldsymbol{v})$ è ottenuto dalla rotazione (in senso antiorario) di $\boldsymbol{v}$ di un angolo $\vartheta = \frac{\pi}{6}$.

(1) Scrivere la matrice A associata a ρ nella base standard di $\mathbb{R}^2$:

(2) Descrivere l' endomorfismo di $\mathbb{R}^2$ associato alla matrice A^2:

Esercizio 10.3. (4 febbraio 2009, appello)
Si consideri la seguente forma quadratica reale $Q\colon \mathbb{R}^3 \to \mathbb{R}$, in funzione del parametro reale k:

$$Q(x, y, z) = x^2 - 2kxy + y^2 + kz^2.$$

(1) Scrivere Q in forma matriciale:

(2) Determinare per quali valori di k la forma Q è definita positiva:

Posto $k = 1$:

(1) scrivere l'equazione canonica di Q:
(2) scrivere la trasformazione ortogonale $X = MX'$ che riduce Q alla forma canonica trovata:

Esercizio 10.4. (19 febbraio 2009, appello)
Si consideri la conica di equazione

$$-2y^2 + 2\sqrt{3}xy - 1 = 0$$

(1) Riconoscere la conica:
(2) Determinare le equazioni del/degli assi di simmetria:
(3) Esplicitare il cambio di riferimento per avere la conica in forma canonica:

Risoluzione. Costruiamo la matrice A che è associata alla conica:

$$A = \begin{pmatrix} 0 & \sqrt{3} & 0 \\ \sqrt{3} & -2 & 0 \\ 0 & 0 & -1 \end{pmatrix} . \tag{9.46}$$

Il determinante di A è

$$\det(A) = -1 \begin{vmatrix} 0 & \sqrt{3} \\ \sqrt{3} & -2 \end{vmatrix} = -1[0 - (\sqrt{3})^2] = 3 \neq 0$$

quindi la conica non è degenere. Inoltre, il determinante della sottomatrice associata alla forma quadratica è

$$\det(A_2) = \begin{vmatrix} 0 & \sqrt{3} \\ \sqrt{3} & -2 \end{vmatrix} = -3 < 0 \,,$$

quindi la conica è un'iperbole.

Per trovare gli assi di simmetria basta trovare le direzioni degli autovettori della matrice A_2; cerchiamone gli autovalori. Calcoliamo il polinomio caratteristico e troviamone le radici (sicuramente reali, poiché la matrice è reale simmetrica).

$$\det(A_2 - \lambda I(2)) = \begin{vmatrix} 0-\lambda & \sqrt{3} \\ \sqrt{3} & -2-\lambda \end{vmatrix} = \lambda(\lambda+2) - 3 = \lambda^2 + 2\lambda - 3 = 0 \,,$$

che ci dà i due autovalori

$$\lambda_1 = -3 \qquad \lambda_2 = 1 \,.$$

Cerchiamo un autovettore associato al primo autovalore:

$$(A_2 - \lambda_1 I)X = \begin{pmatrix} 3 & \sqrt{3} \\ \sqrt{3} & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} .$$

Questo si riduce ad un'unica equazione per le componenti incognite dell'autovettore:

$$\sqrt{3}x + y = 0 \,, \tag{9.47}$$

ossia

$$y = -\sqrt{3}x \,.$$

Un autovettore $\boldsymbol{v}_1$ si ottiene ponendo $x = 1$:

$$\boldsymbol{v}_1 = \begin{pmatrix} 1 \\ -\sqrt{3} \end{pmatrix} . \tag{9.48}$$

Per il secondo autovettore $\boldsymbol{v}_2$ non è necessario alcun calcolo: essendo associato ad un autovalore diverso, deve risultare ortogonale al primo, possiamo scrivere immediatamente:

$$\boldsymbol{v}_2 = \begin{pmatrix} \sqrt{3} \\ 1 \end{pmatrix} . \tag{9.49}$$

e direzioni dei due autovettori coincidono con quelle degli assi di simmetria dell'iperbole; poiché mancano i termini lineari, non è necessaria alcuna traslazione: gli assi di simmetria sono, pertanto, le rette passanti per l'origine date da:

$$y = \frac{\sqrt{3}}{3} x \tag{9.50}$$

e

$$y = -\sqrt{3} x \, . \tag{9.51}$$

Una base di autovettori $\{\boldsymbol{w}_1, \boldsymbol{w}_2\}$ per la nostra matrice si ottiene semplicemente normalizzando i vettori $\boldsymbol{v}_1$, $\boldsymbol{v}_2$; le loro rappresentazioni nella base canonica originaria sono, dunque:

$$\boldsymbol{w}_1 = \begin{pmatrix} \frac{1}{2} \\ -\frac{\sqrt{3}}{2} \end{pmatrix} , \qquad \boldsymbol{w}_2 = \begin{pmatrix} \frac{\sqrt{3}}{2} \\ \frac{1}{2} \end{pmatrix} ,$$

La matrice M di cambio di base che permette di scrivere le coordinate $X = \begin{pmatrix} x \\ y \end{pmatrix}$ in termini delle nuove coordinate $X' = \begin{pmatrix} x' \\ y' \end{pmatrix}$ è proprio formata da colonne dove si riportano le rappresentazioni nnella vecchia base dei vettori della nuova base:

$$M = \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix} . \tag{9.52}$$

Mediante la trasformazione $X = MX'$ possiamo esplicitare il cambio di variabili per scrivere la forma quadratica in forma canonica:

$$x = \frac{1}{2} x' + \frac{\sqrt{3}}{2} y' \tag{9.53a}$$

$$y = -\frac{\sqrt{3}}{2} x' + \frac{1}{2} y' \, . \tag{9.53b}$$

Si lascia allo studente la verifica che, sostituendo le equazioni precedenti nella scrittura della conica, otteniamo

$$-3(x')^2 + (y')^2 = 1 \, ,$$

che è la forma canonica cercata. Ancora, non essendo necessaria alcuna traslazione aggiuntiva, il cambio di variabile richiesto è dato dalle Equazioni (9.53).

Esercizio 10.5. (6 luglio 2009, appello)
Fissato nel piano un sistema di riferimento cartesiano ortogonale $\mathcal{R}(O, \hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}, \hat{\boldsymbol{k}})$, si consideri la conica di equazione

$$y^2 - 2x + 2y + 3 = 0.$$

(1) Riconoscere la conica:
È una parabola.
(2) Scrivere l'equazione canonica della conica:
$Y^2 = 2X$, con $X = x - 1$, $Y = y + 1$.
(3) Scrivere le equazioni di eventuali assi di simmetria della conica:
$Y = 0$, ossia $y = -1$.

Esercizio 10.6. (23 novembre 2009, appello straord.)
Fissato nel piano un sistema di riferimento cartesiano $\mathcal{R}(O, \hat{\boldsymbol{\imath}}, \hat{\boldsymbol{\jmath}}, \hat{\boldsymbol{k}})$ considerare la conica $\mathcal{C}$ di equazione:

$$\mathcal{C}: \ 3x^2 - 4xy + 3y^2 - 1 = 0\,.$$

(1) Classificare la conica $\mathcal{C}$.
È un ellisse.
(2) Scrivere le equazioni degli assi di simmetria di $\mathcal{C}$.

$$x - y = 0 \qquad x + y = 0\,.$$

(3) Scrivere l'equazione canonica di $\mathcal{C}$, fornendo esplicitamente l'espressione dei cambi di variabile per ottenere tale forma (cambio di base).

$$\begin{pmatrix} x \\ y \end{pmatrix} = X \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix} + Y \begin{pmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}$$

$$X^2 + 5Y^2 = 1$$

Esercizio 10.7. (11 febbraio 2010)
Si consideri la forma quadratica reale $Q\colon \mathbb{R}^3 \to \mathbb{R}$ data da

$$Q \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 3x^2 - 2xy + 3y^2 + 4z^2.$$

(1) Scrivere la matrice associata a Q nella base standard di $\mathbb{R}^3$:

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

.

(2) Scrivere l'equazione canonica di Q: $Q \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = 2x'^2 + 4y'^2 + 4z'^2.$

(3) Scrivere l'equazione del cambiamento di base necessario per ottenere la forma canonica:

$$\begin{cases} x' = \frac{1}{\sqrt{2}}(x+y) \\ y' = \frac{1}{\sqrt{2}}(x-y) \\ z' = z \end{cases}$$

(4) Stabilire se Q è definita positiva (o negativa), semidefinita positiva (o negativa), o non definita: definita POSITIVA.

---

11. Esercizi di riepilogo

Svolgere in modo completo gli esercizi proposti.

Esercizio 11.1. (27 gennaio 2010)
Sia $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$ una matrice reale simmetrica di ordine 2 con autovalori 2 e 3, e tale che $A \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 2 \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

(1) Trovare una base di $\mathbb{R}^2$ formata da autovettori di A.
(2) Determinare la matrice A.

Risoluzione.
(• Punto 1) Osserviamo che il testo già ci dice che $\boldsymbol{v}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ è autovettore con autovalore 2.

Poichè la matrice ha due autovalori distinti, un autovettore relativo all'altro autovalore si può trovare immediatamente, perchè deve essere ortogonale al primo, e la sua direzione è determinata, poichè la dimensione dello spazio su cui agisce A è 2: possiamo scrivere

$$\boldsymbol{v}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Abbiamo quindi una base di $\mathbb{R}^2$ formata da autovettori:

$$\mathcal{B} = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\}.$$

(• Punto 2) Scriviamo le due equazioni agli autovalori per i due autovettori:

$$\begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix} \qquad \begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 3 \\ -3 \end{pmatrix}$$

che si traduce nel sistema per le a, b, c:

$$\begin{cases} a+b=2 \\ b+c=2 \\ a-b=3 \\ b-c=-3 \end{cases} .$$

Si verifica subito che il sistema ammette un'unica soluzione:

$$\begin{cases} a=\frac{5}{2} \\ b=-\frac{1}{2} \\ c=\frac{5}{2} \end{cases}$$

e, quindi la matrice richiesta è

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

Esercizio 11.2. (27 gennaio 2010)
Si consideri la seguente matrice A di tipo $(2, 3)$: $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}$.

(1) Trovare una matrice B di tipo $(3, 2)$ tale che $A.B = I_2$.
B è unica?
(2) Verificare che ogni matrice B che soddisfa la relazione $A.B = I_2$, verifica la proprietà:

$$\operatorname{Im} B \cap \operatorname{Ker} A = \{\mathbf{0}\}.$$

Risoluzione.
($\bullet$ Punto 1) Sia

$$B = \begin{pmatrix} a & d \\ b & e \\ c & f \end{pmatrix} \qquad a, b, c, d, e, f \in \mathbb{R} .$$

La condizione richiesta è

$$\begin{aligned} A.B &= \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} a & d \\ b & e \\ c & f \end{pmatrix} \\ &= \begin{pmatrix} a+b+c & d+e+f \\ a+c & d+f \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \end{aligned} \tag{9.54}$$

Otteniamo il sistema di equazioni

$$\begin{cases} a+b+c=1 \\ d+e+f=0 \\ a+c=0 \\ d+f=1 \end{cases}$$

equivalente a

$$\begin{cases} a+c=0 \\ b=1 \\ d+f=1 \\ e=-1 \end{cases} \tag{9.55}$$

Una possibile matrice B che soddisfa la richiesta si ottiene con $a = 1$ e $f = 0$, ed è, quindi

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

Non è unica: basta prendere, per esempio, $a = 0$, $f = 1$ per avere una matrice ugualmente valida, ma differente.
(• Punto 2) Anzitutto, osserviamo che le componenti dei vettori di $\operatorname{Ker} A$ obbedisono alle equazioni

$$\begin{cases} x+y+z=0 \\ x+z=0 \end{cases} \tag{9.56}$$

ossia

$$\begin{cases} x+z=0 \\ y=0 \end{cases} \quad ; \tag{9.57}$$

in altre parole, $\operatorname{Ker} A = \operatorname{Span}\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$.

Le matrici B che soddisfano la richiesta del testo sono una varietà lineare; infatti le soluzioni del sistema (9.55 hanno la forma

$$\begin{pmatrix} a \\ b \\ c \\ d \\ e \\ f \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ -1 \\ 1 \end{pmatrix} + a \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + d \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ -1 \end{pmatrix}$$

con $a, d \in \mathbb{R}$ liberi, o, tornando alle matrici

$$B = \begin{pmatrix} a & d \\ b & e \\ c & f \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 1 & -1 \\ 0 & 1 \end{pmatrix} + a \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ -1 & 0 \end{pmatrix} + d \begin{pmatrix} 0 & 1 \\ 0 & 0 \\ 0 & -1 \end{pmatrix}.$$

L'immagine di B è generata dalle colonne della matrice:

$$\operatorname{Im} B = \operatorname{Span}\left(\begin{pmatrix} a \\ b \\ c \end{pmatrix}, \begin{pmatrix} d \\ e \\ f \end{pmatrix}\right) = \operatorname{Span}\left(\left[\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + a \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}\right], \left[\begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} + d \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}\right]\right).$$

Se un vettore è in $\operatorname{Im} B \cap \operatorname{Ker} A$, devono esistere λ, μ, ν tali che

$$\lambda \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = \mu \left[\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + a \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \right] + \nu \left[\begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} + d \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \right],$$

ossia, riordinando

$$(\lambda - a\mu - d\nu) \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} - \mu \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} - \nu \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

I vettori

$$\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}$$

sono linearmente indipendenti; infatti, basta calcolare il determinante della matrice M le cui colonne sono i vettori

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

per convincersene. Quindi, l'unica combinazione lineare nulla è quella banale, e in particolare $\mu = \nu = 0$, da cui $\lambda = 0$. In altre parole, l'unico vettore nell'intersezione è il vettore nullo.

Esercizio 11.3. (27 gennaio 2010)

Si consideri la seguente matrice quadrata di ordine 2: $A = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$.

(1)) Trovare una matrice 2×2 non nulla B tale che $A \cdot B = \mathbf{0}_{2,2}$.
(2)) Mostrare che il sottoinsieme delle matrici B di ordine 2 tali che $A \cdot B = \mathbf{0}_{2,2}$, è un sottospazio di $\mathcal{M}_{\mathbb{R}}(2)$ e calcolarne la dimensione.

Risoluzione.

(• Punto 1) La soluzione è simile a quella della domanda precedente; una matrice che soddisfi le richieste è

$$\begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}.$$

(• Punto 2) Sia

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

una matrice tale che $A.B = \mathbf{0}_{2,2}$; questo si traduce nel sistema

$$\begin{cases} a - c = 0 \\ -a + c = 0 \\ b - d = 0 \\ -b + d = 0 \end{cases}.$$

È immediato convincersi che due equazioni sono superflue, e le altre indipendenti: quindi la matrice dei coefficienti ha rango 2 e la dimensione del nucleo dell'applicazione associata, ossia la dimensione del sottospazio vettoriale delle matrici B è $4-2=2$.

Esercizio 11.4. (27 gennaio 2010)
Sia $\mathbb{R}_n[x]$ lo spazio vettoriale dei polinomi in x a coefficienti reali con grado $\leq n$. Si consideri l'applicazione $L\colon \mathbb{R}_4[x] \to \mathbb{R}_5[x]$ data da:

$$L(p(x)) = (x-1)p(x).$$

(1)) Verificare che L è un'applicazione lineare.
(2)) Trovare una base per lo spazio $\operatorname{Im} L$.
(3)) Verificare che $\operatorname{Im} L$ è il sottoinsieme di $\mathbb{R}_5[x]$ dei polinomi che ammettono $x=1$ come radice.

Risoluzione.
(• Punto 1) Siano $p(x), q(x) \in \mathbb{R}_4[x]$; calcoliamo $L(p(x)+q(x))$:

$$L(p(x)+q(x)) = (x-1)(p(x)+q(x)) = (x-1)p(x)+(x-1)q(x) = L(p(x))+L(q(x))\,;$$

calcoliamo ora $L(\lambda p(x))$ con $\lambda \in \mathbb{R}$:

$$L(\lambda p(x)) = (x-1)\lambda p(x) = \lambda(x-1)p(x) = \lambda L(p(x))\,.$$

Quindi L è lineare.

L'immagine è generata dalle immagini dei vettori della base del dominio $\mathbb{R}_4[x]$; una base dell'anello dei polinomi di grado inferiore o uguale a 4 è:

$$\mathcal{B} = \left\{1, x, x^2, x^3, x^4\right\}\,.$$

(• Punto 2) Troviamo l'immagine dei vettori della base:

$$L(\mathcal{B}) = \left\{(x-1),\, x(x-1),\, x^2(x-1),\, x^3(x-1),\, x^4(x-1)\right\}\,.$$

Poichè sono tutti polinomi di grado differente, i vettori di $L(\mathcal{B})$ sono linearmente indipendenti, quindi costituiscono una base di $\operatorname{Im}(L)$, che ha dimensione 5.
(• Punto 3) Ogni vettore dell'immagine è dato da una combinazione lineare di vettori di $L(\mathcal{B})$; per come è costruita, questa combinazione lineare contiene il fattore comune $(x-1)$, quindi il polinomio ottenuto o è il polinomio nullo o ammette la radice $x=1$.

Sia ora $p(x)$ un polinomio di $\mathbb{R}_5[x]$ che ammette la radice $x=1$; allora, possiamo scrivere

$$p(x) = (x-1)q(x)\,,$$

dove $\deg(q(x)) \leq 4$, quindi esiste un polinomio $q(x) \in \mathbb{R}_4[x]$ la cui immagine è proprio $p(x)$.

Esercizio 11.5. (27 gennaio 2010)
Considerato lo spazio vettoriale reale $\mathcal{M}_{\mathbb{R}}(3)$ delle matrici quadrate reali di ordine 3, mostrare che il suo sottoinsieme delle matrici simmetriche a traccia nulla

$$T = \{A \in \mathcal{M}_{\mathbb{R}}(3) \mid \operatorname{tr}(A) = 0,\, A = A^{\mathrm{T}}\}$$

è un sottospazio vettoriale, e determinarne dimensione ed una base.

RISOLUZIONE.
Siano $A, B \in \mathcal{M}_{\mathbb{R}}(3)$; allora, $\operatorname{tr}(A+B) = \operatorname{tr}(A) + \operatorname{tr}(B) = 0 + 0 = 0$. Inoltre, per ogni $\lambda \in \mathbb{R}$, $\operatorname{tr}(\lambda A) = \lambda \operatorname{tr}(A) = \lambda \cdot 0 = 0$. Quindi T è chiuso rispetto alla somma fra vettori ed alla moltiplicazione per uno scalare, pertanto è un sottospazio vettoriale.

Una matrice generica di T si può scrivere come

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

con $a + d + f = 0$, ossia $f = -(a + d)$; così facendo, la matrice è simmetrica. Esistono, quindi, 5 parametri liberi: $\dim(T) = 5$. Una base di T è data da:

$$\left\{ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \right\}$$

ESERCIZIO 11.6. (27 gennaio 2010)
Sia $L : \mathbb{R}^3 \to \mathbb{R}^4$ l'applicazione lineare tale che

$$L\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}, \qquad L\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 4 \\ 3 \\ 2 \\ 1 \end{pmatrix}, \qquad L\begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

(1)) Calcolare $L\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$.

(2)) Determinare la matrice rappresentativa di L rispetto alle basi standard di $\mathbb{R}^3$ e $\mathbb{R}^4$.

RISOLUZIONE.
(• Punto 1) Notiamo che

$$\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} - \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

e che conosciamo le immagini degli ultimi due vettori. Per la linearità di L abbiamo

$$L\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = L\left(\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} - \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}\right) = L\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} - L\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 4 \\ 3 \\ 2 \\ 1 \end{pmatrix} - \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \\ -1 \\ -3 \end{pmatrix}.$$

(• Punto 2) Ci servono le immagini dei vettori della base standard del dominio. La definizione di L ci fornisce direttamente $L(\boldsymbol{e}_1)$, e abbiamo appena trovato

$L(\boldsymbol{e}_2)$: manca quella del terzo vettore $\boldsymbol{e}_3$. Osserviamo che la terza equazione di definizione di L può essere scritta come

$$L(-2\boldsymbol{e}_1 + \boldsymbol{e}_2 + \boldsymbol{e}_3) = -2L(\boldsymbol{e}_1) + L(\boldsymbol{e}_2) + L(\boldsymbol{e}_3) = \mathbf{0}_4,$$

da cui $L(\boldsymbol{e}_3) = 2L(\boldsymbol{e}_1) - L(\boldsymbol{e}_2)$. Passando alle rappresentazioni nella base standard di $\mathbb{R}^4$:

$$L\begin{pmatrix}0\\0\\1\end{pmatrix} = 2\begin{pmatrix}1\\2\\3\\4\end{pmatrix} - \begin{pmatrix}3\\1\\-1\\-3\end{pmatrix} = \begin{pmatrix}-1\\3\\7\\11\end{pmatrix}.$$

Le colonne della matrice rappresentativa richiesta sono le immagini della base del dominio nella base del codominio, quindi:

$$A_L = \begin{pmatrix}1 & 3 & -1\\ 2 & 1 & 3\\ 3 & -1 & 7\\ 4 & -3 & 11\end{pmatrix}.$$

Esercizio 11.7. (11 febbraio 2010)
In $\mathbb{R}^4$ si considerino i seguenti vettori dipendenti da un parametro $h \in \mathbb{R}$

$$\boldsymbol{v}_1 = \begin{pmatrix}h\\1\\h^2\\h\end{pmatrix} \qquad \boldsymbol{v}_2 = \begin{pmatrix}1\\1\\1\\1\end{pmatrix} \qquad \boldsymbol{v}_3 = \begin{pmatrix}h\\1\\h^2\\-h\end{pmatrix} \qquad \boldsymbol{v}_4 = \begin{pmatrix}1\\0\\1\\1\end{pmatrix}.$$

Determinare per quali valori del parametro $h \in \mathbb{R}$:

(1) i vettori $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3$ e $\boldsymbol{v}_4$ sono **linearmente indipendenti**:
(2) il vettore $\boldsymbol{v}_4$ appartiene allo $\mathrm{Span}(\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3)$:

Esercizio 11.8. (11 febbraio 2010)
Si consideri la seguente matrice reale:

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

(1) Verificare che il vettore $\boldsymbol{v} = \begin{pmatrix}1\\1\\1\\1\end{pmatrix}$ è un autovettore di A e calcolare l'autovalore corrispondente.
(2) Fissato in $\mathbb{R}^4$ il prodotto scalare standard, determinare una base ortogonale del sottospazio $U = \mathrm{Span}(\boldsymbol{v})^{\perp}$.
(3) Verificare che U è un autospazio di A e calcolare l'autovalore corrispondente.

Esercizio 11.9. (11 febbraio 2010)
Si consideri la seguente matrice $A = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$, e siano U e W gli insiemi così definiti:

$$U = \{M \in \mathcal{M}_{\mathbb{R}}(2) \mid \operatorname{tr}(A \cdot M) = 0\};$$
$$W = \{M \in \mathcal{M}_{\mathbb{R}}(2) \mid A \cdot M = \mathbf{0}\}.$$

(1) Dimostrare che U e W sono sottospazi vettoriali reali.
(2) Calcolare le dimensioni di U e W.
(3) Verificare che W é un sottospazio proprio di U.

Esercizio 11.10. (11 febbraio 2010)
Si considerino le seguenti matrici reali:

$$A = \begin{pmatrix} 2 & 0 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad B = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

(1) Verificare che le matrici A e B hanno lo stesso polinomio caratteristico.
(2) Le matrici sono simili? Giustificare la risposta.
(3) Trovare, se esiste, una matrice M ortogonale di ordine 3 per cui risulta $A = M^{\mathrm{T}}BM$.

Esercizio 11.11. (11 febbraio 2010)
Si considerino le seguenti matrici reali:

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

(1) Verificare che le matrici A e B hanno lo stesso polinomio caratteristico.
(2) Le matrici sono simili? Giustificare la risposta.
(3) Trovare, se esiste una matrice M invertibile di ordine 3 per cui $M^{-1}BM$ risulta diagonale. M può essere ortogonale?

Esercizio 11.12. (11 febbraio 2010)
In $\mathbb{R}^4$ si considerino i seguenti vettori dipendenti da un parametro $h \in \mathbb{R}$

$$\boldsymbol{v}_1 = \begin{pmatrix} h \\ 0 \\ h^2+2h \\ h \end{pmatrix} \qquad \boldsymbol{v}_2 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \qquad \boldsymbol{v}_3 = \begin{pmatrix} h \\ 0 \\ h^2+2h \\ -h-2 \end{pmatrix} \qquad \boldsymbol{v}_4 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \end{pmatrix}.$$

Determinare per quali valori del parametro $h \in \mathbb{R}$:

(1) lo spazio $\operatorname{Span}(\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \boldsymbol{v}_4)$ ha dimensione 2:
(2) lo spazio $\operatorname{Span}(\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \boldsymbol{v}_4)$ ha dimensione 3:

Esercizio 11.13. (11 febbraio 2010)
Considerare le matrici $A = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $C = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$. Siano U e W gli insiemi così definiti:

$$U = \{M \in \mathcal{M}_{\mathbb{R}}(2) \mid \operatorname{tr}(AM) = 0\};$$

$$W = \{M \in \mathcal{M}_{\mathbb{R}}(2) \mid \operatorname{tr}(AM) = \operatorname{tr}(BM) = \operatorname{tr}(CM) = 0\}.$$

(1) Dimostrare che U e W sono sottospazi vettoriali reali.
(2) Calcolare le dimensioni di U e W.

Esercizio 11.14. (11 febbraio 2010)
In $\mathbb{R}^4$ si considerino i seguenti vettori dipendenti da un parametro $h \in \mathbb{R}$

$$\boldsymbol{v}_1 = \begin{pmatrix} h \\ 0 \\ h^2+2h \\ h \end{pmatrix} \qquad \boldsymbol{v}_2 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \qquad \boldsymbol{v}_3 = \begin{pmatrix} h \\ 0 \\ h^2+2h \\ -h-2 \end{pmatrix} \qquad \boldsymbol{v}_4 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \end{pmatrix}.$$

Determinare per quali valori del parametro $h \in \mathbb{R}$:

(1) lo spazio $\operatorname{Span}(\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \boldsymbol{v}_4)$ ha dimensione 2:
(2) lo spazio $\operatorname{Span}(\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \boldsymbol{v}_4)$ ha dimensione 3:

Esercizio 11.15. (11 febbraio 2010)
Considerare le matrici $A = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $C = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$. Siano U e W gli insiemi così definiti:

$$U = \{M \in \mathcal{M}_{\mathbb{R}}(2) \mid \operatorname{tr}(AM) = 0\};$$

$$W = \{M \in \mathcal{M}_{\mathbb{R}}(2) \mid \operatorname{tr}(AM) = \operatorname{tr}(BM) = \operatorname{tr}(CM) = 0\}.$$

(1) Dimostrare che U e W sono sottospazi vettoriali reali.
(2) Calcolare le dimensioni di U e W.

Esercizio 11.16. (9 aprile 2010)
Si consideri la seguente matrice reale quadrata di ordine 3, in funzione del parametro reale h:

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

(1)) Determinare il rango di A al variare del parametro h.
(2)) Determinare gli autovalori della matrice A e le loro molteplicità algebriche PER $h = 0$ e PER $h = 1$.
(3)) POSTO $h = 1$, determinare dimensione, equazione cartesiana e base per ciascun autospazio di A:
(4)) Stabilire se la matrice sia diagonalizzabile per $h = 1$ e per $h = 0$.

Risoluzione.

(• Punto 1) Per $h = 0, 1$ si ha $\mathrm{rg}(A) = 2$; altrimenti, $\mathrm{rg}(A) = 3$.

(• Punto 2) Per $h = 0$ vi sono 3 autovalori reali distinti, ciascuno, quindi, con molteplicità algebrica unitaria:

$$\lambda_1 = 0,\ \lambda_2 = 2 - \sqrt{3},\ \lambda_3 = 2 + \sqrt{3}.$$

Per $h = 1$:

$$\begin{cases} \lambda_1 = 0 & \text{con molteplicità algebrica } \mu_1 = 2 \\ \lambda_2 = 5 & \text{con molteplicità algebrica } \mu_2 = 1 \end{cases}$$

(• Punto 3)

$$\begin{cases} \dim(V_0) = 1;\ V_0\colon x + z = y = 0;\ \mathcal{B}_{V_0} = \left\{ \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \right\} \\ \dim(V_5) = 1;\ V_5\colon 4x - y - z = x - 2y + z = 0;\ \mathcal{B}_{V_5} = \left\{ \begin{pmatrix} 3 \\ 5 \\ 7 \end{pmatrix} \right\} \end{cases}$$

(• Punto 4) Per $h = 0$ SÌ: la matrice è quadrata di ordine 3 ed ha 3 autovalori reali distinti; per $h = 1$ NO: $\dim(V_0) = 1 \neq \mu_1 = 2$ (la molteplicità algebrica differisce da quella geometrica).

Esercizio 11.17. (28 giugno 2010)

Determinare quali tra le seguenti applicazioni $\mathbb{R}^3 \to \mathbb{R}^3$ sono applicazioni lineari, giustificando ogni affermazione. In caso affermativo, determinare una rappresentazione matriciale nella base canonica.

(1)) $L_1 \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x - y \\ 2z \\ xy \end{pmatrix}$.

(2)) $L_2 \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$.

(3)) $L_3 \begin{pmatrix} x \\ y \\ z \end{pmatrix} = (x + y + z) \cdot \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$.

Risoluzione.

(• Punto 1) NO. Infatti

$$L_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + L_1 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \neq L_1 \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

(• Punto 2) NO. Infatti

$$L_2 \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

(• Punto 3) SÌ: per ogni $\alpha, \beta \in \mathbb{R}$ e $\boldsymbol{u}_1, \boldsymbol{u}_2 \in \mathbb{R}^3$

$$\begin{aligned}
L_3(\alpha \boldsymbol{u}_1 + \beta \boldsymbol{u}_2) &= L_3 \begin{pmatrix} \alpha x_1 + \beta x_2 \\ \alpha y_1 + \beta y_2 \\ \alpha z_1 + \beta z_2 \end{pmatrix} \\
&= (\alpha x_1 + \beta x_2 + \alpha y_1 + \beta y_2 + \alpha z_1 + \beta z_2) \cdot \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \\
&= (\alpha(x_1 + y_1 + z_1) + \beta(x_2 + y_2 + z_2)) \cdot \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \alpha L_3(\boldsymbol{u}_1) + \beta L_3(\boldsymbol{u}_2)
\end{aligned}$$

.

$$A_{L_3} = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \\ 3 & 3 & 3 \end{pmatrix}.$$

Esercizio 11.18. (28 giugno 2010)
Determinare quali tra i seguenti sottoinsiemi di $\mathbb{R}^3$ sono sottospazi, giustificando ogni affermazione. In caso affermativo, determinare una base e dimensione del sottospazio.

(1)) $U_1 = \{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 : \ x - y = 2z = 1 \ \}.$

(2)) $U_2 = \{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 : \ (1\ 2\ 3) \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0 \ \}.$

(3)) $U_3 = \{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 : \ xy = z = 0 \ \}.$

Risoluzione.

(• Punto 1) NO. Infatti $\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \notin U_1 \quad (2z = 0 \neq 1).$

(• Punto 1) SÌ: per ogni $\alpha, \beta \in \mathbb{R}$ e $\boldsymbol{u}_1, \boldsymbol{u}_2 \in U_2$

$$\begin{aligned}
&\begin{pmatrix} 1 & 2 & 3 \end{pmatrix} \cdot (\alpha \boldsymbol{u}_1 + \beta \boldsymbol{u}_2) \\
&= \alpha \begin{pmatrix} 1 & 2 & 3 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} + \beta \begin{pmatrix} 1 & 2 & 3 \end{pmatrix} \cdot \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix} = \alpha\, 0 + \beta\, 0 = 0,
\end{aligned}$$

quindi $\alpha \boldsymbol{u}_1 + \beta \boldsymbol{u}_2 \in U_2$.

$$U_2 = \{x + 2y + 3z = 0\}\text{. Da cui } \dim(U_2) = 2; \quad \mathcal{B}_0 = \left\{ \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 0 \\ -1 \end{pmatrix} \right\}.$$

(• Punto 3) NO. Infatti

$$\begin{pmatrix}1\\0\\0\end{pmatrix} \in U_3, \qquad \begin{pmatrix}0\\1\\0\end{pmatrix} \in U_3$$

ma

$$\begin{pmatrix}1\\0\\0\end{pmatrix} + \begin{pmatrix}0\\1\\0\end{pmatrix} = \begin{pmatrix}1\\1\\0\end{pmatrix} \notin U_3 \quad (xy = 1 \neq 0)$$

Esercizio 11.19. (8 luglio 2010)
Considerare le matrici $A = \begin{pmatrix}1 & -1\\-1 & 2\end{pmatrix}$, $B = \begin{pmatrix}1 & 0\\0 & 0\end{pmatrix}$, $C = \begin{pmatrix}0 & 0\\0 & 1\end{pmatrix}$.

Siano ora U, V e W gli insiemi così definiti:

$$U = \{M \in \mathcal{M}_{\mathbb{R}}(2) \mid \operatorname{tr}(AM) = 0\};$$

$$V = \{M \in \mathcal{M}_{\mathbb{R}}(2) \mid BM = C\};$$

$$W = \{M \in \mathcal{M}_{\mathbb{R}}(2) \mid \operatorname{tr}(AM) = \operatorname{tr}(BM) = \operatorname{tr}(CM) = 0\}.$$

(1) Determinare se U, V e W sono sottospazi vettoriali reali.
(2) Per gli insiemi per i quali la risposta è affermativa, calcolarne le dimensioni.

Esercizio 11.20. (14 settembre 2010)
Posto

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

(1) Si verifichi che le colonne di A formano una base ortogonale di $\mathbb{R}^3$.
(2) Si verifichi che per ogni $\boldsymbol{v} \in \mathbb{R}^3$ si ha $||A\boldsymbol{v}||^2 = 9||\boldsymbol{v}||^2$.
(3) Si deduca che se λ è autovalore di A allora $\lambda = 3$ oppure $\lambda = -3$.
(4) Si determinino gli autovalori di A.

Risoluzione.
(• Punto 1) $\det(A) = 27 \neq 0$, quindi i 3 vettori colonna sono linearmente indipendenti, e formano una base di $\mathbb{R}^3$. Basta ora verificare che sono ortogonali; siano

$$\boldsymbol{v}_1 = A^1 = \begin{pmatrix}1\\-2\\-2\end{pmatrix}, \quad \boldsymbol{v}_2 = A^2 = \begin{pmatrix}2\\2\\-1\end{pmatrix}, \quad \boldsymbol{v}_3 = A^3 = \begin{pmatrix}2\\-1\\2\end{pmatrix}.$$

Si ha

$$\begin{aligned}
\langle \boldsymbol{v}_1, \boldsymbol{v}_2\rangle &= 1\cdot 2 + (-2)\cdot 2 + (-2)\cdot(-1) = 2 - 4 + 2 = 0\\
\langle \boldsymbol{v}_1, \boldsymbol{v}_3\rangle &= 1\cdot 2 + (-2)\cdot) - 1) + (-2)\cdot 2 = 2 + 2 - 4 = 0\\
\langle \boldsymbol{v}_2, \boldsymbol{v}_3\rangle &= 2\cdot 2 + (-1)\cdot 2 + (-1)\cdot 2 = 4 - 2 - 2 = 0\,,
\end{aligned}$$

quindi i vettori sono ortogonali.

(• Punto 2) Sia

$$\boldsymbol{v} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

Abbiamo

$$A\boldsymbol{v} = \begin{pmatrix} 1 & 2 & 2 \\ -2 & 2 & -1 \\ -2 & -1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x+2y+2z \\ -2x+2y-z \\ -2x-y+2z \end{pmatrix}.$$

Ora,

(9.58)

$$\begin{aligned} \|A\boldsymbol{v}\|^2 &= 9x^2+9y^2+9z^2 = (x+2y+2z)^2 + (-2x+2y-z)^2 + (-2x-y+2z)^2 \\ &= 9(x^2+y^2+z^2) = 9\|\boldsymbol{v}\|^2 \end{aligned}$$

per ogni scelta di $x, y, z \in \mathbb{R}$, da cui l'asserto.

(• Punto 3) Sia λ autovalore di A e $\boldsymbol{v} \neq \mathbf{0}$ un autovettore associato a λ. Allora

$$\|A\boldsymbol{v}\|^2 = \|\lambda\boldsymbol{v}\|^2 = \lambda^2\|\boldsymbol{v}\|^2 = 9\|\boldsymbol{v}\|^2\,,$$

in base al punto precedente. Quindi, essendo $\boldsymbol{v}$ di modulo non nullo (è autovettore)

$$\lambda^2 = 9\,,$$

ossia $\lambda = \pm 3$.

(• Punto 4) La condizione trovata sopra per λ è necessaria, ma non sufficiente; gli autovalori possono essere solo 3 o -3, ma non è detto che lo siano. Si verifica immediatamente (basta svolgere il calcolo) che, con $\lambda = 3$, $\det(A-3I) = 0$, quindi è autovalore di A. Invece, $\det(A+3I) \neq 0$, quindi $\lambda = -3$ non è autovalore di A.

Esercizio 11.21. (14 settembre 2010)
Considerare l'applicazione $L\colon \mathbb{R}^3 \to \mathbb{R}$ così definita:

$$L\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = \det \begin{pmatrix} 1 & 0 & x \\ 2 & 1 & y \\ 1 & -1 & z \end{pmatrix}$$

(1) Dimostrare che L è un'applicazione lineare.
(2) Scrivere la matrice rappresentativa di L nelle basi standard.
(3) Determinare la/le equazioni cartesiane ed una base per $\operatorname{Ker} L$.
(4) Verificare che L è suriettiva.

Risoluzione.
(• Punto 1) Discende direttamente da una delle proprietà del determinante: siano

$$\boldsymbol{u} = \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} \qquad \text{e} \qquad \boldsymbol{v} = \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix}.$$

Per ogni $\alpha, \beta \in \mathbb{R}$ e $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^3$, vale

$$\begin{aligned} L(\alpha\boldsymbol{u} + \beta\boldsymbol{v}) &= \det \begin{pmatrix} 1 & 0 & \alpha x_1 + \beta x_2 \\ 2 & 1 & \alpha y_1 + \beta y_2 \\ 1 & -1 & \alpha z_1 + \beta z_2 \end{pmatrix} \\ &= \alpha \det \begin{pmatrix} 1 & 0 & x_1 \\ 2 & 1 & y_1 \\ 1 & -1 & z_1 \end{pmatrix} + \beta \det \begin{pmatrix} 1 & 0 & x_2 \\ 2 & 1 & y_2 \\ 1 & -1 & z_2 \end{pmatrix} \\ &= \alpha L(\boldsymbol{u}) + \beta L(\boldsymbol{v})\,. \end{aligned}$$

(• Punto 2) Calcolando esplicitamente l'espressione di L si ottiene, sviluppando il determinante lungo l'ultima colonna:

$$L(\begin{pmatrix} x \\ y \\ z \end{pmatrix}) = -3x + y + z = \begin{pmatrix} -3 & 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix},$$

per cui la matrice associata ad L è:

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

In effetti, le colonne della matrice A_L devono essere formate dai trasformati secondo la L dei vettori della base, rappresentati nelle basi canoniche: si verifica immediatamente il risultato calcolando

$$L(\boldsymbol{e}_1) = (\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}) = \det \begin{pmatrix} 2 & 1 \\ 1 & -1 \end{pmatrix} = -3$$

e così via.

(• Punto 3) Il nucleo della applicazione corrisponde ai vettori

$$\operatorname{Ker} L = \left\{ \boldsymbol{v} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \mid L(\boldsymbol{v}) = 0 \right\},$$

ossia ai vettori per i quali

$$-3x + y + z = 0,$$

che è l'equazione cartesiana del nucleo. Una sua base $\mathcal{B}$ si ottiene riscrivendo l'equazione in forma parametrica

$$\begin{cases} x = \lambda \\ y = \mu \\ z = 3\lambda - \mu \end{cases} \qquad \forall\, \lambda, \mu \in \mathbb{R}$$

e calcolando i vettori che si ottegono con due scelte indipendenti dei valori dei parametri. Questo si ottiene facilmente ponendo tutti i parametri uguali a 0 tranne uno, che viene uguagliato a 1, per tutti i due possibili. Otteniamo

$$\mathcal{B} = \left\{ \begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \right\}$$

(• Punto 4) Dobbiamo verificare che, per ogni vettore nel codominio, ossia, per ogni numero reale δ, esiste un vettore la cui immagine coincide con tale vettore, ossia esiste una scelta di valori x, y, z tali che $-3x + y + z = \delta$. Ciò è vero: basta risolvere in z e otteniamo

$$z = 3x - y + \delta\,.$$

Qualunque sia il valore di delta, e assegnato un valore arbitrario a x ed y, questo consente di calcolare sempre il valore di z richiesto, per soddisfare la richiesta.

Esercizio 11.22. (22 novembre 2010)
Si consideri il seguente sistema lineare:

$$\begin{cases} 2x - 2y - z - t - w = -1 \\ z + 2t + w = 2 \\ -6x + z + t + 2w = 1 \end{cases}$$

(1) Scrivere la matrice A dei coefficienti del sistema.
(2) Discutere la risolvibilità del sistema.
(3) Sia $L_A \colon \mathbb{R}^5 \to \mathbb{R}^3$ l'applicazione lineare che associa ad ogni vettore $X \in \mathbb{R}^5$ il vettore AX. Trovare una base del nucleo $\operatorname{Ker} L_A$.
(4) Risolvere il sistema.

Risoluzione.
(• Punto 1)

$$A = \begin{pmatrix} 2 & -2 & -1 & -1 & -1 \\ 0 & 0 & 1 & 2 & 1 \\ -6 & 0 & 1 & 1 & 2 \end{pmatrix}$$

(• Punto 2) È risolvibile: $\operatorname{rg}(A) = 3$, quindi $\operatorname{rg}(A|B) = 3$, dove B è il vettore colonna dei termini noti.
(• Punto 3)

$$\mathcal{B}_{\operatorname{Ker} L_A} = \left\{ \begin{pmatrix} 1 \\ 1 \\ -6 \\ 0 \\ 6 \end{pmatrix}, \begin{pmatrix} -1 \\ 2 \\ -12 \\ 6 \\ 0 \end{pmatrix} \right\}$$

(• Punto 4)

$$\begin{cases} x = \lambda \\ y = \mu \\ z = -6\mu \\ t = 1 - 2\lambda + 2\mu \\ w = 4\lambda + 2\mu \end{cases} \qquad \lambda, \mu \in \mathbb{R}$$
