

Capitolo 1

INSIEMI, APPLICAZIONI, RELAZIONI E PERMUTAZIONI

Gli argomenti presentati nei primi due paragrafi sono in larga parte già noti allo studente, dai corsi del primo anno.

1. Richiami sugli insiemi

Il concetto di *insieme* è primitivo. Dire "un insieme è una collezione di oggetti" è una tautologia [infatti cos'è una *collezione*, se non un *insieme* di oggetti?].

Diremo che un insieme A è assegnato quando è possibile stabilire se un oggetto x è *elemento* di A [e si scrive $x \in A$] o non è elemento di A [e si scrive $x \notin A$].

Esiste un solo insieme privo di elementi: è l'*insieme vuoto*, denotato \emptyset . Assumiamo inoltre elementarmente noti gli insiemi numerici più importanti: \mathbf{N} , \mathbf{Z} , \mathbf{Q} , \mathbf{R} , \mathbf{C} (cioè l'insieme dei *naturali*, degli *interi*, dei *razionali*, dei *reali* e dei *complessi*).

Per descrivere un insieme A :

- se ne possono scrivere gli elementi, elencandoli tra parentesi graffe e separandoli con virgolette: ad esempio $A = \{a, b, c, \dots\}$.
- si può descrivere (sempre tra parentesi graffe) la legge di appartenenza dei suoi elementi. A tale scopo si fa uso di alcuni ben noti *simboli logici*, cioè:
 - *quantificatori*: $\forall, \exists, \exists!, \nexists$ [risp. *per ogni, esiste, esiste un unico, non esiste*];
 - *implicazioni*: $\implies, \iff, \iff, \not\implies$ [risp. *implica, è implicato, equivale, non implica*];
 - *congiunzioni*: $, , \wedge, \vee, \neg, |$ [risp. *virgola (o separatore), e, oppure, non, tale che*].

La congiunzione "*non*" è talvolta indicata con $/$ (invece di \neg), mentre "*tale che*" è spesso indicata con $:$ (invece di $|$).

Ad esempio, l'insieme \mathbf{P} dei numeri naturali pari può essere descritto in questi modi:

$$\mathbf{P} = \{0, 2, 4, 6, 8, \dots\}, \text{ oppure } \mathbf{P} = \{n \in \mathbf{N} \mid n = 2x, \exists x \in \mathbf{N}\}, \text{ oppure } \mathbf{P} = \{2n, \forall n \in \mathbf{N}\}.$$

Definizione 1. Siano A, B due insiemi. A è detto *sottoinsieme* di B se $a \in B, \forall a \in A$. Si scrive in tal caso $A \subseteq B$ [oppure $B \supseteq A$] e si dice che A è *contenuto in* B [o che B *contiene* A].

Si dice poi che A è *contenuto propriamente* in B o che B *contiene propriamente* A [e si scrive $A \subset B$ o $B \supset A$] se $A \subseteq B$ ed $\exists b \in B \mid b \notin A$.

Ogni insieme A ammette sempre i due sottoinsiemi \emptyset, A , detti *sottoinsiemi banali* di A . Gli altri (eventuali) sottoinsiemi di A sono detti *sottoinsiemi propri*.

Definizione 2. Due insiemi A, B sono detti *uguali* [e si scrive $A = B$] se hanno gli stessi elementi. Si ha quindi:

$$A = B \iff A \subseteq B \text{ e } B \subseteq A.$$

A, B sono detti *diversi* se non sono uguali [e si scrive $A \neq B$].

Osservazione 1. (i) Dalla definizione precedente segue, ad esempio, che $\{a, a\} = \{a\}$.

(ii) Sia $n \in \mathbf{N}$. Se un insieme A è formato da n elementi (a due a due distinti), si scrive $|A| = n$, [oppure $\#(A) = n$] e si dice che A ha *cardinalità* n . In particolare, $|\emptyset| = 0$; viceversa, se $|A| = 0$, allora $A = \emptyset$.

Un insieme A è detto *finito* se ha cardinalità n (per qualche $n \in \mathbf{N}$); altrimenti è detto *infinito*. Ad esempio \mathbf{N} è un insieme infinito. Torneremo sul concetto di cardinalità nel paragrafo 3, per darne una definizione meno intuitiva.

(iii) Si noti che abbiamo utilizzato il simbolo $=$ non solo nel senso della **Def. 2** ma anche per definire un insieme. In effetti, scrivendo $\mathbf{P} = \{0, 2, 4, \dots\}$ abbiamo assegnato il nome \mathbf{P} all'insieme $\{0, 2, 4, \dots\}$. In tal caso è più corretto sostituire $=$ con $:=$, scrivendo quindi $\mathbf{P} := \{0, 2, 4, \dots\}$.

Analogamente, abbiamo usato la doppia implicazione \iff anche per definire un concetto. Ma in tal caso è più corretto sostituirla con \iff ovvero con $\stackrel{\text{def}}{\iff}$. Ad esempio, avremmo dovuto scrivere (trattandosi di una definizione) $A = B \stackrel{\text{def}}{\iff} A \subseteq B$ e $B \subseteq A$.

Definizione 3. Sia X un insieme e siano A, B sottoinsiemi di X . Sono definiti i seguenti insiemi:

$$\begin{aligned} A \cap B &:= \{x \in X : x \in A \text{ e } x \in B\}, \text{ detto intersezione di } A \text{ e } B; \\ A \cup B &:= \{x \in X : x \in A \text{ oppure } x \in B\}, \text{ detto unione di } A \text{ e } B; \\ A - B &:= \{x \in X : x \in A \text{ e } x \notin B\}, \text{ detto differenza di } A \text{ con } B; \\ \mathbf{C}_x(A) &:= X - A = \{x \in X : x \notin A\}, \text{ detto complementare di } A \text{ in } X. \end{aligned}$$

Osservazione 2. (i) Valgono le seguenti ovvie inclusioni:

$$A \cap B \subseteq A \subseteq A \cup B, \quad A \cap B \subseteq B \subseteq A \cup B, \quad A - B \subseteq A.$$

Inoltre $A - B = A \cap \mathbf{C}_x(B)$ e si verifica subito che $A - B_2 \subseteq A - B_1$, se $B_1 \subseteq B_2 \subseteq A$. Si noti infine che $A - B = A \iff A \cap B = \emptyset$.

(ii) L'intersezione e l'unione si generalizzano in questo modo: se $\{A_i\}_{i \in I}$ è una famiglia di sottoinsiemi di X ,

$$\bigcap_{i \in I} A_i := \{x \in X : x \in A_i, \forall i \in I\}, \quad \bigcup_{i \in I} A_i := \{x \in X : x \in A_i, \exists i \in I\}$$

(iii) Valgono le seguenti uguaglianze tra sottoinsiemi A, B, C di un insieme X [note come *formule di De Morgan*], la cui verifica è lasciata per esercizio:

$$\begin{aligned} (A \cap B) \cup C &= (A \cup C) \cap (B \cup C); & (A \cup B) \cap C &= (A \cap C) \cup (B \cap C); \\ A - (B \cup C) &= (A - B) \cap (A - C); & A - (B \cap C) &= (A - B) \cup (A - C). \end{aligned}$$

Definizione 4. Due insiemi A, B sono detti *disgiunti* se $A \cap B = \emptyset$.

Vale il seguente risultato, noto come il *Principio della somma*, che per il momento accetteremo per vero [in quanto non abbiamo ancora una soddisfacente nozione di cardinalità] e che dimostreremo invece nel successivo paragrafo 3.

Proposizione 1 (*Principio della somma*). Se A, B sono insiemi finiti disgiunti, risulta:

$$|A \cup B| = |A| + |B|.$$

Il Principio della somma si generalizza facilmente (procedendo per *induzione*), dal caso di due insiemi al caso di $k \geq 2$ insiemi.

Proposizione 1' (*Principio generalizzato della somma*). Se A_1, A_2, \dots, A_k sono insiemi finiti a due a due disgiunti, risulta:

$$\left| \bigcup_{i=1}^k A_i \right| = |A_1| + |A_2| + \dots + |A_k|.$$

Prima di dimostrare la **Prop. 1'**, richiamiamo brevemente il *principio di induzione*.

Principio di induzione. Sia $k_0 \in \mathbf{N}$ e sia $\mathcal{P} = \mathcal{P}(k)$ un'affermazione da dimostrare, dipendente dal naturale k , $\forall k \in \mathbf{N}$, $k \geq k_0$. Se valgono le due seguenti condizioni:

- (i) $\mathcal{P}(k_0)$ è vera,
- (ii) $\mathcal{P}(n)$ vera $\implies \mathcal{P}(n+1)$ vera $\quad \forall n \geq k_0$,

allora $\mathcal{P}(k)$ è vera, $\forall k \geq k_0$.

La (i) è detta "base induttiva", la (ii) è detta "passo induttivo", l'ipotesi " $\mathcal{P}(n)$ vera" [in (ii)] è detta "ipotesi induttiva".

Si noti che il passo induttivo poteva anche essere formulato in questo modo:

$$(ii) \quad \mathcal{P}(n-1) \text{ vera} \implies \mathcal{P}(n) \text{ vera} \quad \forall n > k_0.$$

Dim. (Prop. 1'). Tale proposizione è definita per tutti i naturali $k \geq 2$ [e dunque sia $k_0 = 2$]. La base induttiva è esattamente la **Prop. 1** [che abbiamo assunto di aver già dimostrato]. Dimostriamo quindi il passo induttivo: sia $n \geq 2$ e consideriamo gli insiemi a due a due disgiunti A_1, A_2, \dots, A_{n+1} . Per ipotesi induttiva:

$$|A_1 \cup A_2 \cup \dots \cup A_n| = |A_1| + \dots + |A_n|.$$

I due insiemi $A_1 \cup A_2 \cup \dots \cup A_n$ e A_{n+1} sono disgiunti; dalla **Prop. 1** segue che

$$|(A_1 \cup A_2 \cup \dots \cup A_n) \cup A_{n+1}| = (|A_1| + \dots + |A_n|) + |A_{n+1}|,$$

cioè

$$|A_1 \cup A_2 \cup \dots \cup A_n \cup A_{n+1}| = |A_1| + \dots + |A_n| + |A_{n+1}|.$$

Dunque il passo induttivo è provato ed il principio d'induzione ci consente di affermare che la proposizione è vera.

Dal principio della somma segue un importante risultato, utile nelle tecniche di conteggio: il *Principio di inclusione - esclusione*, che enunceremo e dimostreremo nel caso di tre insiemi finiti, ma che può essere facilmente generalizzato al caso di un numero finito di insiemi (finiti).

Proposizione 2 (*Principio di inclusione - esclusione*). *Siano A, B, C tre insiemi finiti. Risulta:*

$$|A \cup B \cup C| = |A| + |B| + |C| - (|A \cap B| + |A \cap C| + |B \cap C|) + |A \cap B \cap C|.$$

Dim. Dimostriamo dapprima la seguente formula (che è il principio di inclusione - esclusione relativo a due soli insiemi): se A, B sono due insiemi finiti, si ha:

$$(*) \quad |A \cup B| = |A| + |B| - |A \cap B|.$$

Partiamo dalle due seguenti ovvie uguaglianze insiemistiche:

$$A \cup B = (A - B) \cup B, \quad A = (A - B) \cup (A \cap B).$$

Poiché $(A - B) \cap B = \emptyset$ e $(A - B) \cap (A \cap B) = \emptyset$, possiamo applicare il principio della somma:

$$|A \cup B| = |A - B| + |B|, \quad |A| = |A - B| + |A \cap B|.$$

Sostituendo la seconda uguaglianza nella prima si ottiene:

$$|A \cup B| = -|A \cap B| + |A| + |B|$$

e l'affermazione (*) è così provata.

Siano ora A, B, C tre insiemi finiti arbitrari. Applichiamo (*) prima agli insiemi $A \cup B, C$ e poi agli insiemi A, B . Si ha:

$$(**) \quad |A \cup B \cup C| = |(A \cup B) \cup C| = |A \cup B| + |C| - |(A \cup B) \cap C| = |A| + |B| - |A \cap B| + |C| - |(A \cup B) \cap C|.$$

Esaminiamo l'ultimo addendo di tale sommatoria. In base alle formule di De Morgan,

$$(A \cup B) \cap C = (A \cap C) \cup (B \cap C).$$

Utilizzando nuovamente (*):

$$|(A \cup B) \cap C| = |A \cap C| + |B \cap C| - |(A \cap C) \cap (B \cap C)| = |A \cap C| + |B \cap C| - |(A \cap B \cap C)|.$$

Sostituendo tale uguaglianza in (**) si ottiene:

$$|A \cup B \cup C| = |A| + |B| - |A \cap B| + |C| - |A \cap C| - |B \cap C| + |(A \cap B \cap C)|,$$

cioè la formula cercata.

N.B. Se invece di tre insiemi ne abbiamo ad esempio quattro, la cardinalità della loro unione è ottenuta sommando le cardinalità dei quattro sottinsiemi, sottraendo le cardinalità delle intersezioni di due di essi, sommando poi le cardinalità delle intersezioni di tre di essi e sottraendo infine la

cardinalità dell'intersezione dei quattro insiemi [complessivamente 15 addendi]. Dovrebbe ora essere chiaro come il principio di inclusione - esclusione si generalizzi al caso di un numero finito di insiemi.

Definizione 5. Sia X un insieme. L'insieme dei sottoinsiemi di X è detto *insieme delle parti di X* . È denotato $\mathcal{P}(X)$.

Esempi 1. Se ad esempio $A = \{1\}$, $\mathcal{P}(A) = \{\emptyset, A\}$. Se $B = \{1, 2\}$, $\mathcal{P}(B) = \{\emptyset, \{1\}, \{2\}, B\}$. Se infine $C = \{1, 2, 3\}$, $\mathcal{P}(C) = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, C\}$.

Si noti poi che $\mathcal{P}(\emptyset) = \{\emptyset\}$ ha cardinalità 1; $\mathcal{P}(\mathcal{P}(\emptyset)) = \{\emptyset, \{\emptyset\}\}$ ha cardinalità 2; $\mathcal{P}(\mathcal{P}(\mathcal{P}(\emptyset))) = \{\emptyset, \{\emptyset\}, \{\{\emptyset\}\}, \{\emptyset, \{\emptyset\}\}\}$ ha cardinalità 4, ecc..

Nei precedenti esempi si può osservare che, se $|A| = n$, allora $|\mathcal{P}(A)| = 2^n$. Proviamo tale fatto.

Proposizione 3. Se A è un insieme finito di cardinalità n , $|\mathcal{P}(A)| = 2^n$.

Dim. Si procede per induzione su $n \geq 0$.

Base induttiva. Se $|A| = 0$, $|\mathcal{P}(A)| = 2^0$ [infatti se $|A| = 0$, allora $A = \emptyset$ e $|\mathcal{P}(\emptyset)| = |\{\emptyset\}| = 1 = 2^0$].

Passo induttivo. Sia $n \geq 0$ e, per ogni insieme B tale che $|B| = n$, sia $|\mathcal{P}(B)| = 2^n$. È da provare che, per ogni insieme A tale che $|A| = n + 1$, risulta: $|\mathcal{P}(A)| = 2^{n+1}$.

Si fissi un elemento $a_1 \in A$. In $\mathcal{P}(A)$ consideriamo i due sottoinsiemi:

$$\mathfrak{C} = \{B \in \mathcal{P}(A) : B \ni a_1\}, \quad \mathfrak{D} = \{B \in \mathcal{P}(A) : B \not\ni a_1\}.$$

Ovviamente $\mathfrak{C} \cup \mathfrak{D} = \mathcal{P}(A)$ e $\mathfrak{C} \cap \mathfrak{D} = \emptyset$. Dunque, per il principio della somma, $|\mathcal{P}(A)| = |\mathfrak{C}| + |\mathfrak{D}|$.

Ovviamente $\mathfrak{D} = \{B \subseteq A - \{a_1\}\} = \mathcal{P}(A - \{a_1\})$. Inoltre, per ogni $B \in \mathfrak{C}$, risulta: $B = B_1 \cup \{a_1\}$, con $B_1 \in \mathcal{P}(A - \{a_1\})$. Gli insiemi B sono ovviamente quanti gli insiemi B_1 . Ne segue che \mathfrak{C} ha la stessa cardinalità di $\mathcal{P}(A - \{a_1\})$.

Per ipotesi induttiva, $|\mathcal{P}(A - \{a_1\})| = 2^n$. Dunque $|\mathcal{P}(A)| = |\mathfrak{C}| + |\mathfrak{D}| = 2^n + 2^n = 2^{n+1}$.

Introduciamo ora la nozione di *partizione* di un insieme, che avrà molta importanza nello studio delle relazioni di equivalenza [che vedremo nel paragrafo 3].

Definizione 6. Sia X un insieme non vuoto e sia $\mathfrak{U} = \{A_i\}_{i \in I}$ una famiglia di suoi sottoinsiemi non vuoti. \mathfrak{U} è detta *ricoprimento di X* se $\bigcup_{i \in I} A_i = X$. \mathfrak{U} è detta *partizione di X* se è un ricoprimento di X e se $A_i \cap A_j = \emptyset$, se $i \neq j$ (cioè se i sottoinsiemi A_i sono a due a due disgiunti).

Esempi 2. (i) In \mathbf{N} consideriamo i tre sottoinsiemi

$$\mathbf{P} = 2\mathbf{N} := \{2n, \forall n \in \mathbf{N}\}, \quad 3\mathbf{N} := \{3n, \forall n \in \mathbf{N}\}, \quad 1 + 2\mathbf{N} = \{1 + 2n, \forall n \in \mathbf{N}\}.$$

$\{2\mathbf{N}, 3\mathbf{N}\}$ non è un ricoprimento di \mathbf{N} [infatti ad esempio $5 \notin 2\mathbf{N} \cup 3\mathbf{N}$]. Invece $\{2\mathbf{N}, 1 + 2\mathbf{N}\}$ è una partizione di \mathbf{N} [la partizione "dei pari e dei dispari"].

(ii) La famiglia $\mathfrak{U} = \{\{1\}, p\mathbf{N}, \forall p \text{ numero primo}\}$ è un ricoprimento di \mathbf{N} [infatti, come ben noto, ogni naturale $n \geq 2$ ha un fattore primo p e dunque $n \in p\mathbf{N}$], ma non è una partizione di \mathbf{N} [infatti $p_1, p_2 \in p_1\mathbf{N} \cap p_2\mathbf{N}$, se p_1, p_2 sono primi distinti].

(iii) Ogni insieme X ammette le due seguenti partizioni *banali*:

$$\mathfrak{U} = \{\{x\}, \forall x \in X\}, \quad \mathfrak{V} = \{X\}.$$

Veniamo ora alla definizione di *prodotto cartesiano* di una famiglia finita di insiemi, cominciando dal caso di due insiemi.

Definizione 7. Dati due insiemi A, B , si chiama *prodotto cartesiano di A e B* l'insieme

$$A \times B = \{(a, b), \forall a \in A, \forall b \in B\}.$$

L'elemento $(a, b) \in A \times B$ è detto *coppia (ordinata)* formata da a, b . Gli insiemi A e B sono detti rispettivamente *primo e secondo fattore del prodotto cartesiano*.

Osservazione 3. Le seguenti considerazioni sono pressoché ovvie [e la loro eventuale verifica è lasciata per esercizio].

(i) È evidente che $(a, b) \neq (b, a)$, se $a \neq b$. [Si noti invece che $\{a, b\} = \{b, a\}$: le coppie sono ordinate, gli insiemi non lo sono]. Inoltre: $(a, b) = (c, d) \iff a = c$ e $b = d$.

(ii) Ovviamente $A \times \emptyset = \emptyset \times B = \emptyset$. Inoltre:

$$A \times B = \emptyset \implies A = \emptyset \text{ oppure } B = \emptyset.$$

(iii) Risulta:

$$(A \times B) \cap (A \times C) = A \times (B \cap C), \quad (A \times B) \cup (A \times C) = A \times (B \cup C).$$

(iv) Se $A \subseteq X$ e $B \subseteq Y$, allora $A \times B \subseteq X \times Y$ e si ha:

$$\mathbf{C}_{X \times Y}(A \times B) = (\mathbf{C}_X(A) \times Y) \cup (X \times \mathbf{C}_Y(B)).$$

(v) Assegnati n insiemi A_1, A_2, \dots, A_n , si definisce loro *prodotto cartesiano* l'insieme

$$A_1 \times A_2 \times \dots \times A_n = \{(a_1, a_2, \dots, a_n), \forall a_i \in A_i\}.$$

(vi) Se A è un insieme e $n \geq 2$, si scrive semplicemente A^n in luogo di $\underbrace{A \times A \times \dots \times A}_{n \text{ fattori}}$. Gli elementi di A^n vengono chiamati *n-ple* (*di elementi di A*) [le 2-ple e le 3-ple vengono chiamate risp. *coppie* e *terne* di elementi di A].

Vale il seguente "Principio del prodotto", la cui dimostrazione è una semplice conseguenza del Principio generalizzato della somma.

Proposizione 4 (*Principio del prodotto*). Se A, B sono insiemi finiti, risulta:

$$|A \times B| = |A| \cdot |B|.$$

Dim. Siano $|A| = m, |B| = n$. Posto $A = \{a_1, \dots, a_m\}$, l'insieme $A \times B$ si può scrivere come unione dei sottoinsiemi (a due a due disgiunti) $\{a_1\} \times B, \dots, \{a_m\} \times B$. In base al principio generalizzato della somma ed al fatto (evidente) che $|\{a_i\} \times B| = |B|$:

$$|A \times B| = |\{a_1\} \times B| + \dots + |\{a_m\} \times B| = \underbrace{|B| + \dots + |B|}_{m \text{ addendi}} = m \cdot |B| = |A| \cdot |B|.$$

Di tale principio esiste anche la forma generalizzata, facilmente dimostrabile per via induttiva [e la verifica è rinviata ad un esercizio del paragrafo 3].

Proposizione 4' (*Principio generalizzato del prodotto*). Sia $k \geq 2$ e siano A_1, A_2, \dots, A_k insiemi finiti. Risulta:

$$|A_1 \times A_2 \times \dots \times A_k| = |A_1| \cdot |A_2| \cdot \dots \cdot |A_k|.$$

Facendo uso del prodotto cartesiano, c'è un semplice modo per "disgiungere" due insiemi non disgiunti.

Definizione 8. Dati due insiemi A, B , si chiama *unione disgiunta di A e B* l'insieme

$$A \sqcup B := (A \times \{1\}) \cup (B \times \{2\}).$$

Ad esempio, se $A = \{a, b, c\}$ e $B = \{c, d\}$, allora

$$A \sqcup B = \{(a, 1), (b, 1), (c, 1), (c, 2), (d, 2)\}$$

[mentre $A \cup B = \{a, b, c, d\}$]. Si noti che $A \times \{1\}$ e $B \times \{2\}$ sono in ogni caso disgiunti (indipendentemente da A e B). Se in particolare $|A| = m$ e $|B| = n$, ovviamente $|A \sqcup B| = m + n$.

Nota. Se $A \cap B = \emptyset$, $A \cup B$ viene spesso (ma impropriamente) identificato con $A \sqcup B$.

ESERCIZI PROPOSTI

1.1.1. Sono assegnati tre insiemi A, B, C .

- (i) Verificare che $A - (B - C) = (A - B) \cup (A \cap C)$.
- (ii) Verificare che $(A - B) - C = A - (B \cup C)$.
- (iii) Verificare che $(A - B) - C \subseteq A - (B - C)$ e che tale inclusione può essere propria.

1.1.2. Sono assegnati tre insiemi A, B, C .

- (i) Verificare che $(A \cup B) - C = (A - C) \cup (B - C)$ e che $(A \cap B) - C = (A - C) \cap (B - C)$.
- (ii) Verificare che $A \cap (B - C) = (A \cap B) \cap (A - C)$.
- (iii) Determinare un insieme T tale che $A \cup (B - C) = (A \cup B) \cap T$.

1.1.3 Tra i numeri naturali compresi tra 100 e 999, contare quelli che hanno esattamente due cifre uguali tra loro.

1.1.4. Trovare il numero dei naturali compresi tra 100 e 999 formati da cifre non nulle e a due a due distinte.

1.1.5. Contare i naturali tra 1 e 1000 che non sono divisibili né per 4, né per 5, né per 6.

1.1.6. Sia $a \in \mathbf{R}$. Dimostrare che, $\forall n \geq 1$:

$$a^n - 1 = (a - 1)(a^{n-1} + a^{n-2} + \dots + a + 1).$$

1.1.7. Determinare tutte le possibili partizioni di un insieme X di cardinalità 3.

2. Richiami sulle applicazioni

Cominciamo con una "definizione tautologica" di *applicazione* tra due insiemi.

Definizione 1. Siano A, B due insiemi. Un'applicazione (o funzione) $f : A \rightarrow B$ è una legge che ad ogni elemento $a \in A$ associa uno ed un solo elemento $b \in B$, che è detto immagine di a ed è usualmente denotato $f(a)$.

La tautologia di tale definizione sta nel fatto non abbiamo dato (né si può dare, senza ricorrere ad altri concetti non definiti) una definizione matematica di "legge". Una definizione non tautologica di applicazione si potrebbe in effetti dare interpretando le applicazioni da A a B come particolari sottoinsiemi del prodotto cartesiano $A \times B$. Ma, per non complicare la trattazione, converremo di assumere che anche il concetto di applicazione è primitivo.

Non tutte le "leggi" sono comunque funzioni. Siano infatti $A = \{a, b, c\}$ e $B = \{1, 2, 3, 4\}$ due insiemi. In base alla precedente definizione, le due seguenti leggi f, g non sono applicazioni.

$$f : \begin{cases} a \rightarrow 1 \\ b \rightarrow 2 \\ c \rightarrow 3 \\ \end{cases} \quad g : \begin{cases} a \searrow 1 \\ \quad \quad \quad 2 \\ b \rightarrow 3 \\ c \rightarrow 4 \end{cases}$$

[infatti nella prima l'elemento c non ha alcuna immagine, mentre nella seconda l'elemento a ha due immagini]. Invece è un'applicazione ad esempio la seguente legge h :

$$h : \begin{cases} a \rightarrow 1 \\ b \nearrow 2 \\ c \nearrow 3 \\ \end{cases}$$

Ovviamente due applicazioni $f : A \rightarrow B$, $g : A \rightarrow B$ sono dette *uguali* [e si scrive $f = g$] se risulta $f(a) = g(a)$, $\forall a \in A$. Sono quindi diverse [e si scrive $f \neq g$] se $\exists a \in A \mid f(a) \neq g(a)$.

Osservazione 1. Descriviamo alcune applicazioni "standard".

(1) *Applicazione identica.* Sia A un insieme. È sempre definita l'applicazione

$$\mathbf{1}_A : A \rightarrow A \text{ tale che } \mathbf{1}_A(a) = a, \quad \forall a \in A.$$

L'applicazione $\mathbf{1}_A$ è detta *applicazione identica* o *identità di A* .

(2) *Applicazione d'inclusione.* Sia $A' \subseteq A$. È definita l'applicazione

$$i : A' \hookrightarrow A \text{ tale che } i(a) = a, \quad \forall a \in A'.$$

i è detta *applicazione canonica d'inclusione (del sottoinsieme A' di A in A)*.

(3) *Restrizione di un'applicazione.* Sia $f : A \rightarrow B$ un'applicazione e sia $A' \subseteq A$. L'applicazione

$$f|_{A'} : A' \rightarrow B \text{ tale che } f|_{A'}(a') = f(a'), \quad \forall a' \in A'$$

è detta *restrizione di f ad A'* .

(4) *Funzione caratteristica di un sottoinsieme.* Sia $A' \subseteq A$. L'applicazione

$$\chi_{A'} : A \rightarrow \{0, 1\} \text{ tale che } \chi_{A'}(a) = \begin{cases} 0, & \text{se } a \notin A' \\ 1, & \text{se } a \in A' \end{cases}, \quad \forall a \in A,$$

è detta *funzione caratteristica di A' in A* .

(5) *Proiezioni canoniche.* Siano A, B due insiemi ed $A \times B$ il loro prodotto cartesiano. Le due applicazioni

$$p_1 : A \times B \rightarrow A \text{ tale che } p_1(a, b) = a, \quad p_2 : A \times B \rightarrow B \text{ tale che } p_2(a, b) = b, \quad \forall (a, b) \in A \times B,$$

sono dette rispettivamente *prima e seconda proiezione canonica* [dal prodotto cartesiano $A \times B$ rispettivamente ad A e a B].

Definizione 2. Siano $f : A \rightarrow B$ e $g : B \rightarrow C$ due applicazioni. Si chiama *composizione di f e g* ovvero *prodotto operatorio di f e g* l'applicazione

$$g \circ f : A \rightarrow C \text{ tale che } (g \circ f)(a) = g(f(a)), \quad \forall a \in A.$$

Osservazione 2. (i) Perché possa essere definita la composizione di due funzioni è necessario soltanto che l'insieme di arrivo della prima funzione sia contenuto nell'insieme di partenza della seconda.

(ii) Date tre applicazioni $f : A \rightarrow B$, $g : B \rightarrow C$ e $h : C \rightarrow D$, si verifica facilmente che vale per il prodotto operatorio la *proprietà associativa*, cioè: $h \circ (g \circ f) = (h \circ g) \circ f$.

Infatti, $\forall a \in A$: $(h \circ (g \circ f))(a) = h(g(f(a))) = ((h \circ g) \circ f)(a)$.

(iii) Data $f : A \rightarrow B$, risulta: $f \circ \mathbf{1}_A = f$, $\mathbf{1}_B \circ f = f$.

Infatti, $\forall a \in A$: $(f \circ \mathbf{1}_A)(a) = f(\mathbf{1}_A(a)) = f(a)$, $(\mathbf{1}_B \circ f)(a) = \mathbf{1}_B(f(a)) = f(a)$.

(iv) In generale, $f \circ g \neq g \circ f$. Ad esempio, posto $\begin{cases} f : \mathbf{R} \rightarrow \mathbf{R} \mid f(x) = x^2, \quad \forall x \in \mathbf{R}, \\ g : \mathbf{R} \rightarrow \mathbf{R} \mid g(x) = x + 1, \quad \forall x \in \mathbf{R}, \end{cases}$ risulta:
 $(g \circ f)(x) = x^2 + 1$, $(f \circ g)(x) = (x + 1)^2$, $\forall x \in \mathbf{R}$, e dunque $f \circ g \neq g \circ f$.

Definizione 3. Sia $f : A \rightarrow B$ un'applicazione. Per ogni $A' \subseteq A$, l'insieme

$$f(A') = \{f(a), \quad \forall a \in A'\}$$

è detto *immagine di A' tramite f* . In particolare, l'insieme $f(A)$ è detto *immagine di f* ed è denotato $Im f$ [ovvero $Im(f)$]. Per ogni $b \in B$, l'insieme

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

è detto *controimmagine di b tramite f* . Sono sinonimi di *controimmagine*: *fibra*, *antiimmagine* o *preimmagine*. Più generalmente, per ogni $B' \subseteq B$, l'insieme

$$f^{-1}(B') = \{a \in A \mid f(a) \in B'\}$$

è detto *controimmagine di B' tramite f* . Ovviamente risulta $f^{-1}(Im f) = f^{-1}(B) = A$. Si noti poi che, se $A_1 \subseteq A_2 \subseteq A$, allora $f(A_1) \subseteq f(A_2) \subseteq Im f$; se $B_1 \subseteq B_2 \subseteq B$, allora $f^{-1}(B_1) \subseteq f^{-1}(B_2) \subseteq A$.

Definizione 4. Sia $f : A \rightarrow B$ un'applicazione. f è detta *iniettiva* se risulta, $\forall a_1, a_2 \in A$,

$$f(a_1) = f(a_2) \implies a_1 = a_2$$

[ovvero se risulta: $a_1 \neq a_2 \implies f(a_1) \neq f(a_2)$, cioè se "elementi distinti di A hanno immagini distinte (in B)"]. Invece $f : A \rightarrow B$ è detta *suriettiva* se

$$\forall b \in B, \exists a \in A \mid f(a) = b$$

[ovvero se $Im f = B$, cioè se "ogni elemento di B è immagine di qualche elemento di A "]. Infine $f : A \rightarrow B$ è detta *bijettiva* se è iniettiva e suriettiva. In tal caso diciamo che A, B sono *in biiezione* o *in corrispondenza biunivoca* tramite f .

Tornando agli esempi di **Osserv.1** si verifica subito che l'identità $\mathbf{1}_A : A \rightarrow A$ è sempre bijettiva e che l'applicazione d'inclusione $i : A' \hookrightarrow A$ è sempre iniettiva. Inoltre la funzione caratteristica $\chi_{A'} : A \rightarrow \{0, 1\}$ è suriettiva se $\emptyset \neq A' \subset A$ mentre le due proiezioni canoniche p_1, p_2 sono sempre suriettive. Si noti poi che ogni applicazione $f : A \rightarrow B$ può essere sempre "suriettivizzata": basta definire l'applicazione

$$f_{su} : A \rightarrow Im f \text{ tale che } f_{su}(a) = f(a), \quad \forall a \in A,$$

detta *suriettificazione di f* . Ovviamente f_{su} è suriettiva. Inoltre ogni applicazione $f : A \rightarrow B$ si fattorizza sempre nella forma $f = i \circ f_{su}$, con $f_{su} : A \rightarrow Im f$ (suriettiva) e $i : Im f \hookrightarrow B$ (iniettiva).

L'applicazione h definita all'inizio del paragrafo non è né iniettiva né suriettiva. Relativamente poi alle due applicazioni f, g di **Osserv.2(iv)** si verifica subito che g è bijettiva mentre f non è né

iniettiva né suriettiva. Le due composizioni $f \circ g$ e $g \circ f$ non sono né iniettive né suriettive.

Osservazione 3. Sia $f : A \rightarrow B$ un'applicazione. Ricordato che con la notazione $|X|$ intendiamo la cardinalità di un insieme X , lasciamo per esercizio queste due semplici verifiche:

$$(1) \quad f \text{ è iniettiva} \iff |f^{-1}(b)| \leq 1, \quad \forall b \in B; \quad (2) \quad f \text{ è suriettiva} \iff |f^{-1}(b)| \geq 1, \quad \forall b \in B.$$

Ne segue subito che: f è biiettiva $\iff |f^{-1}(b)| = 1, \quad \forall b \in B$.

Veniamo ora ad un'importante caratterizzazione delle biiezioni.

Proposizione 1. Sia $f : A \rightarrow B$ un'applicazione. Risulta:

$$f \text{ è biiettiva} \iff \exists g : B \rightarrow A \text{ tale che } g \circ f = \mathbf{1}_A \text{ e } f \circ g = \mathbf{1}_B.$$

Se tale g esiste, allora è unica ed è detta *applicazione inversa di f* , denotata f^{-1} .

Dim. (\implies). Essendo f biiettiva, in base alla precedente osservazione: $\forall b \in B, \exists! a \in A$ tale che $f(a) = b$. Si definisce allora

$$g : B \rightarrow A \text{ tale che } g(b) = a, \text{ se } f(a) = b.$$

Risulta:

- $\forall a \in A : (g \circ f)(a) = g(f(a)) = g(b) = a = \mathbf{1}_A(a)$, e dunque $g \circ f = \mathbf{1}_A$.
- $\forall b \in B : (f \circ g)(b) = f(g(b)) = f(a) = b = \mathbf{1}_B(b)$, e dunque $f \circ g = \mathbf{1}_B$.

(\iff). Assumiamo che esista g . Verifichiamo che f è iniettiva. Sia $f(a_1) = f(a_2)$. Allora $g(f(a_1)) = g(f(a_2))$, cioè $\mathbf{1}_A(a_1) = \mathbf{1}_A(a_2)$ ovvero $a_1 = a_2$. Verifichiamo ora che f è suriettiva. Per ogni $b \in B$ risulta: $b = \mathbf{1}_B(b) = (f \circ g)(b) = f(g(b))$ e quindi $b \in \text{Im } f$. Dunque $\text{Im } f = B$.

Ora verifichiamo l'unicità di g . Sia $g_1 : B \rightarrow A$ tale che $g_1 \circ f = \mathbf{1}_A$ e $f \circ g_1 = \mathbf{1}_B$. Si tratta di verificare che $g_1 = g$. Si ha, $\forall b \in B : b = \mathbf{1}_B(b) = (f \circ g)(b) = (f \circ g_1)(b)$ e dunque $f(g(b)) = f(g_1(b))$. Essendo f iniettiva, allora $g(b) = g_1(b)$, da cui $g = g_1$.

Si osserva subito che $\mathbf{1}_A$ è biiettiva, con inversa se stessa. Se poi $f : A \rightarrow B$ è biiettiva, anche $f^{-1} : B \rightarrow A$ lo è (ha inversa f). Se infine $f : A \rightarrow B$ e $g : B \rightarrow C$ sono applicazioni biiettive, anche $g \circ f : A \rightarrow C$ è biiettiva e risulta $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$. Infatti si verifica facilmente che $(g \circ f) \circ (f^{-1} \circ g^{-1}) = \mathbf{1}_C$ e $(f^{-1} \circ g^{-1}) \circ (g \circ f) = \mathbf{1}_A$.

Osservazione 4. Si noti che con la stessa notazione f^{-1} abbiamo denotato due oggetti diversi tra loro, su cui bisogna riflettere con attenzione.

$f^{-1} : B \rightarrow A$ denota la funzione inversa di un'applicazione biiettiva $f : A \rightarrow B$; invece $f^{-1}(b)$ o $f^{-1}(B')$ sono insiemi, che denotano la controimmagine di un elemento b o di un sottoinsieme B' di B , rispetto ad una qualsiasi applicazione $f : A \rightarrow B$.

Quando f è biiettiva la controimmagine $f^{-1}(b)$ è formata da un solo elemento (di A) e quindi può essere identificata con l'immagine dell'elemento b dell'applicazione inversa f^{-1} .

Definizione 5. Le applicazioni biiettive di un insieme finito non vuoto A in sé vengono di solito chiamate *permutazioni* di A .

Se A è un insieme formato da un solo elemento, c'è un'unica permutazione di A , ed è l'identità $\mathbf{1}_A$. Se invece $|A|=2$, ad esempio $A = \{a, b\}$ [con $a \neq b$], le permutazioni di A sono esattamente 2, cioè

$$\mathbf{1}_A : \begin{cases} a \rightarrow a \\ b \rightarrow b, \end{cases} \quad f : \begin{cases} a \rightarrow b \\ b \rightarrow a. \end{cases}$$

Se poi $|A|=3$, ad esempio $A = \{a, b, c\}$ [con a, b, c a due a due diversi], si può verificare che le permutazioni di A sono 6, cioè

$$\begin{aligned} \mathbf{1}_A : & \begin{cases} a \rightarrow a \\ b \rightarrow b \\ c \rightarrow c, \end{cases} & f_2 : & \begin{cases} a \rightarrow b \\ b \rightarrow a \\ c \rightarrow c, \end{cases} & f_3 : & \begin{cases} a \rightarrow c \\ b \rightarrow b \\ c \rightarrow a, \end{cases} \\ f_4 : & \begin{cases} a \rightarrow a \\ b \rightarrow c \\ c \rightarrow b, \end{cases} & f_5 : & \begin{cases} a \rightarrow b \\ b \rightarrow c \\ c \rightarrow a, \end{cases} & f_6 : & \begin{cases} a \rightarrow c \\ b \rightarrow a \\ c \rightarrow b. \end{cases} \end{aligned}$$

Dedicheremo allo studio delle permutazioni su un insieme finito l'intero paragrafo 4. In questo paragrafo ci occuperemo invece di due problemi di conteggio:

- quante sono le applicazioni tra due insiemi finiti,
- quanti sono i sottoinsiemi di una data cardinalità in un insieme finito.

Cominciamo con il primo problema, introducendo una notazione.

Notazione. Dati due insiemi A, B , con

$$B^A := \{f : A \rightarrow B\}$$

denotiamo l'insieme di tutte le possibili applicazioni da A a B . Si noti in particolare che:

- se $B \neq \emptyset$, B^\emptyset ha un solo elemento [cioè l'inclusione canonica $i : \emptyset \hookrightarrow B$];
- se $A \neq \emptyset$, $\emptyset^A = \emptyset$ [in quanto manca l'immagine di ogni elemento di A].

Si noti infine che conviene considerare \emptyset^\emptyset indeterminato [infatti contrastano tra loro due fatti: esiste l'inclusione canonica $i : \emptyset \hookrightarrow \emptyset$, ma non esistono applicazioni prive di immagini].

Si osservi che la "strana" notazione B^A è giustificata dalla proposizione che segue

Proposizione 2. Siano A, B due insiemi finiti tali che $|A| = m$ e $|B| = n$. Allora $|B^A| = n^m$.

Dim. Sia $A = \{a_1, \dots, a_m\}$. Ogni applicazione $f \in B^A$ è completamente individuata dalla m -pla

$$(f(a_1), \dots, f(a_m)) \in B^m.$$

Ogni elemento $f(a_i)$ di tale m -pla può essere scelto in $n = |B|$ modi diversi [indipendentemente dalla scelta degli altri elementi della m -pla]. In base al principio generalizzato del prodotto, si hanno complessivamente $\underbrace{n \cdot n \cdot \dots \cdot n}_m = n^m$ possibili scelte per f . Dunque $|B^A| = n^m$.

Esempio 1. Siano $A = \{1, 2\}$ e $B = \{a, b, c\}$. Ogni $f : A \rightarrow B$ è completamente individuata dalla coppia $(f(1), f(2)) \in B \times B$. Gli elementi $f(1), f(2)$ possono essere scelti ciascuno in tre modi diversi. Dunque B^A è formato da $9 = 3^2$ applicazioni, associate alle nove coppie:

$$(a, a), (a, b), (a, c), (b, a), (b, b), (b, c), (c, a), (c, b), (c, c).$$

Si noti invece che A^B è formato da $8 = 2^3$ applicazioni, associate alle otto terne:

$$(1, 1, 1), (1, 1, 2), (1, 2, 1), (1, 2, 2), (2, 1, 1), (2, 1, 2), (2, 2, 1), (2, 2, 2).$$

Proposizione 3. Siano A, B due insiemi finiti.

- (i) Se $|A| > |B|$, non esistono applicazioni iniettive da A a B .
- (ii) Se $|A| < |B|$, non esistono applicazioni suriettive da A a B .
- (iii) Se $|A| = |B|$, ogni applicazione iniettiva da A a B è anche suriettiva e ogni applicazione suriettiva da A a B è anche iniettiva.

Dim. Se A, B sono insiemi finiti, valgono i seguenti fatti, intuitivamente ovvi:

$$\text{se } A \subseteq B, |A| \leq |B|; \quad \text{se } A \subseteq B \text{ e } |A| = |B|, \text{ allora } A = B.$$

Sia ora $f : A \rightarrow B$ un'applicazione [tra insiemi finiti]. In base alle precedenti osservazioni, si ha:

- $|Im f| \leq |B|$ [perché $Im f \subseteq B$];
- $|Im f| \leq |A|$ [perché $Im f = \{f(a_1), \dots, f(a_m)\}$, se $A = \{a_1, \dots, a_m\}$];
- $f : A \rightarrow B$ è iniettiva $\iff |Im f| = |A|$;

- $f : A \rightarrow B$ è suriettiva $\iff |Im f| = |B|$.

Si ha quindi:

(i) Se $f : A \rightarrow B$ è iniettiva, allora $|A| = |Im f| \leq |B|$. Se quindi $|A| > |B|$, non esistono applicazioni iniettive da A a B .

(ii) Se $f : A \rightarrow B$ è suriettiva, allora $|B| = |Im f| \leq |A|$. Se quindi $|A| < |B|$, non esistono applicazioni suriettive da A a B .

(iii) Sia $|A| = |B|$. Si ha: f è iniettiva $\iff |Im f| = |A| \iff |Im f| = |B| \iff f$ è suriettiva.

Nota. La precedente affermazione (i) è una formulazione del cosiddetto *Principio dei cassetti* (o *Principio delle gabbie di piccioni*): *se m oggetti vanno messi in n cassetti e $m > n$, un cassetto deve contenere almeno due oggetti*.

Proposizione 4. Siano A, B insiemi finiti (non vuoti), con $|A| = m$, $|B| = n$. Sia $1 \leq m \leq n$. Il numero delle applicazioni iniettive da A a B è $n(n-1)\dots(n-m+1)$.

Dim. Sia $A = \{a_1, \dots, a_m\}$ e sia $f : A \rightarrow B$ un'applicazione iniettiva. L'elemento $f(a_1)$ può essere scelto in B in n modi distinti. Per ogni $k = 2, \dots, m$, risulta che $f(a_k) \in B - \{f(a_1), \dots, f(a_{k-1})\}$. Dunque $f(a_k)$ può essere scelto in $n - (k - 1)$ modi distinti. In base al principio generalizzato del prodotto, le applicazioni iniettive da A a B sono $n(n-1)\dots(n-m+1)$.

Corollario 1. Siano A, B insiemi finiti (non vuoti), con la stessa cardinalità $n \geq 1$. Il numero delle applicazioni biettive da A a B è $n(n-1)\dots\cdot 2\cdot 1$.

Dim. Segue dalla **Prop. 4** e dalla **Prop. 3(iii)**.

Definizione 6. Per ogni $n \in \mathbf{N}$, $n \geq 1$, il numero $n(n-1)\dots\cdot 2\cdot 1$ è chiamato n fattoriale ed è denotato $n!$. Si definisce poi $0! = 1$.

In base al **Cor. 1**, se A è un insieme finito con cardinalità $n \geq 1$, le applicazioni biettive di A in A [cioè le *permutezioni* di A] sono $n!$.

Dopo aver affrontato il problema di contare le applicazioni tra insiemi finiti, affrontiamo il problema di contare i sottoinsiemi di una data cardinalità di un insieme finito. Introduciamo la seguente definizione.

Definizione 7. Sia A un insieme finito con n elementi. Per ogni $k \in \mathbf{N}$, tale che $0 \leq k \leq n$, si chiama coefficiente binomiale di n su k il numero dei sottoinsiemi di A formati da k elementi. Tale numero è denotato $\binom{n}{k}$.

Osservazione 5. Sia $A = \{a_1, a_2, \dots, a_n\}$. Allo scopo di ottenere una formula che calcoli $\binom{n}{k}$ [in funzione di n e k], premettiamo le seguenti elementari osservazioni:

$\binom{n}{0} = 1$ [infatti \emptyset è l'unico sottoinsieme di A con 0 elementi];

$\binom{n}{n} = 1$ [infatti A è l'unico sottoinsieme di A con n elementi];

$\binom{n}{1} = n$ [infatti $\{a_1\}, \dots, \{a_n\}$ sono tutti e soli i sottoinsiemi di A con 1 elemento];

$\binom{n}{n-1} = n$ [infatti $A - \{a_1\}, \dots, A - \{a_n\}$ sono tutti e soli i sottoinsiemi di A con $n-1$ elementi].

Proposizione 5. Per ogni $n, k \in \mathbf{N}$, tali che $1 \leq k \leq n$, risulta:

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}.$$

Dim. Si fissi in A un arbitrario elemento a_1 . Per ogni sottoinsieme B di A formato da k elementi, si hanno due alternative: $a_1 \in B$ oppure $a_1 \notin B$.

Se $a_1 \in B$, gli altri $k - 1$ elementi di B formano un insieme B_1 di cardinalità $n - 1$, contenuto in $A - \{a_1\}$. Dunque si hanno $\binom{n-1}{k-1}$ possibili sottoinsiemi B_1 (e quindi B).

Se $a_1 \notin B$, i k elementi di B vanno scelti in $A - \{a_1\}$. Dunque si hanno $\binom{n-1}{k}$ possibili sottoinsiemi B .

Complessivamente i possibili sottoinsiemi B di A aventi cardinalità k sono $\binom{n-1}{k-1} + \binom{n-1}{k}$.

Osservazione 6. La proposizione precedente afferma che è possibile calcolare $\binom{n}{k}$ conoscendo i coefficienti binomiali $\binom{n-1}{k-1}$ [o meglio due di essi: $\binom{n-1}{k}$ e $\binom{n-1}{k-1}$]. Se ordiniamo su righe i binomiali con lo stesso coefficiente "alto" n e su colonne i binomiali con lo stesso coefficiente "basso" k , otteniamo il seguente triangolo, detto *triangolo di Tartaglia* o *triangolo di Pascal*:

$$\begin{array}{ccccccc} & & \binom{0}{0} & & & & \\ & & \binom{1}{0} & \binom{1}{1} & & & \\ & & \binom{2}{0} & \binom{2}{1} & \binom{2}{2} & & \\ & & \binom{3}{0} & \binom{3}{1} & \binom{3}{2} & \binom{3}{3} & \\ & & & & & & \\ & & \binom{k}{0} & \binom{k}{1} & \dots & \dots & \dots & \binom{k}{k} \\ & & & & & & & \\ & & \dots & \dots & & & & \end{array}$$

Tenuto conto di **Osserv.** 5 e di **Prop.** 5, i valori numerici delle prime righe del triangolo di Tartaglia sono:

$$\begin{array}{ccccccccc} & & 1 & & & & & & \\ & & 1 & 1 & & & & & \\ & & 1 & 2 & 1 & & & & \\ & & 1 & 3 & 3 & 1 & & & \\ & & 1 & 4 & 6 & 4 & 1 & & \\ & & 1 & 5 & 10 & 10 & 5 & 1 & \\ & & 1 & 6 & 15 & 20 & 15 & 6 & 1 & \\ & & 1 & 7 & 21 & 35 & 35 & 21 & 7 & 1 \\ & & & & \dots & & & & \end{array}$$

Proposizione 6. (*Formula del binomiale*). Per ogni $n, k \in \mathbf{N}$, tali che $1 \leq k \leq n$, risulta:

$$\binom{n}{k} = \frac{n(n-1) \dots (n-k+1)}{k!}.$$

Dim. Sia A un insieme con n elementi e sia B un insieme con k elementi (con $1 \leq k \leq n$). Otterremo la formula cercata contando in due modi diversi le applicazioni iniettive da B ad A .

- (a) Dalla **Prop. 4**, la cardinalità delle applicazioni iniettive da B ad A è $n(n-1) \dots (n-k+1)$.
- (b) Ogni applicazione iniettiva $f : B \rightarrow A$ si fattorizza nella forma

$$f = i \circ f_{su},$$

dove $i : \text{Im } f \hookrightarrow A$ è l'inclusione canonica e $f_{su} : B \rightarrow \text{Im } f$ è la suriettivizzazione di f [ed è biiettiva, essendo f iniettiva]. $\text{Im } f$ è un sottoinsieme di k elementi di A : dunque può essere scelto in $\binom{n}{k}$ modi. f_{su} è una biiezione tra B e $\text{Im } f$: dunque può essere scelta in $k!$ modi. Dal principio del prodotto, si hanno per l'applicazione iniettiva f $\binom{n}{k} k!$ possibili scelte.

Dunque si ha: $\binom{n}{k} k! = n(n-1) \dots (n-k+1)$ e da ciò segue la formula cercata.

Proposizione 7. Per ogni $n, k \in \mathbf{N}$, tali che $0 \leq k \leq n$, risulta:

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} = \binom{n}{n-k}.$$

Dim. Se $k = 0$, il risultato è evidente. Sia $1 \leq k \leq n$. Si ha:

$$\binom{n}{k} = \frac{n(n-1)\dots(n-k+1)}{k!} \cdot \frac{(n-k)!}{(n-k)!} = \frac{n!}{k!(n-k)!}.$$

Ne segue che:

$$\binom{n}{n-k} = \frac{n!}{(n-k)!(n-(n-k))!} = \frac{n!}{(n-k)!k!} = \binom{n}{k}.$$

Osservazione 7. (i) Il motivo per cui $\binom{n}{k}$ è chiamato *coefficiente binomiale* discende dal fatto che vale la seguente formula [dimostrabile per induzione, cfr. gli esercizi di questo paragrafo]:

$$(x+y)^n = \sum_{k=0}^n \binom{n}{k} x^{n-k} y^k.$$

I coefficienti binomiali sono quindi i coefficienti dello sviluppo della potenza n -sima del binomio $x+y$.

(ii) In base alla **Def.3** ed alla **Prop. 1.3**, si ha subito:

$$\sum_{k=0}^n \binom{n}{k} = 2^n.$$

Concludiamo il paragrafo con la definizione di *operazione su un insieme*.

Definizione 8. Chiamiamo *operazione su un insieme non vuoto A* ogni applicazione da $A \times A$ ad A .

Ad esempio sono operazioni sugli insiemi numerici $\mathbf{N}, \mathbf{Z}, \mathbf{Q}, \mathbf{R}$ l'addizione e la moltiplicazione. Ad esempio

$$+ : \mathbf{Z} \times \mathbf{Z} \rightarrow \mathbf{Z} \text{ tale che } +((n, m)) = n + m, \quad \forall n, m \in \mathbf{Z},$$

è l'usuale addizione tra numeri interi.

Un altro esempio di operazione è la *composizione* sull'insieme delle biiezioni di ogni insieme non vuoto A [infatti abbiamo già osservato che la composizione $f \circ g$ di due biiezioni $f, g : A \rightarrow A$ è ancora una biiezione di A]. Se dunque indichiamo con $\mathcal{S}(A)$ l'insieme delle biiezioni di A in sé,

$$\circ : \mathcal{S}(A) \times \mathcal{S}(A) \rightarrow \mathcal{S}(A) \text{ tale che } \circ((f, g)) = g \circ f, \quad \forall f, g \in \mathcal{S}(A),$$

è un'operazione su $\mathcal{S}(A)$.

Le due operazioni che abbiamo considerato sono ovviamente ben diverse tra loro. Però hanno in comune tre importanti proprietà: sono entrambe *operazioni associative* [infatti $(n+m)+p = n+(m+p)$, $(f \circ g) \circ h = f \circ (g \circ h)$], hanno entrambe un "elemento neutro" [si tratta di 0 per l'addizione e dell'applicazione identica $\mathbf{1}_A$ per il prodotto operatorio], ogni elemento ammette un "elemento reciproco", cioè un elemento che, operando con esso, restituisce l'elemento neutro [ad esempio $n + (-n) = 0$, $f \circ f^{-1} = \mathbf{1}_A$].

Torneremo su tali proprietà nel **Cap. 2**, quando introdurremo la definizione di *gruppo*.

ESERCIZI PROPOSTI

1.2.1. Sia $f : A \rightarrow B$ un'applicazione. Siano A' e B' sottoinsiemi non vuoti rispettivamente di A e di B . Verificare che:

- (i) $f^{-1}(f(A')) \supseteq A'$. Se f è iniettiva, $f^{-1}(f(A')) = A'$.
- (ii) $f(f^{-1}(B')) \subseteq B'$. Se f è suriettiva, $f(f^{-1}(B')) = B'$.

1.2.2. Determinare due insiemi finiti A, B e due applicazioni $f : A \rightarrow B$, $g : B \rightarrow A$ tali che g ed f non sono biiettive, ma $g \circ f = \mathbf{1}_A$.

1.2.3. Dati gli insiemi $A = \{a, b, c\}$ e $B = \{1, 2\}$, quante sono e come si possono scrivere le

applicazioni suriettive da A a B ? E le applicazioni iniettive da B ad A ?

1.2.4. Sia $f : \mathbf{R} \rightarrow \mathbf{R}$ tale che $f(x) = \sin(x)$, $\forall x \in \mathbf{R}$. Determinare $f^{-1}([-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}])$, dove $[-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}]$ è l'intervallo chiuso di estremi $-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}$.

1.2.5. Sia $f : \mathbf{R} - \{1\} \rightarrow \mathbf{R}$ l'applicazione tale che $f(x) = \frac{x^2}{x-1}$, $\forall x \in \mathbf{R}, x \neq 1$. Per ogni $r \in \mathbf{R}$, determinare $f^{-1}(r)$.

1.2.6. Verificare, per induzione su $n \geq 0$, che per ogni $x, y \in \mathbf{R}$ risulta:

$$(x+y)^n = \sum_{k=0}^n \binom{n}{k} x^{n-k} y^k.$$

1.2.7. Per ogni $n \geq 1$, si ponga $S_n := 1 + 2 + 3 + \dots + n = \sum_{k=1}^n k$. Verificare che risulta:

$$S_n = \binom{n+1}{2}, \quad \forall n \geq 1.$$

1.2.8. Per ogni $n \geq 1$, si ponga

$$\Sigma_n := 1 + 3 + 5 + \dots + (2n-1) \quad [\text{somma dei primi } n \text{ numeri dispari}].$$

Verificare che risulta: $\Sigma_n = n^2$, $\forall n \geq 1$.

3. Relazioni su un insieme

Per introdurre il concetto di *relazione su un insieme*, conviene partire dal concetto di *grafico* associato alla relazione.

Definizione 1. Sia A un insieme non vuoto. Ogni sottoinsieme \mathfrak{R} di $A \times A$ è detto *grafico di una relazione ρ su A* [associata a \mathfrak{R}]. Tale relazione ρ è così definita: presa comunque una coppia $(a_1, a_2) \in A \times A$, si dice che a_1 e a_2 sono in relazione ρ [e si scrive $a_1 \rho a_2$] se $(a_1, a_2) \in \mathfrak{R}$. Si dice invece che a_1 e a_2 non sono in relazione ρ [e si scrive $a_1 \not\rho a_2$] se $(a_1, a_2) \notin \mathfrak{R}$.

Le relazioni su un insieme A sono quindi quante i sottoinsiemi di $A \times A$. Ogni insieme non vuoto A possiede in particolare le seguenti tre relazioni:

- la *relazione di uguaglianza* (o *relazione identica*) su A : $a \rho b \iff a = b, \forall a, b \in A$ [il grafico è l'insieme $\{(a, a), \forall a \in A\}$, detto *diagonale di A*].
- la *relazione caotica*: $a \rho b \iff a, b \in A$ [il grafico è l'insieme $A \times A$].
- la *relazione vuota*: $a \not\rho b, \forall a, b \in A$ [il grafico è l'insieme vuoto].

Se $A = \{a_1, a_2, \dots, a_n\}$ è un insieme finito, ogni relazione ρ su A può essere rappresentata in "forma cartesiana" (o "matriciale"), ponendo

$$a_i \times a_j = \begin{cases} 0, & \text{se } a_i \not\rho a_j \\ 1, & \text{se } a_i \rho a_j. \end{cases}$$

Ad esempio, se $A = \{a, b, c, d\}$ ed il grafico di ρ è $\{(a, a), (b, b), (c, c), (a, d), (c, d)\}$, allora ρ è rappresentata con la seguente tavola di valori 0, 1:

ρ	a	b	c	d
a	1	0	0	1
b	0	1	0	0
c	0	0	1	1
d	0	0	0	0

Definizione 2. Una relazione ρ su A è detta:

- *riflessiva*, se $a \rho a, \forall a \in A$;
- *simmetrica*, se $a \rho b \implies b \rho a, \forall a, b \in A$;
- *transitiva*, se $a \rho b$ e $b \rho c \implies a \rho c, \forall a, b, c \in A$;
- *antisimmetrica*, se $a \rho b$ e $b \rho a \implies a = b, \forall a, b \in A$
- *totale*, se risulta $a \rho b$ oppure $b \rho a, \forall a, b \in A$.

Una relazione riflessiva, simmetrica e transitiva è detta *relazione di equivalenza*. Una relazione riflessiva e transitiva è detta *relazione di pre-ordine*. Una relazione riflessiva, antisimmetrica e transitiva è detta *relazione di ordine*. Infine una relazione di ordine che è anche totale è detta *relazione di ordine totale*.

Esempi 1. (i) La relazione di uguaglianza è riflessiva, simmetrica, transitiva, antisimmetrica, ma non totale (se $|A| \geq 2$). Dunque è una relazione di equivalenza e di ordine (non totale).

(ii) La relazione caotica è riflessiva, simmetrica, transitiva, totale, ma non antisimmetrica. La relazione vuota è simmetrica, antisimmetrica, transitiva [in modo banale], ma non è riflessiva né totale.

(iii) La relazione ρ definita sopra non ha alcuna di queste proprietà.

(iv) Negli insiemi numerici $\mathbf{N}, \mathbf{Z}, \mathbf{Q}, \mathbf{R}$ è definita la *relazione di disuguaglianza* \leq , in questo modo: in \mathbf{N} e \mathbf{Z} : $a \leq b \stackrel{\text{def}}{\iff} b = a + t, \exists t \in \mathbf{N}$; in \mathbf{Q} ed \mathbf{R} : $a \leq b \stackrel{\text{def}}{\iff} b = a + t, \exists t \geq 0$.

Si tratta di una relazione di ordine totale.

Esaminiamo ora tre relazioni che ci interesseranno particolarmente nel seguito: la *relazione di divisibilità in \mathbf{N}* (o in \mathbf{Z}), la *relazione di congruenza modulo un intero* e la *relazione di equipotenza tra insiemi*.

Definizione 3. In \mathbf{N} introduciamo la seguente relazione di divisibilità. Presi comunque $a, b \in \mathbf{N}$:

$$a | b \stackrel{\text{def}}{\iff} b = at, \exists t \in \mathbf{N},$$

[$a | b$ si legge: a divide b , oppure a è un divisore di b , oppure anche b è un multiplo di a . Se a non divide b si scrive $a \nmid b$].

Proposizione 1. La relazione di divisibilità in \mathbf{N} è una relazione di ordine (non totale).

Dim. Presi comunque $a, b, c \in \mathbf{N}$ sono verificate le proprietà:

riflessiva: $a | a$ [infatti $a = a \cdot 1$];

transitiva: $a | b, b | c \implies a | c$ [infatti $b = at, c = bs \implies c = ast$];

antisimmetrica: $a | b, b | a \implies a = b$ [infatti $b = at, a = bs \implies b = bst \implies st = 1 \implies s = 1 \implies b = a$].

La relazione di divisibilità su \mathbf{N} è quindi una relazione d'ordine. Ma non è totale: ad esempio $2 \nmid 3$ e $3 \nmid 2$. Non è neppure simmetrica: infatti $a | b \not\implies b | a$ [ad esempio $2 | 4$ ma $4 \nmid 2$].

N.B. La relazione di divisibilità può essere introdotta anche in \mathbf{Z} [con definizione analoga: $a | b \iff b = at, \exists t \in \mathbf{Z}$]. In tal caso è solo una relazione di pre-ordine, in quanto non è verificata l'antisimmetria [ad esempio $2 | -2, -2 | 2$ ma $2 \neq -2$].

Definizione 4. Sia n un intero ≥ 2 . Si chiama *relazione di congruenza modulo n* la seguente relazione su \mathbf{Z} . Presi comunque $a, b \in \mathbf{Z}$:

$$a \equiv b \pmod{n} \stackrel{\text{def}}{\iff} n | b - a$$

[cioè $b - a = nt, \exists t \in \mathbf{Z}$]. Si dice in tal caso che a è *conguente* (o *congruo*) a b modulo n . In luogo di $a \equiv b \pmod{n}$ si può anche scrivere $a \equiv b \pmod{n}$ oppure $a \equiv_n b$.

Proposizione 2. La relazione \equiv_n è una relazione di equivalenza su \mathbf{Z} .

Dim. Presi comunque $a, b, c \in \mathbf{Z}$, sono verificate le proprietà

riflessiva: $a \equiv a \pmod{n}$ [infatti $a - a = n \cdot 0$];

simmetrica: $a \equiv b \pmod{n} \implies b \equiv a \pmod{n}$ [infatti $b - a = nt \implies a - b = n(-t)$];

transitiva: $a \equiv b \pmod{n}, b \equiv c \pmod{n} \implies a \equiv c \pmod{n}$ [infatti, se $b - a = nr, c - b = ns$ (con $r, s \in \mathbf{Z}$), allora $c - a = (c - b) + (b - a) = n(s + r)$ e quindi $n | c - a$].

Definizione 5. Due insiemi X, Y sono detti *equipotenti* (e si scrive $X \sim Y$) se esiste un'applicazione biiettiva $f : X \rightarrow Y$. La relazione \sim è detta *relazione di equipotenza* (nella famiglia di tutti gli insiemi).

Proposizione 3. La relazione di equipotenza tra insiemi è una relazione di equivalenza.

Dim. Presi comunque tre insiemi X, Y, Z , sono verificate le proprietà

riflessiva: $X \sim X$ [tramite la biiezione identica $\mathbf{1}_X$];

simmetrica: $X \sim Y \implies Y \sim X$ [se $f : X \rightarrow Y$ è biiettiva, anche $f^{-1} : Y \rightarrow X$ lo è];

transitiva: $X \sim Y, Y \sim Z \implies X \sim Z$ [se $f : X \rightarrow Y, g : Y \rightarrow Z$ sono biiettive, anche $g \circ f : X \rightarrow Z$ lo è].

Ci occuperemo ora esclusivamente delle *relazioni di equivalenza*, cioè delle relazioni riflessive, simmetriche e transitive.

Definizione 6. Sia ρ una relazione di equivalenza su A . Per ogni $a \in A$, il sottoinsieme di A

$$[a] = [a]_\rho := \{x \in A : a \rho x\}$$

è detto *classe di equivalenza di a modulo ρ* . Poiché ρ è simmetrica, $[a] = \{x \in A : x \rho a\}$.

Proposizione 4. Sia ρ una relazione di equivalenza su A . Si ha:

- (i) $a \in [a], \forall a \in A$.
- (ii) $[a] = [b] \iff a \rho b$.
- (iii) $[a] \cap [b] = \emptyset \iff a \not\rho b$.
- (iv) Le classi di equivalenza modulo ρ (a due a due distinte) formano una partizione di A .

Dim. (i) Da $a \rho a$ segue che $a \in [a]$.

(ii) (\implies). Poiché $b \in [b] = [a]$, allora $a \rho b$.

(\iff). Sia $x \in [a]$. Si ha: $x \rho a, a \rho b$ e quindi, per transitività, $x \rho b$, cioè $x \in [b]$. Dunque $[a] \subseteq [b]$. In modo analogo si verifica che $[b] \subseteq [a]$.

(iii) Dimostreremo, equivalentemente, che $[a] \cap [b] \neq \emptyset \iff a \rho b$.

(\implies). Se $x \in [a] \cap [b]$, allora $a \rho x, x \rho b$ e quindi (per transitività) $a \rho b$.

(\iff). Segue da (ii) e (i).

(iv) La famiglia

$$\mathfrak{U} = \mathfrak{U}_\rho = \{[a], \forall a \in A\},$$

formata da tutte le classi di equivalenza [a due a due distinte] di A (modulo ρ), è un ricoprimento di A [in base a (i)]; inoltre due classi distinte sono disgiunte [in base a (ii) e (iii)]. Ne segue che \mathfrak{U} è una partizione di A .

Osservazione 1. Dalla proposizione precedente segue che ogni relazione di equivalenza ρ induce una partizione \mathfrak{U}_ρ .

Vale anche il viceversa. Sia infatti $\mathfrak{U} = \{A_i, i \in I\}$ una partizione di A (formata quindi da sottoinsiemi non vuoti); possiamo definire su A la seguente relazione $\rho = \rho_{\mathfrak{U}}$:

$$a \rho b \iff a, b \in A_i, \exists i \in I.$$

Si verifica subito che ρ è una relazione di equivalenza su A . Le classi di equivalenza *modulo ρ* sono gli insiemi $A_i \in \mathfrak{U}$. Si verifichi poi che $\mathfrak{U}_{\rho_{\mathfrak{U}}} = \mathfrak{U}$ e che $\rho_{\mathfrak{U}_{\rho}} = \rho$.

Definizione 7. Sia ρ una relazione di equivalenza su A . Si chiama *insieme quoziante di A modulo ρ* l'insieme - denotato A/ρ - formato dalle classi di equivalenza (a due a due distinte) di A modulo ρ , cioè

$$A/\rho = \{[a], \forall a \in A\}.$$

L'applicazione

$$\pi : A \rightarrow A/\rho, \text{ tale che } \pi(a) = [a], \forall a \in A,$$

è ovviamente suriettiva ed è chiamata *proiezione canonica di A su A/ρ* .

Delle classi di equivalenza e dell'insieme quoziante modulo la relazione di congruenza introdotta nella **Def. 4** parleremo diffusamente nel prossimo capitolo.

Consideriamo invece la relazione di equipotenza \sim introdotta nella **Def. 5**. Dato un insieme X , la sua classe di equivalenza $[X]_\sim$, denotata usualmente $|X|$, è detta *cardinalità di X* . $|X|$ è formata da tutti gli insiemi Y tali che esiste una biiezione $f : X \rightarrow Y$.

In particolare, ogni insieme finito formato da n elementi (a due a due distinti) è equipotente a $\{1, 2, \dots, n\}$ ovvero a $\{0, 1, \dots, n-1\}$. Se infatti $A = \{a_1, a_2, \dots, a_n\}$, l'applicazione

$$f : \{0, 1, \dots, n-1\} \rightarrow A \text{ tale che } f(i) = a_{i+1}, \quad \forall i = 0, \dots, n-1,$$

è certamente biiettiva.

Per indicare le cardinalità finite si utilizzano usualmente i numeri naturali - come del resto abbiamo già fatto nel primo paragrafo - in questo modo:

$$0 := |\emptyset|, \quad 1 := |\{0\}|, \quad 2 := |\{0, 1\}|, \quad \dots, \quad k := |\{0, 1, \dots, k-1\}|, \quad \dots$$

Le cardinalità degli insiemi \mathbf{N} ed \mathbf{R} sono dette rispettivamente *cardinalità del numerabile* e *cardinalità del continuo*.

Utilizzando la definizione di cardinalità data sopra siamo ora in grado di dimostrare il *Principio della somma*, presentato nel primo paragrafo (cfr. **Prop. 1.1**).

Dim. (*Principio della somma*). Siano A, B due insiemi finiti disgiunti, di cardinalità rispettivamente m, n . Sia $A = \{a_1, \dots, a_m\}$ e $B = \{b_1, \dots, b_n\}$. Si consideri l'applicazione $f : \{1, \dots, m+n\} \rightarrow A \cup B$ così definita:

$$f(1) = a_1, \dots, f(m) = a_m, \quad f(m+1) = b_1, \dots, f(m+n) = b_n.$$

Tale applicazione è ovviamente biiettiva e quindi $|A \cup B| = |\{1, \dots, m+n\}| = m+n = |A| + |B|$.

Torniamo allo studio delle relazioni di equivalenza, verificando per prima cosa che ogni applicazione definisce una relazione di equivalenza sul suo insieme di partenza.

Definizione 8. Ad ogni applicazione $f : A \rightarrow B$ resta "canonicamente" associata una relazione ρ_f su A , così definita:

$$a_1 \rho_f a_2 \iff f(a_1) = f(a_2), \quad \forall a_1, a_2 \in A.$$

Si verifica subito che ρ_f è una relazione di equivalenza su A , detta *relazione di equivalenza associata ad f* .

Si noti che, se f è iniettiva, la relazione ρ_f è la relazione identica su A (e viceversa). Inoltre

$$[a]_{\rho_f} = \{x \in A : f(x) = f(a)\} = f^{-1}(f(a)), \quad \forall a \in A.$$

Ne segue che l'insieme quoziente di A modulo ρ_f è dato da

$$A /_{\rho_f} = \{f^{-1}(b), \quad \forall b \in \text{Im}(f)\}.$$

Vedremo ora come attraverso questo insieme quoziente sia possibile "iniettivizzare" una qualsiasi applicazione f .

Proposizione 5. Sia $f : A \rightarrow B$ un'applicazione e sia ρ_f la relazione di equivalenza associata ad f . È ben definita l'applicazione

$$F : A /_{\rho_f} \rightarrow B \text{ tale che } F([a]) = f(a), \quad \forall [a] \in A /_{\rho_f}.$$

Inoltre F è iniettiva.

Dim. Dimostrare che F è "ben definita", significa dimostrare che la definizione di F non dipende dal rappresentante scelto in ogni classe, cioè che

$$[a] = [a_1] \implies F([a]) = F([a_1]).$$

Infatti: $[a] = [a_1] \iff a \rho_f a_1 \iff f(a) = f(a_1) \iff F([a]) = F([a_1]).$

Dalle precedenti implicazioni (lette da destra a sinistra) segue che $F([a]) = F([a_1]) \implies [a] = [a_1]$, cioè che F è iniettiva.

L'applicazione F agisce dunque come f , ma è definita su A/ρ_f (e non su A). La sua prerogativa è quella di trattare tutti gli elementi di A aventi la stessa immagine come se si trattasse di un unico elemento: dunque "iniettività" f .

N.B. Perché per iniettivizzare f non abbiamo semplicemente "rimpicciolito" A , eliminando da A tutti gli elementi che hanno la stessa immagine tranne uno? In qualche caso si potrebbe anche fare, ma quale elemento conservare? Si porrebbe dunque un problema di scelta e la costruzione non sarebbe "standard".

Osservazione 2. (i) Si noti che F è l'unica applicazione tale che $F \circ \pi = f$, cioè tale che il seguente diagramma (di insiemi e applicazioni)

$$\begin{array}{ccc} A & \xrightarrow{f} & B \\ \pi \downarrow & \nearrow F & \\ A/\rho_f & & \end{array}$$

è commutativo [nel senso che "il passaggio da un insieme ad un altro è indipendente da ogni possibile percorso (che usi le applicazioni assegnate)". Infatti, se $\tilde{F} : A/\rho_f \rightarrow B$ verifica $\tilde{F} \circ \pi = f$, allora $\tilde{F}([a]) = (\tilde{F} \circ \pi)(a) = f(a) = F([a])$, $\forall [a] \in A/\rho_f$. Dunque $\tilde{F} = F$.

(ii) Risulta: F è suriettiva $\iff f$ è suriettiva. Infatti:

$$\begin{aligned} F \text{ è suriettiva} &\iff \forall b \in B, \exists [a] \in A/\rho_f \text{ tale che } F([a]) = b \iff \\ &\iff \forall b \in B, \exists a \in A \text{ tale che } f(a) = b \iff f \text{ è suriettiva.} \end{aligned}$$

Concludiamo con il seguente risultato, che ci consentirà di esprimere in maniera "standard" ogni applicazione come prodotto operatorio di tre applicazioni: una suriettiva, una biiettiva ed una iniettiva

Proposizione 6. (*Teorema di decomposizione delle applicazioni*). Sia $f : A \rightarrow B$ un'applicazione tra insiemi. Esiste un'unica biiezione $\varphi : A/\rho_f \rightarrow \text{Im}(f)$ tale che $f = i \circ \varphi \circ \pi$, cioè tale che rende commutativo il diagramma:

$$\begin{array}{ccc} A & \xrightarrow{f} & B \\ \pi \downarrow & & \uparrow i \\ A/\rho_f & \xrightarrow{\varphi} & \text{Im}(f) \end{array}$$

Dim. In base alla **Prop. 5**, l'applicazione

$$F : A/\rho_f \rightarrow B \text{ tale che } F \circ \pi = f$$

è iniettiva. La sua suriettificazione F_{su} è quindi biiettiva. Si osserva poi subito che $\text{Im } f = \text{Im } F$. Posto $\varphi := F_{su}$ e considerata l'inclusione canonica $i : \text{Im } f \hookrightarrow B$, si ha, $\forall a \in A$:

$$(i \circ \varphi \circ \pi)(a) = (i \circ \varphi)([a]) = i(\varphi([a])) = i(F([a])) = i(f(a)) = f(a).$$

Quindi $f = i \circ \varphi \circ \pi$.

L'unicità di φ è evidente [si verifichi che se $i \circ \psi \circ \pi = i \circ \varphi \circ \pi$, allora $\psi = \varphi$].

Ad esempio, assegnata l'applicazione

$$f : \mathbf{R} \rightarrow \mathbf{R} \text{ tale che } f(x) = x^2 - 1, \quad \forall x \in \mathbf{R},$$

vogliamo decomporla nel senso descritto nella proposizione precedente. Si ha:

$$\text{Im } f = \{r \in \mathbf{R} : r \geq -1\} =: [-1, \infty) \text{ [intervallo di } \mathbf{R}\text{].}$$

Infatti: $r \in \text{Im } f \iff \exists x \in \mathbf{R} : x^2 - 1 = r \iff \exists x \in \mathbf{R} : x^2 = r + 1 \iff r + 1 \geq 0.$

Inoltre, $\forall x, y \in \mathbf{R} : x \rho_f y \iff x^2 - 1 = y^2 - 1 \iff y = \pm x.$ Dunque

$$[x] = \{\pm x\}, \quad \forall x \in \mathbf{R}.$$

L'insieme $\mathbf{R}/_{\rho_f}$ si identifica con l'insieme $\mathbf{R}^{\geq 0}$ dei reali ≥ 0 e risulta: $f = i \circ \varphi \circ \pi,$ con

$$\pi : \mathbf{R} \rightarrow \mathbf{R}/_{\rho_f} \text{ tale che } \pi(x) = [x], \quad \forall x \in \mathbf{R};$$

$$\varphi : \mathbf{R}/_{\rho_f} \rightarrow [-1, \infty) \text{ tale che } \varphi([x]) = x^2 - 1, \quad \forall [x] \in \mathbf{R}/_{\rho_f};$$

$$i : [-1, \infty) \hookrightarrow \mathbf{R} \text{ tale che } i(x) = x, \quad \forall x \in [-1, \infty).$$

ESERCIZI PROPOSTI

1.3.1. Sia $A = \{a, b, c\}$ un insieme (di cardinalità 3). Sia ρ la relazione su A avente grafico

$$\mathfrak{R} = \{(a, a), (a, b), (a, c), (b, b), (b, c), (c, c)\}.$$

Verificare se ρ è una relazione d'ordine totale su $A.$

1.3.2. Dimostrare il principio generalizzato del prodotto: *se $k \geq 2$ e A_1, A_2, \dots, A_k sono insiemi finiti, si ha:*

$$|A_1 \times A_2 \times \dots \times A_k| = |A_1| \cdot |A_2| \cdot \dots \cdot |A_k|.$$

1.3.3. Determinare una biiezione tra \mathbf{N} e \mathbf{Z} [ciò dimostra che $|\mathbf{Z}| = |\mathbf{N}|$, cioè che \mathbf{Z} è numerabile].

4. Permutazioni di un insieme finito

Considerato un insieme finito non vuoto X , studieremo l'insieme $\mathbf{S}(X)$ delle permutazioni di X .

La prima osservazione da fare è che non importa il nome e la natura degli elementi di X , mentre è ovviamente importante come gli elementi di X vengono "mossi" dalle singole permutazioni. È lecito quindi ed è prassi comune identificare tali elementi con i numeri naturali $1, 2, 3, \dots$. Se quindi X ha cardinalità $n \geq 1$, potremo identificare X con $\{1, 2, 3, \dots, n\}$ e denotare $\mathbf{S}(X)$ semplicemente con \mathbf{S}_n .

Dunque \mathbf{S}_n denota l'insieme delle permutazioni di ogni insieme di n elementi. Per indicare le permutazioni di \mathbf{S}_n vengono spesso usate lettere dell'alfabeto greco come σ, τ ecc..

In base al **Coroll. 1** di **Cap. 1.2**, \mathbf{S}_n ha cardinalità $n!$. Dobbiamo però stabilire la strategia da adottare per poter scrivere tutte queste $n!$ permutazioni.

Possiamo procedere come segue. Prima consideriamo tutte le permutazioni che trasformano 1 in 1. Si tratta di $(n - 1)!$ permutazioni [quelle che permutano gli elementi dell'insieme $\{2, \dots, n\}$]. Poi consideriamo quelle che trasformano 1 in 2. Si tratta ancora di $(n - 1)!$ permutazioni [quelle che permutano gli elementi dell'insieme $\{1, 3, \dots, n\}$]. Procediamo in questo modo fino ad ottenere n insiemi (a due a due disgiunti) formati ciascuno da $(n - 1)!$ permutazioni, cioè complessivamente (per il principio generalizzato della somma) $n!$ permutazioni.

Ora cerchiamo un modo efficiente per indicare come una permutazione $\sigma \in \mathbf{S}_n$ agisce sui singoli elementi di X . Ovviamente, ponendo σ_i al posto di $\sigma(i)$, possiamo scrivere

$$\sigma : \begin{cases} 1 \longrightarrow \sigma_1 \\ 2 \longrightarrow \sigma_2 \\ \vdots \\ n \longrightarrow \sigma_n. \end{cases}$$

Ma per economia di spazio è preferibile scrivere la stessa σ nella forma

$$\sigma = \begin{pmatrix} 1 & 2 & \dots & n \\ \sigma_1 & \sigma_2 & \dots & \sigma_n \end{pmatrix},$$

convenendo quindi che ogni elemento della prima riga sia mandato da σ in quello disposto esattamente al di sotto. Ad esempio la permutazione identica $\mathbf{1}_x$ di \mathbf{S}_4 si scrive nella forma

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

Nel seguito del paragrafo troveremo però un modo "più economico" (ad una sola riga) per scrivere le permutazioni.

Due permutazioni possono ovviamente essere "composte", secondo l'operazione di prodotto operatorio (o composizione). Se quindi $\sigma, \tau \in \mathbf{S}_n$, allora

$$\tau \circ \sigma = \begin{pmatrix} 1 & 2 & \dots & n \\ \tau_{\sigma_1} & \tau_{\sigma_2} & \dots & \tau_{\sigma_n} \end{pmatrix}.$$

Se ad esempio $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix}$, $\tau = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 1 & 3 \end{pmatrix} \in \mathbf{S}_4$, allora

$$\tau \circ \sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 1 & 3 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 3 & 2 \end{pmatrix}.$$

Osserviamo poi che, nel calcolare $\tau \circ \sigma$, prima agisce σ e poi τ . Quindi, come sopra evidenziato, occorre seguire "a ritroso" l'azione di $\tau \circ \sigma$ sugli elementi $1, 2, \dots, n$. Poiché ciò è in contrasto con

la nostra abitudine di leggere da sinistra verso destra, scriveremo $\sigma\tau$ al posto di $\tau\circ\sigma$, per cui

$$\sigma\tau = \begin{pmatrix} 1 & 2 & 3 & 4 \\ \downarrow & & & \\ 2 & 3 & 4 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 \\ \downarrow & & & \\ 2 & 4 & 1 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 3 & 2 \end{pmatrix} = \tau\circ\sigma.$$

Dunque abbiamo convenuto di definire

$$\sigma\tau = \tau\circ\sigma, \quad \forall \sigma, \tau \in S_n.$$

Tale convenzione non è generalmente adottata nei testi matematici che si occupano di questi argomenti. Molti preferiscono mantenere la notazione con il prodotto operatorio e leggere a ritroso l'azione della composizione sui singoli elementi.

Sappiamo che ogni permutazione σ , in quanto applicazione biiettiva, è dotata di inversa σ^{-1} . Assegnata σ possiamo subito ottenere σ^{-1} osservando che, se $\sigma : i \rightarrow \sigma_i$, allora $\sigma^{-1} : \sigma_i \rightarrow i$. Dunque per determinare σ^{-1} basta associare ad ogni elemento la sua immagine guardando σ dal basso verso l'alto. Ad esempio, se

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix}, \text{ allora } \sigma^{-1} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 2 & 3 \end{pmatrix}.$$

[Si verifichi che $\sigma^{-1}\sigma = \mathbf{1}_x = \sigma\sigma^{-1}$].

Si osserva subito che, $\forall \sigma \in S_n$, risulta:

$$\sigma\mathbf{1}_x = \sigma = \mathbf{1}_x\sigma.$$

Dunque la permutazione identica $\mathbf{1}_x$ funge da "elemento neutro" rispetto all'operazione di composizione. Un'altro fatto da osservare è che in generale $\sigma\tau \neq \tau\sigma$, cioè che la composizione di permutazioni non è in generale un'operazione commutativa. Ad esempio scelti in S_3 :

$$\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \quad \tau = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix},$$

si ha:

$$\sigma\tau = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}, \quad \tau\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}.$$

Per semplificare le notazioni porremo nel seguito, $\forall n \geq 1$:

$$\sigma^n := \underbrace{\sigma \cdot \dots \cdot \sigma}_{n \text{ fattori}}, \quad \sigma^0 := \mathbf{1}_x, \quad \sigma^{-n} := (\sigma^{-1})^n.$$

Ne segue che, $\forall n, m \in \mathbf{Z}$: $\sigma^{n+m} = \sigma^n \cdot \sigma^m$.

Come promesso, presentiamo ora un modo più economico di scrivere le permutazioni. Abbiamo bisogno di introdurre certe permutazioni speciali, dette *cicli*.

Definizione 1. Sia $\sigma \in S_n$ e sia k un intero tale che $1 \leq k \leq n$. La permutazione σ è detta *ciclo* o *k-ciclo* o *ciclo di lunghezza k* se $\exists c_1, c_2, \dots, c_k \in X$, a due a due distinti, tali che

$$\sigma(c_1) = c_2, \quad \sigma(c_2) = c_3, \quad \dots, \quad \sigma(c_{k-1}) = c_k, \quad \sigma(c_k) = c_1 \quad \text{e} \quad \sigma(t) = t, \quad \forall t \in X, \quad t \neq c_1, c_2, \dots, c_k.$$

Tale *k-ciclo* σ è usualmente denotato (c_1, c_2, \dots, c_k) , ovvero, più brevemente, $(c_1 c_2 \dots c_k)$. I 2-cicli vengono anche chiamati *trasposizioni*.

Diremo poi, per brevità, che i naturali c_1, \dots, c_k sono "elementi" del *ciclo* $(c_1 c_2 \dots c_k)$. Infine, due cicli di S_n , di lunghezza ≥ 2 , sono detti *cicli disgiunti* se non hanno elementi in comune.

Ad esempio, considerata in S_5 la permutazione

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 4 & 5 & 1 \end{pmatrix},$$

si osserva subito che essa è il 5-ciclo $(1 2 3 4 5)$. Tale ciclo può essere anche scritto in altre forme:

infatti $(1\ 2\ 3\ 4\ 5) = (2\ 3\ 4\ 5\ 1) = (3\ 4\ 5\ 1\ 2) = (4\ 5\ 1\ 2\ 3) = (5\ 1\ 2\ 3\ 4)$.

Più in generale possiamo dire che, se $2 \leq k \leq n$, ogni k -ciclo si scrive in k modi diversi [basta iniziare la scrittura da uno qualsiasi dei suoi k elementi e scriverli tutti, mantenendone l'ordine, con una sorta di "rotazione oraria"].

Di 1-cicli ne esiste uno solo e coincide con la permutazione identica $\mathbf{1}_X$. Infatti, $\forall t \in X$, l'1-ciclo (t) fissa ogni elemento di X . Ovviamente $(t) = (s)$, $\forall s \in X$.

È facile, assegnato un ciclo di S_n , scriverlo in forma di permutazione. Ad esempio, considerato il 4-ciclo $\sigma = (3\ 5\ 4\ 1) \in S_6$, risulta:

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & & 5 & 1 & 4 & \end{pmatrix}$$

e quindi, scrivendo anche le immagini degli altri due elementi (che restano fissi), si ottiene

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 2 & 5 & 1 & 4 & 6 \end{pmatrix}.$$

Osservazione 1. (i) Tutte le permutazioni di S_1, S_2, S_3 sono cicli. Infatti:

$$S_1 = \{(1)\}, \quad S_2 = \{(1), (1\ 2)\}, \quad S_3 = \{(1), (1\ 2), (1\ 3), (2\ 3), (1\ 2\ 3), (1\ 3\ 2)\}.$$

Si noti che i cicli scritti sopra vanno ovviamente interpretati nel corrispondente S_n . Ad esempio $(1\ 2) \in S_2$ è la permutazione $\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$, mentre $(1\ 2) \in S_3$ è la permutazione $\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$.

(ii) Non ogni permutazione $\sigma \in S_n$ è un ciclo. Ad esempio la permutazione

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix} \in S_4$$

non è un ciclo. Si vede subito però che essa contiene due cicli disgiunti, e cioè le due trasposizioni $(1\ 2)$ e $(3\ 4)$. In effetti, essa è proprio il prodotto di queste due trasposizioni. Infatti

$$(1\ 2)(3\ 4) = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 3 & 4 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 4 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix} = \sigma.$$

Tale fatto non è casuale, come messo in luce nella successiva **Prop.1**.

(iii) Come si ottengono i cicli (a due a due disgiunti) di una permutazione non identica $\sigma \in S_n$? Si apra un ciclo con l'elemento $1 \in X$ e si considerino successivamente gli elementi

$$\sigma(1), \sigma(\sigma(1)) = \sigma^2(1), \sigma(\sigma(\sigma(1))) = \sigma^3(1), \dots$$

Si continui con questa procedura finché tali elementi sono $\neq 1$; appena si incontrerà l'elemento 1 si interrompa il procedimento, e si otterrà così il ciclo

$$\gamma_1 := (1, \sigma(1), \sigma(\sigma(1)), \sigma(\sigma(\sigma(1))), \dots)$$

Si passi ora ad aprire un altro ciclo a partire dal più piccolo naturale di X rimasto estraneo al ciclo γ_1 . Si ottiene nello stesso modo un nuovo ciclo γ_2 , disgiunto dal precedente. Dopo un numero finito di passi, tutti gli elementi di X si troveranno all'interno di un (unico) ciclo ed il procedimento termina. Ad esempio, se

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 2 & 1 & 3 & 4 & 7 & 5 & 6 \end{pmatrix} \in S_7,$$

si ottengono i cicli $(1\ 2)$, (3) , (4) , $(5\ 7\ 6)$. Si può verificare, eseguendo il prodotto di tali cicli, che

$$(1\ 2)(3)(4)(5\ 7\ 6) = (1\ 2)(5\ 7\ 6) = \sigma.$$

[si noti che gli 1-cicli (3) , (4) possono essere omessi dal prodotto, in quanto coincidono con $\mathbf{1}_X$]. Dunque σ è prodotto dei suoi cicli disgiunti di lunghezza ≥ 2 .

Proposizione 1. *Ogni permutazione non identica $\sigma \in S_n$ è prodotto dei suoi cicli disgiunti di lunghezza ≥ 2 .*

Dim. Sia $\sigma \neq \mathbf{1}_X$ e siano $\gamma_1, \dots, \gamma_t$ tutti i cicli (a due a due disgiunti) di σ di lunghezza ≥ 2 . Bisogna verificare che $\sigma = \gamma_1 \dots \gamma_t$, cioè che, $\forall x \in X$, risulta $\sigma(x) = (\gamma_1 \dots \gamma_t)(x)$.

Se x non appartiene ad alcuno dei cicli γ_i , allora $\sigma(x) = x = (\gamma_1 \dots \gamma_t)(x)$. Altrimenti, sia γ_i l'unico k -ciclo di σ contenente x . Allora:

$$\gamma_i = (x, \sigma(x), \sigma^2(x), \dots, \sigma^{k-1}(x)).$$

Si ha:

$$(\gamma_1 \dots \gamma_t)(x) = (\gamma_t \circ \dots \circ \gamma_1)(x) = (\gamma_t \circ \dots \circ \gamma_i)(x) = (\gamma_t \circ \dots \circ \gamma_{i+1})(\sigma(x)) = \dots = \sigma(x),$$

in quanto x non appartiene ai cicli $\gamma_1, \dots, \gamma_{i-1}$ e $\sigma(x)$ non appartiene ai cicli $\gamma_{i+1}, \dots, \gamma_t$ (in quanto appartiene a γ_i). Dunque è provato che $\sigma = \gamma_1 \dots \gamma_t$.

Osservazione 2. (i) Due cicli disgiunti di S_n commutano. Siano infatti $\gamma_1 = (c_1 c_2 \dots c_k)$ e $\gamma_2 = (d_1 d_2 \dots d_h)$ due cicli disgiunti di S_n . Si tratta di verificare che

$$(\gamma_1 \gamma_2)(x) = (\gamma_2 \gamma_1)(x), \quad \forall x \in X.$$

Se $x \in X - \{c_1, \dots, c_k, d_1, \dots, d_h\}$, allora $\gamma_1(x) = \gamma_2(x) = x$ e dunque $(\gamma_1 \gamma_2)(x) = (\gamma_2 \gamma_1)(x) = x$. Se invece $x \in \{c_1, \dots, c_k\}$, si ha:

$$(\gamma_1 \gamma_2)(x) = (\gamma_2 \circ \gamma_1)(x) = \gamma_2(\gamma_1(x)) = \gamma_1(x) \quad [\text{perché } \gamma_1(x) \notin \{d_1, \dots, d_h\}],$$

$$(\gamma_2 \gamma_1)(x) = (\gamma_1 \circ \gamma_2)(x) = \gamma_1(\gamma_2(x)) = \gamma_1(x) \quad [\text{perché } x \notin \{d_1, \dots, d_h\}].$$

Se infine $x \in \{d_1, \dots, d_h\}$, si ottiene ancora, con analoghe considerazioni, che $(\gamma_1 \gamma_2)(x) = (\gamma_2 \gamma_1)(x)$.

(ii) Ogni k -ciclo ($k \geq 2$) è sempre esprimibile come prodotto di $k-1$ trasposizioni non disgiunte. Infatti si verifica con calcolo diretto che:

$$(c_1 c_2 \dots c_k) = (c_1 c_2)(c_1 c_3) \dots (c_1 c_k).$$

Ne segue che ogni permutazione σ è esprimibile come prodotto di sole trasposizioni (a due a due non disgiunte).

(iii) Si verifica con calcolo diretto che l'inverso del k -ciclo $(c_1 c_2 \dots c_{k-1} c_k)$ è il k -ciclo $(c_k c_{k-1} \dots c_2 c_1)$. Infatti

$$(c_1 c_2 \dots c_{k-1} c_k)(c_k c_{k-1} \dots c_2 c_1) = (c_1)(c_2) \dots (c_{k-1})(c_k) = \mathbf{1}_X.$$

Dalla **Prop.1** e dall'**Osserv.2(ii)** segue che ogni permutazione è rappresentabile come prodotto di trasposizioni (a due a due non necessariamente disgiunte). Ma tale rappresentazione non è unica. Ad esempio il 4-ciclo $\sigma = (1 4 2 3) \in S_4$ si può indifferentemente scrivere nella forma $(1 2)(1 3)(2 4)$, ovvero $(1 4)(1 2)(1 3)$, ovvero $(1 2)(2 3)(1 3)(1 2)(2 4)$, ed in tanti altri modi.

C'è però una proprietà significativa da sottolineare [per la dimostrazione, che non è immediata, si rinvia ad esempio a www.mat.uniroma1.it/people/campbella, **Appunti di Algebra 1**, Cap. 4.3, pag. 153]. Eccola.

Proposizione 2. Assegnata una permutazione $\sigma \in S_n$, il numero delle trasposizioni di cui σ è prodotto è sempre pari o sempre dispari.

Ad esempio abbiamo appena visto che la permutazione $\sigma = (1 4 2 3) \in S_4$ è prodotto di tre oppure di cinque trasposizioni. La precedente proposizione ci dice che non avrebbe potuto essere prodotto di un numero pari di trasposizioni. Si noti poi che la permutazione identica $\mathbf{1}_x \in S_n$ ($n \geq 2$) è prodotto di un numero pari di trasposizioni. Infatti ad esempio $\mathbf{1}_x = (1 2)(1 2)$.

Definizione 2. Una permutazione $\sigma \in S_n$ è detta *di classe pari* se è esprimibile come prodotto di un numero pari di trasposizioni. Altrimenti è detta *di classe dispari*. L'insieme delle permutazioni di classe pari è denotato A_n .

Osservazione 3. (i) Se una permutazione $\sigma \in S_n$ è prodotto di t cicli $\gamma_1, \dots, \gamma_t$, aventi lunghezze rispettivamente k_1, \dots, k_t , la parità di σ è data dalla parità del numero $\sum_{i=1}^t (k_i - 1)$ [infatti γ_i è prodotto di $k_i - 1$ trasposizioni (cfr. **Osserv. 2(ii)**)].

(ii) Ci chiediamo come, assegnata una permutazione $\sigma = \begin{pmatrix} 1 & 2 & \dots & n \\ \sigma_1 & \sigma_2 & \dots & \sigma_n \end{pmatrix} \in S_n$, se ne possa calcolare la parità *direttamente*, senza cioè ricorrere alla sua scrittura come prodotto di cicli.

Per ogni $k = 1, \dots, n$, si pone

$$I(\sigma, k) := |\{\sigma_j : j > k \text{ e } \sigma_j < \sigma_k\}|.$$

[Tale numero è detto *k-simo numero delle inversioni di* σ ; si tratta della cardinalità dei σ_j minori di σ_k che seguono σ_k (nella seconda riga di σ)]. Ovviamente $I(\sigma, n) = 0$. Si pone poi:

$$I(\sigma) := I(\sigma, 1) + \dots + I(\sigma, n-1)$$

[detto *numero delle inversioni di* σ]. Si potrebbe verificare che:

$$\sigma \text{ è di classe pari} \iff I(\sigma) \text{ è un numero pari.}$$

Ad esempio, sia $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 6 & 1 & 5 & 3 & 4 & 2 \end{pmatrix} \in S_6$. Si ha:

$$I(\sigma, 1) = 5, I(\sigma, 2) = 0, I(\sigma, 3) = 3, I(\sigma, 4) = 1, I(\sigma, 5) = 1,$$

e quindi $I(\sigma) = 5 + 0 + 3 + 1 + 1 = 10$. Pertanto σ è di classe pari. In effetti risulta:

$$\sigma = (1\ 6\ 2)(3\ 5\ 4) = (1\ 6)(1\ 2)(3\ 5)(3\ 4).$$

Vogliamo ora elencare le 24 permutazioni di S_4 , scrivendole direttamente come prodotto dei loro cicli disgiunti.

È conveniente suddividerle rispetto a quella che viene chiamata *struttura ciclica*. Tra le permutazioni di S_4 ci potranno essere [oltre alla permutazione identica (1)] 2-cicli, 3-cicli, 4-cicli e coppie di 2-cicli disgiunti. Si ottengono le 24 permutazioni:

(1) [unico 1-ciclo (classe pari)];

$(1\ 2), (1\ 3), (1\ 4), (2\ 3), (2\ 4), (3\ 4)$ [2-cicli (classe dispari)];

$(1\ 2)(3\ 4), (1\ 3)(2\ 4), (1\ 4)(2\ 3)$ [coppie di 2-cicli disgiunti (classe pari)];

$\begin{cases} (1\ 2\ 3), (1\ 2\ 4), (1\ 3\ 4), (2\ 3\ 4) \\ (1\ 3\ 2), (1\ 4\ 2), (1\ 4\ 3), (2\ 4\ 3) \end{cases}$ [3-cicli (classe pari)];

$(1\ 2\ 3\ 4), (1\ 2\ 4\ 3), (1\ 3\ 2\ 4), (1\ 3\ 4\ 2), (1\ 4\ 2\ 3), (1\ 4\ 3\ 2)$ [4-cicli (classe dispari)].

[Si noti che per ottenere tutti i 4-cicli abbiamo fissato come primo elemento 1 e poi permutato gli altri tre elementi; per elencare i 3-cicli abbiamo scritto sulla prima riga tutte le sequenze crescenti di tre interi (da 1 a 4) e sotto, in corrispondenza di ciascun 3-ciclo, il suo inverso].

Osservazione 4. Il lettore attento avrà probabilmente osservato che in S_4 esistono 12 permutazioni di classe pari ed altrettante di classe dispari. Si sarà quindi chiesto se tale uguaglianza è un fatto casuale o no.

La risposta è che non si tratta di un fatto casuale. Dimostriamolo.

Se moltiplichiamo una permutazione σ per un 2-ciclo ne cambiamo la parità, cioè trasformiamo una permutazione pari in una dispari (e viceversa). Consideriamo allora la seguente applicazione

$$\varphi : S_n \rightarrow S_n \text{ tale che } \varphi(\sigma) = (1\ 2)\sigma, \quad \forall \sigma \in S_n.$$

L'applicazione φ è biiettiva. Infatti $\varphi \circ \varphi$ è l'applicazione identica su S_n [essendo $(\varphi \circ \varphi)(\sigma) = (1\ 2)((1\ 2)\sigma) = \sigma$] e dunque φ ammette inversa (se stessa).

Poiché φ trasforma l'insieme A_n delle permutazioni pari nell'insieme complementare $S_n - A_n$ delle permutazioni dispari (e viceversa), allora $|A_n| = |S_n - A_n|$. Il numero delle permutazioni pari (e

delle dispari) è $\frac{n!}{2}$.

ESERCIZI PROPOSTI

1.4.1. Sia $\sigma = (1\ 3\ 2\ 4)(5\ 6) \in \mathbf{S}_6$. Determinare il minimo intero $k \geq 1$ tale che $\sigma^k = \mathbf{1}_x$.

1.4.2. Stesso esercizio con $\sigma = (1\ 3\ 2)(4\ 5) \in \mathbf{S}_5$.

1.4.3. Scrivere tutte le permutazioni di \mathbf{S}_5 che contengono il ciclo $(1\ 2)$.

1.4.4. Assegnate le permutazioni $\sigma_1 = (1\ 3\ 4\ 2)$, $\sigma_2 = (2\ 5)(3\ 4) \in \mathbf{S}_5$, determinare la permutazione $\tau \in \mathbf{S}_5$ tale che $\sigma_2 = \tau \sigma_1$.

1.4.5. Scrivere la "tavola pitagorica" di \mathbf{S}_3 , cioè la tavola 6×6 formata da tutti i prodotti $\sigma_i \sigma_j$, al variare di σ_i, σ_j in \mathbf{S}_3 .

1.4.6. (i) Verificare che se $\sigma \in \mathbf{S}_n$ è una permutazione di classe dispari, non esiste alcuna permutazione $\alpha \in \mathbf{S}_n$ tale che $\alpha^2 = \sigma$.

(ii) Determinare $\alpha \in \mathbf{S}_6$ tale che $\alpha^2 = (1\ 2\ 3)(4\ 5\ 6)$.

(iii) Spiegare perché non esiste $\alpha \in \mathbf{S}_6$ tale che $\alpha^2 = (1\ 2)(3\ 4\ 5\ 6)$.

1.4.7. (i) Verificare che $\forall \sigma \in \mathbf{S}_n \exists k \in \mathbf{N}, k \geq 1$ tale che $\sigma^k = \mathbf{1}_x$.

(ii) Verificare che $\forall \sigma, \tau \in \mathbf{S}_n$ l'equazione (di primo grado) $\sigma X = \tau$ ammette una ed una sola soluzione (in \mathbf{S}_n).

1.4.8. Determinare per quali $\sigma \in \mathbf{S}_4$ l'equazione $X^2 = \sigma$ è risolubile.

1.4.9. In \mathbf{S}_5 sono assegnati un 3-ciclo σ ed un 2-ciclo τ , disgiunti tra loro. Sia \mathbf{H} l'insieme formato dalle permutazioni di \mathbf{S}_5 ottenibili come prodotti finiti di σ e di τ . Determinare le permutazioni di \mathbf{H} e verificare se \mathbf{H} è un sottogruppo di \mathbf{S}_5 .

1.4.10. (i) Quanti sono i 3-cicli di \mathbf{S}_6 ?

(ii) Quante sono le permutazioni di \mathbf{S}_6 che sono prodotto di due 3-cicli disgiunti?

Capitolo 2

STRUTTURE ALGEBRICHE

1. Gruppi, anelli, campi e spazi vettoriali

In questo paragrafo torneremo sul concetto di "operazione su un insieme" e presenteremo la definizione e le proprietà di base di alcune importanti *strutture algebriche* [cioè insiemi con una o più operazioni], fornendone poi qualche esempio.

Come già detto nel capitolo precedente, un'operazione su un insieme A è un'applicazione che in generale denoteremo $\ast : A \times A \rightarrow A$. Per adeguarci all'uso comune, scrivereemo poi, $\forall (a, b) \in A \times A$, $a \ast b$ in luogo di $\ast((a, b))$.

Definizione 1. Un'operazione \ast su A è:

- (1) *associativa* se $(a \ast b) \ast c = a \ast (b \ast c)$, $\forall a, b, c \in A$;
- (2) *dotata di elemento neutro* e se $\exists e \in A$ tale che $a \ast e = a = e \ast a$, $\forall a \in A$;
- (3) *dotata di reciproco di ogni elemento* se $\forall a \in A$, $\exists a' \in A$ tale che $a \ast a' = e = a' \ast a$;
- (4) *commutativa* se $a \ast b = b \ast a$, $\forall a, b \in A$.

Esempi 1. (i) L'addizione $+$ su \mathbf{N} verifica le proprietà (1), (2), (4) ma non (3) [l'elemento neutro è 0; non esiste il reciproco di alcun $n \geq 1$].

(ii) La moltiplicazione \cdot su \mathbf{N} verifica le proprietà (1), (2), (4) ma non (3) [l'elemento neutro è 1; non esiste il reciproco di alcun $n \geq 2$].

(iii) L'addizione $+$ su \mathbf{Z} (e su \mathbf{Q} ed \mathbf{R}) verifica le proprietà (1), (2), (3) e (4) [l'elemento neutro è 0; il reciproco di n è $-n$].

(iv) La moltiplicazione \cdot su \mathbf{Z} verifica le proprietà (1), (2), (4) ma non (3) [l'elemento neutro è 1; non esiste il reciproco di alcun $n \neq \pm 1$].

(v) La moltiplicazione \cdot su $\mathbf{Q}^{\circ} := \mathbf{Q} - \{0\}$ e su $\mathbf{R}^{\circ} := \mathbf{R} - \{0\}$ verifica le proprietà (1), (2), (3) e (4) [l'elemento neutro è 1; il reciproco di ogni $x \in \mathbf{Q}^{\circ}$ (o \mathbf{R}°) è $\frac{1}{x}$].

Nota. Come osservato, negli esempi precedenti relativi all'addizione, l'elemento neutro è 0 ed il reciproco è detto *opposto*; relativamente alla moltiplicazione, l'elemento neutro è 1 ed il reciproco è detto *inverso*.

Converremo nel seguito che, per ognuno degli insiemi numerici $A = \mathbf{N}, \mathbf{Z}, \mathbf{Q}, \mathbf{R}$, con A° si intende l'insieme $A - \{0\}$.

(vi) Per ogni insieme non vuoto X , il prodotto operatorio \circ sull'insieme $\mathcal{S}(X)$ [delle biiezioni di X in sé] verifica le proprietà (1), (2) e (3). L'elemento neutro è l'applicazione identica $\mathbf{1}_x$ ed il reciproco di una biiezione f è la biiezione f^{-1} . In base a quanto visto nel precedente capitolo, se X ha almeno tre elementi, il prodotto operatorio \circ non verifica la proprietà (4).

Definizione 2. Si chiama *gruppo* ogni coppia (A, \ast) tale che A è un insieme non vuoto (detto *insieme sostegno del gruppo*) ed \ast è un'operazione su A verificante le proprietà (1), (2), (3), cioè associativa, dotata di elemento neutro e dotata di reciproco di ogni elemento. Un gruppo (A, \ast) è detto *commutativo* (o *abeliano*) se \ast verifica (4), cioè è commutativa. Infine, la cardinalità $|A|$ dell'insieme sostegno A è detta *ordine del gruppo* (A, \ast) .

Abbiamo dunque i gruppi commutativi $(\mathbf{Z}, +)$, $(\mathbf{Q}, +)$, $(\mathbf{R}, +)$, $(\mathbf{Q}^{\circ}, \cdot)$, $(\mathbf{R}^{\circ}, \cdot)$ ed i gruppi non commutativi $(\mathcal{S}(X), \circ)$, $\forall X$ con $|X| \geq 3$. Ricordiamo che se $|X| = n$, $\mathcal{S}(X)$ è stato denotato \mathcal{S}_n . Il

gruppo (S_n, \circ) è detto *gruppo simmetrico su n elementi*. Si noti che $(N, +)$, (Z, \cdot) , (Q, \cdot) , (R, \cdot) non sono gruppi [in quanto non tutti i loro elementi ammettono reciproco].

Vedremo a breve altri esempi di gruppi.

Proposizione 1. In ogni gruppo $(A, *)$:

- (1) L'elemento neutro è unico.
- (2) Il reciproco di ogni elemento è unico.
- (3) Vale la legge di cancellazione $\begin{cases} \text{a sinistra: } a * b = a * c \Rightarrow b = c \\ \text{a destra: } a * b = c * b \Rightarrow a = c. \end{cases}$
- (4) Il reciproco di un prodotto è il prodotto dei reciproci, in ordine inverso.

Dim. (1) Siano e, e' due elementi neutri di $(A, *)$. Allora

$$\begin{cases} e * e' = e' * e = e' & \text{essendo } e \text{ elemento neutro,} \\ e * e' = e' * e = e & \text{essendo } e' \text{ elemento neutro.} \end{cases}$$

Dunque $e = e'$.

(2) Siano a', a'' due reciproci di a . Allora $\begin{cases} a * a' = a' * a = e \\ a * a'' = a'' * a = e. \end{cases}$ Dunque, utilizzando tali uguaglianze, la proprietà associativa ed il fatto che e è elemento neutro:

$$a' = a' * e = a' * (a * a'') = (a' * a) * a'' = e * a'' = a''.$$

(3) Verifichiamo la legge di cancellazione a sinistra [per quella a destra si procede in modo analogo]. Moltiplicando l'uguaglianza $a * b = a * c$ a sinistra per a' [reciproco di a], si ottiene:

$$a' * (a * b) = a' * (a * c) \Rightarrow (a' * a) * b = (a' * a) * c \Rightarrow e * b = e * c \Rightarrow b = c.$$

(4) Dimostriamo che $(a * b)' = b' * a'$, $\forall a, b \in A$. Per l'unicità del reciproco, basta verificare che:

$$(a * b) * (b' * a') = e = (b' * a') * (a * b).$$

Verifichiamo la prima uguaglianza [per l'altra si procede in modo analogo]. Si ha:

$$(a * b) * (b' * a') = ((a * b) * b') * a' = (a * (b * b')) * a' = (a * e) * a' = a * a' = e.$$

In modo analogo si verifica che $(a_1 * a_2 * \dots * a_n)' = a'_n * \dots * a'_2 * a'_1$, $\forall a_1, a_2, \dots, a_n \in A$.

Definizione 3. Si chiama sottogruppo di un gruppo $A = (A, *)$ ogni sottoinsieme non vuoto $B \subseteq A$ tale che $(B, *)$ è un gruppo (rispetto alla stessa operazione $*$ di A , ovviamente ristretta agli elementi di B). Per indicare che B è un sottogruppo di A scriveremo $B \leq A$ (anziché $B \subseteq A$).

Si verifica subito che A ed $\{e\}$ sono sottogruppi del gruppo A , detti sottogruppi banali di A . Gli altri sottogruppi (se ne esistono) sono detti sottogruppi propri di A .

Esempi 2. (i) Ad esempio, Z è un sottogruppo di $(Q, +)$, Z e Q sono sottogruppi di $(R, +)$. Inoltre Q^+ è un sottogruppo di (R, \cdot) . Indicato con Q^+ l'insieme dei razionali positivi, si osserva subito che Q^+ è un sottogruppo di (Q, \cdot) [si noti che il prodotto di razionali positivi è positivo e che l'inverso di un razionale positivo è positivo]. Analogamente, R^+ è un sottogruppo di (R, \cdot) .

(ii) In S_n consideriamo il sottoinsieme Σ formato dalle sole permutazioni di $X = \{1, 2, \dots, n\}$ che fissano un dato elemento, ad esempio l'elemento $1 \in X$. Si può subito verificare che Σ verifica le proprietà (1), (2), (3) della definizione di gruppo. Dunque Σ è un sottogruppo di S_n .

Un altro importante sottogruppo di S_n è il sottogruppo A_n delle permutazioni di classe pari, detto *gruppo alterno su n elementi*.

Osservazione 1. Nello studio "astratto" dei gruppi si usa per lo più la *notazione moltiplicativa*. Un gruppo viene tradizionalmente indicato con (G, \cdot) , il suo elemento neutro con 1 (o 1_G) (ed è detto *unità di G*) ed il reciproco di un elemento $g \in G$ con g^{-1} (ed è detto *inverso di g*).

Talvolta però viene anche usata la *notazione additiva* $(G, +)$ [e ciò avviene soprattutto nello studio dei gruppi commutativi]. In tal caso l'elemento neutro si indica con 0 (o 0_G) (ed è detto *zero di G*).

ed il reciproco di un elemento $g \in G$ con $-g$ (ed è detto *opposto di g*).

Spesso in uno stesso insieme coesistono due (o più) operazioni. Ad esempio, in $\mathbf{Z}, \mathbf{Q}, \mathbf{R}$ (ed anche in \mathbf{N}) sono definite sia l'addizione che la moltiplicazione. Le due operazioni non sono indipendenti, ma sono legate dalle leggi distributive. Questa situazione giustifica l'introduzione di una più ricca struttura algebrica, quella di *anello*. Si tratta di un insieme dotato di due operazioni che, per semplificare le notazioni, denoteremo con $+$ e \cdot .

Definizione 4. Si chiama *anello* ogni terna $(A, +, \cdot)$ tale che: A è un insieme non vuoto (detto *sostegno dell'anello*), $+$ e \cdot sono due operazioni su A (dette *somma* e *prodotto* di A), verificanti i seguenti assiomi:

- $(A, +)$ è un gruppo commutativo [con elemento neutro $0 = 0_A$];
- il prodotto \cdot è associativo: $(ab)c = a(bc)$, $\forall a, b, c \in A$;
- valgono le due leggi distributive tra somma e prodotto:

$$a(b+c) = ab + ac, \quad (a+b)c = ac + bc, \quad \forall a, b, c \in A.$$

Definizione 5. Un anello $(A, +, \cdot)$ è detto *anello unitario* se il prodotto \cdot ha elemento neutro (detto *unità di A* e denotato 1 o 1_A), cioè $a \cdot 1 = 1 \cdot a = a$, $\forall a \in A$. Si noti che l'unità, se esiste, è unica (cfr. Prop. 1(1)).

Un anello $(A, +, \cdot)$ è detto *anello commutativo* se il prodotto \cdot è commutativo, cioè se $ab = ba$, $\forall a, b \in A$. Un anello $(A, +, \cdot)$ è detto *campo* se (A^*, \cdot) è un gruppo commutativo [ovviamente $A^* := A - \{0\}$]. Dunque un campo è un anello commutativo unitario tale che ogni $a \in A^*$ ammette inverso $a^{-1} \in A$. I campi sono spesso denotati con la lettera K (o lettere contigue).

Osservazione 2. (i) $(\mathbf{Z}, +, \cdot)$ è un anello commutativo unitario [abbr. *c.u.*], ma non è un campo. Infatti (\mathbf{Z}^*, \cdot) non è un gruppo [soltanto $1, -1$ ammettono inverso in \mathbf{Z}]. Invece $(\mathbf{Q}, +, \cdot)$ e $(\mathbf{R}, +, \cdot)$ sono campi.

(ii) In ogni anello $(A, +, \cdot)$ risulta:

$$a \cdot 0 = 0 \cdot a = 0, \quad \forall a \in A.$$

Infatti $a \cdot 0 = a(0+0) = a \cdot 0 + a \cdot 0$. Dunque $a \cdot 0 + 0 = a \cdot 0 + a \cdot 0$ e, dalla legge di cancellazione (per la somma), segue che $a \cdot 0 = 0$. Analogamente si verifica che $0 \cdot a = 0$.

(iii) In ogni anello $(A, +, \cdot)$ valgono le tre seguenti regole di calcolo:

$$a(-b) = -(ab) = (-a)b; \quad (-a)(-b) = ab; \quad a(b-c) = ab - ac, \quad \forall a, b, c \in A.$$

Per verificare la prima regola, basta osservare che

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

Per la seconda [applicando la prima]: $(-a)(-b) = -((a)(-b)) = -(-(ab)) = ab$. Per la terza infine [tenuto conto che si pone, per definizione: $x - y := x + (-y)$], si ha:

$$a(b-c) = a(b + (-c)) = ab + a(-c) = ab + (-ac) = ab - ac.$$

(iv) Tutti gli anelli sinora considerati sono commutativi; nel seguito ne troveremo anche di non commutativi. Negli anelli finora considerati risulta:

$$ab = 0 \implies a = 0 \text{ o } b = 0.$$

Ma esistono anche anelli in cui la condizione $ab = 0$ non implica necessariamente $a = 0$ o $b = 0$. Tali anelli sono detti *anelli non integri*. Quelli verificanti la condizione scritta sopra sono invece detti *anelli integri*. Un anello c.u. ed integro è detto *dominio* o *dominio d'integrità*. Ad esempio $(\mathbf{Z}, +, \cdot)$ è un dominio d'integrità. Anche i campi sono domini d'integrità [se infatti $ab = 0$ e $a \neq 0$, allora $b = 1_A b = (a^{-1}a)b = a^{-1}(ab) = a^{-1}0 = 0$]. Scopriremo in seguito esempi di anelli non integri, non unitari e non commutativi.

(v) In un anello unitario $(A, +, \cdot)$ diremo che un elemento a è *invertibile* se esiste $b \in A$ tale che $ab = ba = 1$. Tale elemento è unico [cfr. Prop. 1(2)] ed è chiamato *inverso di a* (e denotato

a^{-1}). Ad esempio in $(\mathbf{Z}, +, \cdot)$ gli elementi invertibili sono due: $1, -1$; in un campo $(K, +, \cdot)$ tutti gli elementi non nulli sono invertibili.

Se denotiamo con $\mathcal{U}(A)$ l'insieme degli elementi invertibili di A , si verifica subito che $(\mathcal{U}(A), \cdot)$ è un gruppo. Se infatti $a_1, a_2 \in \mathcal{U}(A)$ [con inversi rispettivamente b_1, b_2] allora anche $a_1 a_2 \in \mathcal{U}(A)$ [con inverso $b_2 b_1$]; ovviamente $1 \in \mathcal{U}(A)$ e se $a \in \mathcal{U}(A)$ anche il suo inverso $a^{-1} \in \mathcal{U}(A)$.

Definizione 6. Si chiama *sottonello* di un anello $(A, +, \cdot)$ ogni sottoinsieme non vuoto $B \subseteq A$ tale che $(B, +, \cdot)$ è un anello (rispetto alle stesse operazioni $+, \cdot$ di A , ovviamente ristrette agli elementi di B). Se A e B sono campi, si dirà che B è un *sottocampo* di A .

Si noti infine che $\{0\}$ ed A sono sottoanelli di A , detti *sottoanelli banali* di A . Gli altri sottoanelli di A (se ne esistono) sono detti *sottoanelli propri* di A .

Ad esempio \mathbf{Z} è un sottoanello di $(\mathbf{Q}, +, \cdot)$, che a sua volta è un sottocampo di $(\mathbf{R}, +, \cdot)$. L'insieme \mathbf{P} degli interi pari è un sottoanello (non unitario) di $(\mathbf{Z}, +, \cdot)$.

Un'altra struttura algebrica di cui ci occuperemo molto diffusamente nel corso è quella di *spazio vettoriale su un campo* K . Il termine spazio vettoriale fa riferimento al ben noto concetto di "vettore".

Denotiamo con \mathcal{V} l'insieme dei vettori (di un piano o dello spazio ordinario). Lo studente ha imparato dalla fisica a sommare due vettori e a moltiplicare un vettore per un numero reale.

La somma di vettori [che si esegue con la nota "regola del parallelogramma"] assegna a \mathcal{V} struttura di gruppo commutativo.

La moltiplicazione di un vettore per un numero reale [che "allunga" o "accorcia" un vettore ed eventualmente lo cambia di verso (ma non di direzione)] è un'applicazione $\mathbf{R} \times \mathcal{V} \rightarrow \mathcal{V}$ [usualmente chiamata "moltiplicazione di un vettore per uno scalare"]. Tale applicazione non è un'operazione, nel senso della definizione data sopra (ma spesso ci si riferisce ad essa col nome di "operazione esterna"). È legata alla somma da importanti relazioni, quali ad esempio $c(\underline{u} + \underline{v}) = c\underline{u} + c\underline{v}$, $\forall c \in \mathbf{R}$, $\forall \underline{u}, \underline{v} \in \mathcal{V}$.

Lo studente riconoscerà che le principali relazioni tra queste due operazioni tra vettori sono gli assiomi della seguente definizione di K -spazio vettoriale (astratto).

Definizione 7. Sia $(K, +, \cdot)$ un campo. Un insieme non vuoto V è detto *K -spazio vettoriale* se è dotato di un'operazione $+$ [detta *somma*], rispetto a cui $(V, +)$ è un gruppo commutativo e se è definita un'applicazione $K \times V \rightarrow V$ [detta *moltiplicazione per uno scalare*], tale che:

- $(c + d)\underline{v} = c\underline{v} + d\underline{v}$, $\forall c, d \in K$, $\forall \underline{v} \in V$;
- $c(\underline{v}_1 + \underline{v}_2) = c\underline{v}_1 + c\underline{v}_2$, $\forall c \in K$, $\forall \underline{v}_1, \underline{v}_2 \in V$;
- $(cd)\underline{v} = c(d\underline{v})$, $\forall c, d \in K$, $\forall \underline{v} \in V$;
- $1 \cdot \underline{v} = \underline{v}$, $\forall \underline{v} \in V$.

Gli elementi di V sono detti *vettori* mentre gli elementi di K sono detti *scalari* [seguendo una consolidata abitudine della fisica, denoteremo spesso i vettori con lettere sottolineate].

Veniamo ad un importante esempio di K -spazio vettoriale. Denotiamo con K^n il prodotto cartesiano $K \times K \times \dots \times K$ di n copie di K . I suoi elementi sono del tipo (c_1, c_2, \dots, c_n) , con $c_1, c_2, \dots, c_n \in K$. Come già fatto nel **Cap. 1**, tali elementi sono detti *n-ple di elementi di K* . Talvolta, per brevità, scriveremo \underline{c} in luogo di (c_1, c_2, \dots, c_n) .

Su K^n è definita la seguente operazione $+ : K^n \times K^n \rightarrow K^n$:

$$(c_1, c_2, \dots, c_n) + (d_1, d_2, \dots, d_n) = (c_1 + d_1, c_2 + d_2, \dots, c_n + d_n), \quad \forall (c_1, c_2, \dots, c_n), (d_1, d_2, \dots, d_n) \in K^n.$$

Tale operazione è detta *somma (componente per componente) di n-ple*. Si verifica con facilità che $(K^n, +)$ è un gruppo commutativo. L'elemento neutro è la *n-pla nulla* $(0, 0, \dots, 0)$. Il reciproco della *n-pla* $\underline{c} = (c_1, c_2, \dots, c_n)$ è la *n-pla* $-\underline{c} = (-c_1, -c_2, \dots, -c_n)$, detta *n-pla opposta di \underline{c}* .

Su K^n è inoltre definita la seguente *moltiplicazione per uno scalare* $\cdot : K \times K^n \rightarrow K^n$ tale che

$$a(c_1, c_2, \dots, c_n) = (a c_1, a c_2, \dots, a c_n), \quad \forall a \in K, \quad \forall (c_1, c_2, \dots, c_n) \in K^n.$$

Lasciamo allo studente la semplice verifica dei quattro assiomi di tale operazione esterna (dalla precedente definizione di spazio vettoriale). Concludiamo così che l'insieme K^n delle *n-ple* è un

K-spazio vettoriale.

Se $n = 1$, le 1-ple si identificano con gli elementi di K e la moltiplicazione per uno scalare con l'usuale moltiplicazione del campo K . Dunque un campo K è anche un *K*-spazio vettoriale.

Osservazione 3. In K^n possiamo definire anche il *prodotto (componente per componente)* di n -ple, ponendo

$$(c_1, c_2, \dots, c_n) \cdot (d_1, d_2, \dots, d_n) = (c_1 d_1, c_2 d_2, \dots, c_n d_n), \quad \forall (c_1, c_2, \dots, c_n), (d_1, d_2, \dots, d_n) \in K^n.$$

Si verifica con facilità che $(K^n, +, \cdot)$ è un anello commutativo unitario [con *unità* $\underline{1} = (1, 1, \dots, 1)$]. Se $n \geq 2$, tale anello non è integro [ad esempio $(0, 1, 0, \dots, 0) \cdot (1, 0, 0, \dots, 0) = (0, 0, \dots, 0)$].

Nel prossimo paragrafo studieremo l'insieme delle *matrici a valori su un campo K*: si tratta di un altro esempio di *K*-spazio vettoriale.

Veniamo ora alla definizione di *sottospazio vettoriale*.

Definizione 8. Sia V un *K*-spazio vettoriale. Un sottoinsieme non vuoto $W \subseteq V$ è detto *K-sottospazio vettoriale* di V se W è un *K*-spazio vettoriale (rispetto alle stesse operazioni di V , opportunamente ristrette agli elementi di W).

Si verifica subito che V ed $\{\underline{0}\}$ sono *K*-sottospazi vettoriali di V , detti *sottospazi vettoriali banali*. Altri esempi di sottospazi vettoriali saranno visti in seguito.

Osservazione 4. Potrà capitarc ci nel seguito di indicare la moltiplicazione per uno scalare scrivendo lo scalare a destra, invece che a sinistra, cioè di considerare il vettore $\underline{v}c$ in luogo di $c\underline{v}$. È evidente che si tratta di un'imprecisione, ma tale imprecisione non comporta conseguenze fatali. Ciò dipende dal fatto che K è commutativo.

Ad esempio, se prima eseguiamo il prodotto di \underline{v} per lo scalare b e poi moltiplichiamo il vettore ottenuto per a , otteniamo "a sinistra" il vettore $a(\underline{b}\underline{v})$ ed "a destra" il vettore $(\underline{v}b)a$. D'altra parte $a(\underline{b}\underline{v}) = (\underline{a}\underline{b})\underline{v}$, mentre $(\underline{v}b)a = \underline{v}(ba)$. Poiché $ba = ab$, allora $(\underline{v}b)a$ si può sostituire con $a(\underline{b}\underline{v})$.

Concludiamo il paragrafo presentando un importante esempio di struttura algebrica: l'*anello dei polinomi in una indeterminata ed a coefficienti in un campo K*.

Si chiama *polinomio P nell'indeterminata X ed a coefficienti in un campo K* ogni espressione formale del tipo

$$P = P(X) = a_0 + a_1 X + a_2 X^2 + \dots + a_n X^n \quad [= \sum_{i=0}^n a_i X^i],$$

con $n \in \mathbf{N}$, $a_0, \dots, a_n \in K$ (detti *coefficienti di P*) e X *indeterminata* (o *incognita*) (su K): si tratta di un simbolo soggetto soltanto alle seguenti regole di calcolo:

$$0X = 0, \quad 1X = X, \quad X^0 = 1, \quad X^1 = X, \quad X^h X^k = X^{h+k}, \quad \forall h, k \in \mathbf{N}.$$

Si noti che tra i polinomi ci sono anche gli elementi di K , detti *(polinomi) costanti*. Tra questi in particolare ci sono il *polinomio nullo* 0 ed il *polinomio unità* 1.

Ad ogni polinomio non nullo P è assegnato un *grado*, denotato $\deg(P)$: si tratta dell'esponente massimo di X , tra i vari addendi non nulli di P . Ai polinomi costanti $c \in K$ viene perciò attribuito grado 0, mentre al polinomio nullo 0 non può essere assegnato alcun grado [perché non ci sono addendi non nulli di 0] e si conviene allora di attribuirgli grado $-\infty$.

L'insieme di tutti i polinomi in X a coefficienti in K viene denotato $K[X]$.

È noto allo studente come due polinomi si sommino e si moltiplichino tra loro. Presi comunque $P = \sum_{i=0}^n a_i X^i$, $Q = \sum_{j=0}^m b_j X^j \in K[X]$ e supposto ad esempio, per fissare le idee, che sia $n \leq m$, si pone:

$$P + Q = (a_0 + b_0) + (a_1 + b_1)X + \dots + (a_n + b_n)X^n + b_{n+1}X^{n+1} + \dots + b_m X^m.$$

$$PQ = a_0 b_0 + (a_0 b_1 + a_1 b_0)X + (a_0 b_2 + a_1 b_1 + a_2 b_0)X^2 + \dots + a_n b_m X^{n+m}.$$

Si può facilmente verificare che $(K[X], +)$ è un gruppo commutativo [con elemento neutro il polinomio nullo 0 e con opposto di P il polinomio $-P = \sum_{i=0}^n (-a_i)X^i$].

Si potrebbe poi verificare che il prodotto tra polinomi è associativo [cioè che $(PQ)R = P(QR)$, $\forall P, Q, R \in K[X]$] e che valgono le leggi distributive tra somma e prodotto. Ne segue che $(K[X], +, \cdot)$ è un anello.

Inoltre è facile vedere che il prodotto è commutativo ed ha unità (il polinomio 1). Infine si verifica che $K[X]$ è un anello integro, cioè tale che $PQ = 0 \implies P = 0$ o $Q = 0$.

Verifichiamo quest'ultimo fatto. Se per assurdo P, Q fossero entrambi non nulli, avrebbero grado ≥ 0 . Poiché il grado del prodotto di due polinomi non nulli è la somma dei gradi dei due fattori [ciò che segue subito dalla definizione di prodotto], allora $\deg(PQ) = \deg(0) \geq 0$: assurdo.

Da quanto precede si ha quindi che $(K[X], +, \cdot)$ è un dominio d'integrità. Lasciamo infine allo studente il compito di verificare che $K[X]$ è anche un K -spazio vettoriale, con moltiplicazione per uno scalare così definita:

$$cP = \sum_{i=0}^n ca_i X^i, \quad \forall c \in K, \quad \forall P = \sum_{i=0}^n a_i X^i \in K[X].$$

N.B. È evidente che è possibile considerare polinomi a coefficienti in una struttura algebrica meno "perfetta" di un campo, ad esempio su un anello A . In tal caso $A[X]$ è un anello, ma in generale non è un dominio d'integrità.

ESERCIZI PROPOSTI

2.1.1. Sia (G, \cdot) un gruppo e siano $a, b \in G$. Verificare che, se $ab = 1_G$, allora $ba = 1_G$ [cioè se b è "inverso a destra" di a , è anche "inverso a sinistra" di a].

2.1.2. Si consideri in S_5 il sottoinsieme

$$\Sigma = \Sigma_{\{1, 2\}} := \{\sigma \in S_5 : \sigma(\{1, 2\}) = \{1, 2\}\}.$$

Verificare che Σ è un sottogruppo di S_5 e scriverne gli elementi.

2.1.3. Sia K un campo ed n un intero ≥ 2 . In K^n si consideri il sottoinsieme

$$W = \{(c_1, c_2, \dots, c_n) \in K^n : c_1 = 0\}.$$

Verificare se W è un sottospazio vettoriale di K^n e se è un sottoanello di K^n .

2. Matrici a valori su un campo

In questo e nel successivo paragrafo presenteremo due importanti esempi di strutture algebriche, di cui ci occuperemo largamente nel corso.

Definizione 1. Sia K un campo e siano m, n due interi positivi. Si chiama matrice a valori in K ad m righe ed n colonne [cioè di tipo (m, n)] ogni insieme ordinato A di mn elementi di K , disposti su m righe ed n colonne, che indicheremo genericamente nella forma:

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

[dove il primo indice di ogni elemento è detto indice di riga ed il secondo indice di colonna]. L'insieme delle matrici (a valori in K) ad m righe ed n colonne verrà indicato con $\mathfrak{M}_{m,n}(K)$. Ovviamente $\mathfrak{M}_{1,1}(K)$ è identificabile con K .

Studieremo in questo paragrafo la struttura algebrica dell'insieme $\mathfrak{M}_{m,n}(K)$.

Sia $A \in \mathfrak{M}_{m,n}(K)$. Per abbreviare le notazioni, scriveremo talvolta $A = (a_{ij})$. L'elemento a_{ij} [situato sulla i -esima riga e sulla j -esima colonna di A (per $i = 1, \dots, m$, $j = 1, \dots, n$)] sarà talvolta denotato anche $(A)_{ij}$. Inoltre indicheremo con $A^{(i)}$ la i -esima riga $(a_{i1} a_{i2} \dots a_{in})$ di A

e con $A_{(j)}$ la j -esima colonna $\begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix}$ di A . Pertanto:

$$A = (A_{(1)} A_{(2)} \dots A_{(n)}) = \begin{pmatrix} A^{(1)} \\ A^{(2)} \\ \vdots \\ A^{(m)} \end{pmatrix}.$$

Dunque la matrice A è la "riga delle sue n colonne" ed è la "colonna delle sue m righe".

N.B. Aver detto che una matrice A è la "riga delle sue colonne" contrasta con il fatto che abbiamo definito soltanto matrici a valori in un campo. Ma è evidente che una matrice è semplicemente un "contenitore" di oggetti. Nulla ci vieta di considerare matrici a valori in un insieme (o in una struttura algebrica) più generale. Dunque, dire che A è la "riga delle sue n colonne" significa interpretare A in $\mathfrak{M}_{1,n}(\mathfrak{M}_{m,1}(K))$. Analogamente, dire che A è la "colonna delle sue m righe" significa interpretare A in $\mathfrak{M}_{m,1}(\mathfrak{M}_{1,n}(K))$.

Osservazione 1. Per ogni intero $n \geq 1$, consideriamo il prodotto cartesiano K^n di n copie di K . Ovviamente esiste una corrispondenza biunivoca tra gli insiemi K^n , $\mathfrak{M}_{1,n}(K)$ e $\mathfrak{M}_{n,1}(K)$.

Si vedrà nel seguito che è conveniente identificare K^n con $\mathfrak{M}_{n,1}(K)$, cioè identificare la n -pla $\underline{a} = (a_1, \dots, a_n) \in K^n$ con la corrispondente "matrice colonna" $\mathbf{a} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \in \mathfrak{M}_{n,1}(K)$. Denoteremo le matrici colonna con lettere in grassetto e le n -ple con lettere sottolineate; tuttavia alcune volte (soprattutto nell'ultima parte del corso) per non appesantire le notazioni indicheremo matrici colonna ed n -ple nello stesso modo.

Definizione 2. Sia $A \in \mathfrak{M}_{m,n}(K)$. Si chiama matrice trasposta di A la matrice $B \in \mathfrak{M}_{n,m}(K)$ così definita:

$$b_{ij} = a_{ji}, \quad \forall i = 1, \dots, n, \quad \forall j = 1, \dots, m.$$

La matrice B verrà denotata con ${}^t A$. Ovviamente ${}^t({}^t A) = A$.

Si noti che la i -sima riga [risp. colonna] della trasposta di A coincide con la trasposta della i -sima colonna [risp. riga] di A , cioè:

$$({}^t A)^{(i)} = {}^t(A_{(i)}) \quad \text{e} \quad ({}^t A)_{(j)} = {}^t(A^{(j)}).$$

Infatti:

$$({}^t A)^{(i)} = (b_{i1} \ b_{i2} \dots b_{im}) = (a_{1i} \ a_{2i} \dots a_{ni}) = \begin{pmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{ni} \end{pmatrix} = {}^t(A_{(i)}).$$

In modo analogo si verifica l'altra uguaglianza.

Definizione 3. Una matrice $A \in \mathfrak{M}_{m,n}(K)$ è detta *matrice quadrata* (di ordine n) se $m = n$. La n -pla $(a_{11}, a_{22}, \dots, a_{nn})$ è detta *diagonale* di A . L'insieme $\mathfrak{M}_{n,n}(K)$ verrà denotato, più semplicemente, con $\mathfrak{M}_n(K)$.

Una matrice quadrata $A = (a_{ij}) \in \mathfrak{M}_n(K)$ è detta *triangolare superiore* [rispett. *inferiore*] se $a_{ij} = 0$, $\forall i > j$ [rispett. $a_{ij} = 0$, $\forall i < j$].

Una matrice quadrata $A \in \mathfrak{M}_n(K)$ è detta *matrice diagonale* se è triangolare superiore e triangolare inferiore [cioè se $a_{ij} = 0$, $\forall i \neq j$].

Una matrice diagonale A è detta *matrice scalare* se risulta: $a_{11} = a_{22} = \dots = a_{nn}$. Una particolare matrice scalare è la *matrice unità* I_n , tale che $a_{11} = \dots = a_{nn} = 1$.

Infine, una matrice quadrata $A \in \mathfrak{M}_n(K)$ è detta *simmetrica* se $A = {}^t A$ [cioè se $a_{ij} = a_{ji}$] ed è detta *antisimmetrica* se $A = -({}^t A)$ [cioè se $a_{ij} = -a_{ji}$].

Definizione 4. In $\mathfrak{M}_{m,n}(K)$ sono definite la somma di matrici ed il prodotto di una matrice per uno scalare:

$$+ : \mathfrak{M}_{m,n}(K) \times \mathfrak{M}_{m,n}(K) \rightarrow \mathfrak{M}_{m,n}(K) \quad \text{tale che:}$$

$$(A, B) \rightarrow A + B, \quad \text{con} \quad (A + B)_{ij} = (A)_{ij} + (B)_{ij};$$

$$\cdot : K \times \mathfrak{M}_{m,n}(K) \rightarrow \mathfrak{M}_{m,n}(K) \quad \text{tale che:}$$

$$(c, A) \rightarrow cA, \quad \text{con} \quad (cA)_{ij} = c(A)_{ij}.$$

L'insieme $\mathfrak{M}_{m,n}(K)$, dotato delle operazioni sopra definite, è un K -spazio vettoriale. In particolare, l'elemento neutro della somma è la *matrice nulla* $\mathbf{0}$ [tale che $\mathbf{0}_{ij} = 0$, $\forall i, j$]; l'opposto della matrice A è la matrice $-A$ [tale che $(-A)_{ij} = -(A)_{ij}$, $\forall i, j$]. Le verifiche degli assiomi di K -spazio vettoriale sono lasciate al lettore.

L'insieme delle matrici quadrate $\mathfrak{M}_n(K)$, è dotato anche di struttura di anello. Per descrivere tale struttura dobbiamo definire un'opportuna operazione di "prodotto" tra matrici.

Definizione 5. Considerati $A = (a_1 \ a_2 \ \dots \ a_n) \in \mathfrak{M}_{1,n}(K)$ e $B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} \in \mathfrak{M}_{n,1}(K)$, si definisce *prodotto (righe per colonne) di A per B* l'elemento

$$AB = (a_1 \ a_2 \ \dots \ a_n) \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} := a_1 b_1 + a_2 b_2 + \dots + a_n b_n \in K.$$

Più generalmente, se $A \in \mathfrak{M}_{m,n}(K)$ e $B \in \mathfrak{M}_{n,p}(K)$, si chiama *matrice prodotto (righe per colonne) di A per B* la matrice:

$$AB = \begin{pmatrix} A^{(1)}B_{(1)} & A^{(1)}B_{(2)} & \dots & A^{(1)}B_{(p)} \\ A^{(2)}B_{(1)} & A^{(2)}B_{(2)} & \dots & A^{(2)}B_{(p)} \\ \vdots & \vdots & \vdots & \vdots \\ A^{(m)}B_{(1)} & A^{(m)}B_{(2)} & \dots & A^{(m)}B_{(p)} \end{pmatrix} \in \mathfrak{M}_{m,p}(K),$$

dove $A^{(i)}B_{(j)} = \sum_{k=1}^n a_{ik}b_{kj}$ [per $i = 1, \dots, m$, $j = 1, \dots, p$].

[Si noti che il prodotto AB è definito soltanto se le colonne di A sono quante le righe di B . Se A e B sono matrici quadrate dello stesso ordine, AB e BA sono ovviamente sempre definite e sono matrici quadrate di quello stesso ordine].

Proposizione 1. (i) Il prodotto righe per colonne è associativo, cioè:

$$(AB)C = A(BC), \quad \forall A \in \mathfrak{M}_{m,n}(K), \forall B \in \mathfrak{M}_{n,p}(K), \forall C \in \mathfrak{M}_{p,q}(K).$$

(ii) Valgono le seguenti proprietà:

$$\begin{aligned} (A+B)C &= AC + BC, \quad \forall A, B \in \mathfrak{M}_{m,n}(K), \forall C \in \mathfrak{M}_{n,p}(K); \\ A(B+C) &= AB + AC, \quad \forall A \in \mathfrak{M}_{m,n}(K), \forall B, C \in \mathfrak{M}_{n,p}(K); \\ AI_n &= A = I_m A, \quad \forall A \in \mathfrak{M}_{m,n}(K); \\ (cA)B &= c(AB) = A(cB), \quad \forall c \in K, \forall A \in \mathfrak{M}_{m,n}(K), \forall B \in \mathfrak{M}_{n,p}(K); \\ {}^t(A+B) &= {}^tA + {}^tB, \quad \forall A, B \in \mathfrak{M}_{m,n}(K); \\ {}^t(AB) &= {}^tB {}^tA, \quad \forall A \in \mathfrak{M}_{m,n}(K), \forall B \in \mathfrak{M}_{n,p}(K). \end{aligned}$$

(iii) Per ogni intero $n \geq 1$, $\mathfrak{M}_n(K)$ è un anello unitario, rispetto alla somma ed al prodotto righe per colonne. Se $n \geq 2$, tale anello è non commutativo e non integro.

Dim. (i) Per dimostrare l'associatività del prodotto righe per colonne, cioè che

$$((AB)C)_{ij} = (A(BC))_{ij}, \quad \forall i = 1, \dots, m, \quad \forall j = 1, \dots, q,$$

occorre un calcolo diretto. Si ha:

$$\begin{aligned} ((AB)C)_{ij} &= (AB)^{(i)}C_{(j)} = ((AB)_{i1} \dots (AB)_{ip}) \begin{pmatrix} c_{1j} \\ \vdots \\ c_{pj} \end{pmatrix} = (A^{(i)}B_{(1)} \dots A^{(i)}B_{(p)}) \begin{pmatrix} c_{1j} \\ \vdots \\ c_{pj} \end{pmatrix} \\ &= \sum_{k=1}^p A^{(i)}B_{(k)} c_{kj}. \text{ Poiché } A^{(i)}B_{(k)} = \sum_{t=1}^n a_{it}b_{tk}, \text{ allora:} \\ ((AB)C)_{ij} &= \sum_{k=1}^p \sum_{t=1}^n a_{it}b_{tk} c_{kj}. \end{aligned}$$

D'altra parte:

$$\begin{aligned} (A(BC))_{ij} &= A^{(i)}(BC)_{(j)} = (a_{i1} \dots a_{in}) \begin{pmatrix} (BC)_{1j} \\ \vdots \\ (BC)_{nj} \end{pmatrix} = (a_{i1} \dots a_{in}) \begin{pmatrix} B^{(1)}C_{(j)} \\ \vdots \\ B^{(n)}C_{(j)} \end{pmatrix} = \\ &= \sum_{t=1}^n a_{it}B^{(t)}C_{(j)}. \text{ Poiché } B^{(t)}C_{(j)} = \sum_{k=1}^p b_{tk}c_{kj}, \text{ allora:} \\ (A(BC))_{ij} &= \sum_{t=1}^n a_{it} \sum_{k=1}^p b_{tk}c_{kj} = \sum_{t=1}^n \sum_{k=1}^p a_{it}b_{tk}c_{kj}. \end{aligned}$$

Le due sommatorie sopra ottenute coincidono [infatti, in base alla commutatività della somma, è possibile scambiare l'ordine di sommazione].

(ii) Le verifiche delle varie proprietà sono lasciate allo studente.

(ii) Che $\mathfrak{M}_n(K)$ sia un gruppo rispetto alla somma già lo sappiamo [in quanto si tratta di un K -spazio vettoriale]. In base ad (i) e alle prime due proprietà di (ii), il prodotto righe per colonne è associativo e verifica le due leggi distributive (destra e sinistra). Dunque $(\mathfrak{M}_n(K), +, \cdot)$ è un anello.

Essendo $m = n$, dalla terza proprietà di (ii), $AI_n = A = I_n A$, $\forall A \in \mathfrak{M}_n(K)$: dunque l'anello è unitario, con unità la matrice unità I_n .

Resta da verificare l'ultima affermazione. A tale scopo conviene introdurre [e possiamo farlo, più generalmente, in $\mathfrak{M}_{m,n}(K)$] la seguente definizione di *matrice elementare di posto* (h, k) , denotata

E^{hk} . Si ponga, per ogni $h = 1, \dots, m$, $k = 1, \dots, n$:

$$(E^{hk})_{ij} = \begin{cases} 0 & \text{se } (i, j) \neq (h, k), \\ 1 & \text{se } (i, j) = (h, k) \end{cases}$$

[in altri termini, E^{hk} è una matrice in $\mathfrak{M}_{m,n}(K)$ con un unico elemento non nullo, quello di posto (h, k) , che ha valore 1]. Lasciamo allo studente il compito di verificare che in $\mathfrak{M}_n(K)$ (con $n \geq 2$) risulta:

$$E^{12} E^{21} = E^{11}, \text{ mentre } E^{21} E^{12} = E^{22}.$$

Da ciò segue che l'anello $\mathfrak{M}_n(K)$ non è commutativo. Risulta inoltre:

$$E^{11} E^{22} = \mathbf{0} \text{ (matrice nulla di } \mathfrak{M}_n(K)).$$

Dunque l'anello $\mathfrak{M}_n(K)$ non è integro.

Osservazione 2. Le matrici elementari $E^{hk} \in \mathfrak{M}_{m,n}(K)$, introdotte nella precedente dimostrazione, godono di un'importante ed immediata proprietà. Risulta, per ogni $A \in \mathfrak{M}_{m,n}(K)$:

$$\sum_{h=1}^m \sum_{k=1}^n a_{hk} E^{hk} = A$$

[cioè ogni matrice $A \in \mathfrak{M}_{m,n}(K)$ è esprimibile come somma di matrici elementari di $\mathfrak{M}_{m,n}(K)$, ciascuna moltiplicata per un opportuno scalare]. Ad esempio, $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \in \mathfrak{M}_{2,3}(\mathbf{Q})$ coincide con

$$E^{11} + 2E^{12} + 3E^{13} + 4E^{21} + 5E^{22} + 6E^{23}.$$

Osservazione 3. Sappiamo che gli elementi invertibili di un anello unitario formano un gruppo. Dunque, nel caso dell'anello $\mathfrak{M}_n(K)$, le matrici quadrate invertibili di ordine n formano un gruppo, che è denotato $\mathbf{GL}_n(K)$ ed è chiamato *gruppo generale lineare di ordine n su K*. Vedremo nel prossimo capitolo che le matrici invertibili sono caratterizzate dal fatto di avere *determinante* non nullo.

ESERCIZI PROPOSTI

2.2.1. Assegnate le due matrici (dipendenti da un parametro $a \in \mathbf{R}$)

$$A = \begin{pmatrix} 1 & 0 & 1 \\ a & 0 & 1 \end{pmatrix} \in \mathfrak{M}_{2,3}(\mathbf{R}), \quad B = \begin{pmatrix} 1 & a \\ a & 0 \\ 1 & 2 \end{pmatrix} \in \mathfrak{M}_{3,2}(\mathbf{R}),$$

- (i) Determinare gli eventuali $a \in \mathbf{R}$ per cui AB è una matrice triangolare superiore.
- (ii) Determinare gli eventuali $a \in \mathbf{R}$ per cui BA è una matrice simmetrica.

2.2.2. Siano $A, B \in \mathfrak{M}_n(K)$ due matrici simmetriche. Verificare che

la matrice AB è simmetrica $\iff A$ e B commutano.

2.2.3. Assegnata la matrice $A = \begin{pmatrix} 0 & 1 \\ 2 & 2 \\ -1 & 0 \end{pmatrix} \in \mathfrak{M}_{3,2}(\mathbf{R})$, verificare che le due matrici ${}^t\!AA$ e $A{}^t\!A$

sono simmetriche. È vero, più in generale, che per ogni $A \in \mathfrak{M}_{m,n}(K)$, le due matrici ${}^t\!AA$ e $A{}^t\!A$ sono simmetriche? È vero che, per ogni $A \in \mathfrak{M}_n(K)$, risulta: ${}^t\!AA = A{}^t\!A$?

2.2.4. Verificare che le matrici simmetriche di $\mathfrak{M}_n(K)$ formano un sottospazio vettoriale di $\mathfrak{M}_n(K)$. Formano un sottoanello di $\mathfrak{M}_n(K)$?

2.2.5. Si consideri in $\mathfrak{M}_2(\mathbf{R})$ la matrice $A = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$. Verificare che $A \in \mathbf{GL}_2(\mathbf{R})$, cioè che esiste $B \in \mathfrak{M}_2(\mathbf{R})$ tale che $AB = I_2 = BA$.

3. Classi resto modulo un intero

In questo paragrafo studieremo la struttura algebrica dell'insieme quoziante $\mathbf{Z}_{/\equiv_n}$, dove \equiv_n è la relazione di congruenza modulo n , introdotta nella **Def. 4 del Cap. 1.3**. Ma prima di far ciò è opportuno ricordare alcune definizioni ed alcuni risultati certamente noti allo studente: la *divisione euclidea in \mathbf{Z}* , il *massimo comun divisore di due (o più) interi* ed il *teorema fondamentale dell'Aritmetica*. [Una trattazione più completa di tali concetti può essere trovata ad esempio in www.mat.uniroma1.it/people/campanella, **Appunti di Algebra 1, Cap. 2**, paragrafi.1,2,3].

La divisione euclidea in \mathbf{Z} altro non è che l'usuale divisione tra numeri interi, che tutti abbiamo imparato ad eseguire già nella scuola primaria. Che sia sempre possibile eseguire tale divisione è oggetto del seguente risultato (che non dimostreremo), la cui dimostrazione poggia sul *Principio del minimo* [che afferma: *ogni sottoinsieme non vuoto di \mathbf{N} è dotato di minimo (cioè del più piccolo elemento), rispetto alla relazione di disegualanza \leq*].

Teorema 1. Siano $a, b \in \mathbf{Z}$, $b \neq 0$. Esiste un'unica coppia $(q, r) \in \mathbf{Z} \times \mathbf{Z}$ tale che

$$a = bq + r, \quad 0 \leq r < |b|.$$

Gli interi q, r sono detti rispettivamente *quoziante* ed *resto* della divisione euclidea di a per b , mentre a, b ne sono rispettivamente *dividendo* e *divisore*.

Nel **Cap. 1.3** abbiamo definito la relazione di divisibilità in \mathbf{Z} , che ora ricordiamo: $\forall a, b \in \mathbf{Z}$:

$$a | b \iff b = at, \exists t \in \mathbf{Z}.$$

Si può poi subito verificare che

$$a | b \iff a\mathbf{Z} \supseteq b\mathbf{Z}$$

[dove ovviamente $a\mathbf{Z} := \{at, \forall t \in \mathbf{Z}\}$ è l'insieme dei multipli interi di a e $b\mathbf{Z} := \{bt, \forall t \in \mathbf{Z}\}$ è l'insieme dei multipli interi di b]. Verifichiamo tale equivalenza:

$$(\implies) a | b \implies b = at, \exists t \in \mathbf{Z} \implies b \in a\mathbf{Z} \implies b\mathbf{Z} \subseteq a\mathbf{Z}.$$

$$(\impliedby) b\mathbf{Z} \subseteq a\mathbf{Z} \implies b = b \cdot 1 \in a\mathbf{Z} \implies b = at, \exists t \in \mathbf{Z} \implies a | b.]$$

Rimarchiamo il fatto (forse utile nella pratica) che il termine "divide" tra interi corrisponde al termine "contiene" tra gli insiemi dei multipli di tali interi.

Osservazione 1. Come osservato in **Cap. 1.3**, la relazione $|$ su \mathbf{Z} è una relazione di pre-ordine (cioè rilessiva e transitiva) non totale. Valgono inoltre i seguenti semplici fatti (la cui verifica è lasciata per esercizio):

(1) ogni $a \in \mathbf{Z}$ ammette come divisori $\pm a, \pm 1$ detti *divisori banali* di a [”banali”, in quanto ci sono sempre]. Gli altri (eventuali) divisori di a sono detti *divisori propri* di a .

Ad esempio 6 ha (oltre ai quattro divisori banali) anche quattro divisori propri: $\pm 2, \pm 3$; invece ad esempio 5 (così come ogni numero primo) ammette soltanto i quattro divisori banali.

(2) per ogni $a \in \mathbf{Z}$: $a | 0$ e $1 | a$. Inoltre: $0 | a \iff a = 0$, mentre: $a | 1 \iff a = \pm 1$.

(3) $a | b \iff ac | bc, \forall c \in \mathbf{Z} \iff ac | bc, \exists c \in \mathbf{Z}, c \neq 0$.

(4) $a | b$ e $a | c \iff a | bx + cy, \forall x, y \in \mathbf{Z}$.

(5) $a | b$ e $b | a \implies b = \pm a$.

Ogni studente pensa di conoscere la definizione di *massimo comun divisore* di due numeri interi. Ma probabilmente ne conosce non la definizione, bensì una regola per calcolarlo. Ad esempio, assegnati gli interi 36, 60, il massimo comun divisore è 12. Perché? Si può pensare che basta scrivere le fattorizzazioni dei due numeri come prodotto di primi:

$$36 = 2^2 \cdot 3^2, \quad 60 = 2^2 \cdot 3 \cdot 5$$

e poi affermare che il massimo comun divisore è dato dal prodotto dei primi comuni alle due fattorizzazioni [cioè 2, 3], presi con il minimo esponente che compare nelle due fattorizzazioni.

È vero che il massimo comun divisore si calcola in questo modo, ma per accettare questa regola come definizione bisogna non solo saper fattorizzare un intero (... e questo è un gran bel problema!) ma aver anche dimostrato che ogni intero $\neq 0, \pm 1$ ammette una (ed una sola) fattorizzazione come prodotto di numeri primi [tale fatto è vero ed è noto come *Teorema fondamentale dell'Aritmetica*]. La definizione di massimo comun divisore che conviene dare è perciò un'altra.

Definizione 1. Siano $a, b \in \mathbf{Z}$, non entrambi nulli. Si chiama *massimo comun divisore* [abbreviato *MCD*] di a, b ogni intero $d \geq 1$ verificante le due condizioni:

- (i) $d | a$ e $d | b$;
- (ii) per ogni intero $d' \geq 1$ tale che $d' | a$ e $d' | b$, risulta che $d' | d$.

Il *MCD* di a, b viene denotato con $MCD(a, b)$. Se $MCD(a, b) = 1$, a e b sono detti *coprimi* (o *relativamente primi*).

Si noti che la (i) afferma che d è un comune divisore di a, b ; la (ii) afferma che d è il più grande tra i divisori comuni di a, b [infatti, se $d' | d$, certo $d' \leq d$ (essendo d, d' entrambi positivi)].

A questa definizione va fatto seguire un teorema di esistenza e unicità del *MCD*, la cui dimostrazione sfrutta il Principio del minimo ed utilizza la divisione euclidea.

Teorema 2. (*Esistenza ed unicità del MCD*). Se $a, b \in \mathbf{Z}$ sono non entrambi nulli, $MCD(a, b)$ esiste ed è unico.

Dim. (Esistenza). Utilizzeremo il *Principio del minimo*, che, come già ricordato, afferma che ogni sottoinsieme non vuoto di \mathbf{N} possiede un elemento minimo. In \mathbf{N} definiamo il seguente sottoinsieme

$$S = \{n \in \mathbf{N} : n > 0 \text{ e } n = ax + by, \exists x, y \in \mathbf{Z}\}.$$

Essendo tale insieme non vuoto (come facilmente si verifica), è dotato di minimo. Indichiamo con d tale minimo. Poiché $d \in S$, per opportuni $s, t \in \mathbf{Z}$, sia $d = as + bt$. Utilizzando la divisione euclidea tra numeri interi, verificheremo che d è un *MCD* di a, b .

Dividiamo a per d . Otteniamo $a = dq + r$, con $0 \leq r < d$. Allora

$$r = a - dq = a - (as + bt)q = a(1 - sq) + b(-tq).$$

Se fosse $r > 0$, allora $r \in S$, ma ciò contraddirrebbe la minimalità di d in S . Dunque $r = 0$, cioè $a = dq$. Pertanto d è un divisore di a . In modo del tutto analogo si verifica che d è anche un divisore di b .

Sia ora d' un altro divisore positivo di a e b . Dobbiamo verificare che d' è un divisore di d . Sia $a = d'u$ e $b = d'v$, per opportuni $u, v \in \mathbf{Z}$. Allora

$$d = as + bt = d'u s + d'v t = d'(us + vt).$$

Si conclude quindi che d' è un divisore di d .

(Unicità). Siano d, \tilde{d} due *MCD* di a, b . Da $\tilde{d} = MCD(a, b)$ segue che $d | \tilde{d}$; da $d = MCD(a, b)$ segue che $\tilde{d} | d$. Pertanto $d | \tilde{d}$ e $\tilde{d} | d$. Dall'**Osserv. 1(5)** segue che $\tilde{d} = \pm d$. Ma \tilde{d}, d sono entrambi positivi e quindi $\tilde{d} = d$.

Dalla definizione di *MCD* segue subito che $MCD(a, b) = MCD(b, a)$ e che $MCD(\pm a, \pm b) = MCD(a, b)$; inoltre $MCD(a, 0) = |a|$, $\forall a \in \mathbf{Z}$.

Dalla dimostrazione dell'esistenza del *MCD* segue che, se $d = MCD(a, b)$, d si può scrivere come una combinazione a coefficienti interi di a e b . Tale uguaglianza viene chiamata *identità di Bézout*. Precisiamo il tutto nel seguente corollario.

Corollario 1. (*Identità di Bézout*). Siano $a, b \in \mathbf{Z}$ non entrambi nulli. Se $d = MCD(a, b)$, esistono $x_0, y_0 \in \mathbf{Z}$ tali che

$$d = ax_0 + by_0 \quad [\text{identità di Bézout per } a, b].$$

Ad esempio, essendo $12 = MCD(36, 60)$, si ha:

$$12 = 36 \cdot 2 + 60 \cdot (-1).$$

Come siamo venuti a capo di tale identità? Prima di verificare che non si è proceduto a caso, diciamo che l'identità ottenuta non è unica. Ad esempio si ha anche: $12 = 36 \cdot (-58) + 60 \cdot 35$. Anzi, ci sono infinite identità di Bézout per due qualsiasi interi. Infatti, $\forall c \in \mathbf{Z}$:

$$d = ax_0 + by_0 = ax_0 + by_0 \pm abc = a(x_0 + bc) + b(y_0 - ac).$$

Per calcolare il MCD di due interi e calcolare un'identità di Bézout si ricorre ad un celeberrimo algoritmo: l'algoritmo euclideo delle divisioni successive. In che cosa consiste? Ci serve una premessa.

Lemma 1. Siano $a, b \in \mathbf{Z}$, con $b \neq 0$. Sia $a = bq + r$, con $0 \leq r < |b|$. Risulta:

$$MCD(a, b) = MCD(b, r).$$

Dim. Siano $d := MCD(a, b)$ e $d_1 := MCD(b, r)$. Basta dimostrare che: $d \mid d_1$ e $d_1 \mid d$.

Infatti:

- se $d \mid a$ e $d \mid b$, allora $d \mid a - bq = r$. Dunque $d \mid b$ e $d \mid r$. Pertanto $d \mid d_1$.
- se $d_1 \mid b$ e $d_1 \mid r$, allora $d_1 \mid bq + r = a$. Dunque $d_1 \mid a$ e $d_1 \mid b$. Pertanto $d_1 \mid d$.

Assumiamo che sia $a > b \geq 0$ [se così non fosse avremmo o casi banali o casi facilmente riconducibili a questo]. L'algoritmo euclideo delle divisioni successive consiste in una successione finita di divisioni euclidee (a partire dalla divisione di a per b), in modo che il divisore ed il resto (se non nullo) diventino rispettivamente dividendo e divisore della divisione successiva. Il procedimento si interrompe non appena si ottiene resto nullo. Dunque l'algoritmo è articolato nei seguenti passi:

(1^o) $a = bq_1 + r_1$, $0 \leq r_1 < b$. Se $r_1 > 0$, si procede con il passo successivo.

(2^o) $b = r_1 q_2 + r_2$, $0 \leq r_2 < r_1$. Se $r_2 > 0$, si procede con il passo successivo.

(3^o) $r_1 = r_2 q_3 + r_3$, $0 \leq r_3 < r_2$. Se $r_3 > 0$, si procede con il passo successivo.

⋮
⋮

Poiché $b > r_1 > r_2 > r_3 > \dots$, $\exists n \in \mathbf{N}$ tale che $r_n > 0$ e $r_{n+1} = 0$. Ciò significa che gli ultimi due passi dell'algoritmo sono

(n^o) $r_{n-2} = r_{n-1} q_n + r_n$, $0 < r_n < r_{n-1}$.

(n+1^o) $r_{n-1} = r_n q_{n+1} + 0$.

Dal **Lemma 1** segue:

$$MCD(a, b) = MCD(b, r_1) = MCD(r_1, r_2) = MCD(r_2, r_3) = \dots = MCD(r_{n-1}, r_n) = MCD(r_n, 0) = r_n.$$

Dunque $r_n = MCD(a, b)$. Il MCD è quindi l'ultimo resto non nullo dell'algoritmo.

Per ottenere un'identità di Bézout si procede in questo modo. Si isolano gli n resti ottenuti nelle divisioni successive. Per ricordarsi di non eseguire semplificazioni numeriche, si conviene di scrivere tra parentesi quadre gli interi a, b ed i resti r_k . Si ottengono pertanto le seguenti uguaglianze:

(1^o) $[r_1] = [a] - q_1[b]$.

(2^o) $[r_2] = [b] - q_2[r_1]$.

(3^o) $[r_3] = [r_1] - q_3[r_2]$.

⋮
⋮

$$(\mathbf{n}^0) \quad [r_n] = [r_{n-2}] - q_n[r_{n-1}].$$

Si osserva subito che, $\forall k = 1, 2, \dots, n$, $[r_k]$ è combinazione di $[r_{k-1}]$ e $[r_{k-2}]$ (convenendo in particolare di porre $r_0 = b$, $r_{-1} = a$). A partire da $k = 2$ (se $r_2 \neq 0$), con successive sostituzioni si può quindi esprimere ogni $[r_k]$ come combinazione lineare di $[a]$ e $[b]$, con coefficienti che sono funzioni di q_1, \dots, q_k .

In conclusione, si otterrà $[r_n]$ in funzione di $[a]$, $[b]$. Eliminando le parentesi quadre, si ottiene, come richiesto, un'identità di Bézout relativa ad a , b (con $a \geq b > 0$).

Esempio. Calcolare il MCD e un'identità di Bézout per $a = -123$, $b = -39$.

Si ha: $123 > 39 > 0$. Risulta:

$$\begin{aligned} 123 &= 39 \cdot 3 + 6 & [6] &= [123] - [39] \cdot 3 \\ 39 &= 6 \cdot 6 + 3 & [3] &= [39] - [6] \cdot 6 \\ 6 &= 3 \cdot 2 + 0. & & \end{aligned}$$

Dunque $MCD(123, 39) = 3$ e

$$[3] = [39] - ([123] - [39] \cdot 3) \cdot 6 = [39] - 6[123] + 18[39] = -6[123] + 19[39].$$

Da ciò segue $3 = -6 \cdot 123 + 19 \cdot 39$ ed, essendo $a = -123$, $b = -39$, si ottiene l'identità di Bézout

$$3 = 6 \cdot a - 19 \cdot b.$$

Per concludere le premesse al paragrafo occorre enunciare il *Teorema fondamentale dell'Aritmetica*, a cui premettiamo la definizione di numero primo.

Definizione 2. Sia $p \in \mathbf{Z}$, $p \geq 2$. p è detto *numero primo* se p ha soltanto i quattro divisori banali ± 1 , $\pm p$ (cioè p non ha divisori propri).

Il Teorema fondamentale dell'Aritmetica viene dimostrato solitamente per induzione (su $n \geq 2$) per gli interi positivi. Poi, come conseguenza, si può dimostrare la versione sugli interi ed infine la nota formula che ci permette di calcolare il MCD di due interi a partire dalla loro fattorizzazione. Ci limiteremo ad enunciare i tre risultati.

Teorema 3. (*Teorema Fondamentale dell'Aritmetica* (in \mathbf{N})).

- (1) Ogni naturale $n \geq 2$ è prodotto di un numero finito di numeri primi.
- (2) Se per ogni $n \geq 2$ poniamo:

$$n = p_1^{h_1} p_2^{h_2} \dots p_s^{h_s}, \text{ con } \begin{cases} s \geq 1 \\ p_1, \dots, p_s \text{ primi distinti} \\ h_1, \dots, h_s \geq 1, \end{cases}$$

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

Corollario 2. (*Teorema Fondamentale dell'Aritmetica* (in \mathbf{Z})). Sia $a \in \mathbf{Z}$, $a \neq 0$, $a \neq \pm 1$. L'intero a si scrive in modo unico (a meno dell'ordine dei fattori) nella forma

$$a = \pm p_1^{h_1} p_2^{h_2} \dots p_s^{h_s},$$

dove: $s \geq 1$, p_1, \dots, p_s sono numeri primi distinti, $h_1, \dots, h_s \geq 1$, e vale il segno $+$ se $a > 0$, vale il segno $-$ se $a < 0$.

Corollario 3. Siano $a, b \in \mathbf{Z}$, $a, b \neq 0, \pm 1$. Se

$$a = \pm p_1^{h_1} p_2^{h_2} \dots p_s^{h_s}, \quad b = \pm p_1^{k_1} p_2^{k_2} \dots p_s^{k_s}$$

con p_1, \dots, p_s numeri primi e $h_i, k_i \geq 0$, allora

$$MCD(a, b) = p_1^{d_1} p_2^{d_2} \dots p_s^{d_s}, \text{ con } d_i := \min\{h_i, k_i\} \quad (\forall i = 1, \dots, s).$$

Nota. Si osservi che, avendo assunto $h_i, k_i \geq 0$, è stato possibile esprimere a, b come prodotto degli stessi primi [ad esempio, posto $a = 36, b = 60$, allora $a = 2^2 3^2 5^0, b = 2^2 3^1 5^1$].

Ora veniamo all'oggetto di questo paragrafo. Ricordiamo la definizione di relazione di congruenza, già data nella **Def. 4 del Cap. 1.3**: *Sia n un intero ≥ 2 . Si chiama relazione di congruenza modulo n la seguente relazione su \mathbf{Z} : presi comunque $a, b \in \mathbf{Z}$,*

$$a \equiv b \pmod{n} \iff n \mid b - a.$$

Sappiamo che si tratta di una relazione di equivalenza su \mathbf{Z} . Osserviamo ora il seguente fatto.

Proposizione 1. *Sia $n \geq 2$ e siano $a, b \in \mathbf{Z}$. Risulta:*

$$a \equiv b \pmod{n} \iff a, b, \text{ divisi per } n, \text{ hanno lo stesso resto.}$$

Dim. (\Leftarrow) Se $a = nq_1 + r$ e $b = nq_2 + r$, allora $b - a = n(q_2 - q_1)$. Dunque $n \mid b - a$, cioè $a \equiv b \pmod{n}$.

(\Rightarrow) Sia $a \equiv b \pmod{n}$ e quindi $b - a = nt, \exists t \in \mathbf{Z}$. Dividiamo a e b per n . Si ha:

$$a = nq_1 + r_1, \quad b = nq_2 + r_2, \quad \text{con } 0 \leq r_1 < n, \quad 0 \leq r_2 < n.$$

Si tratta di verificare che $r_1 = r_2$. Risulta:

$$nt = b - a = n(q_2 - q_1) + (r_2 - r_1).$$

Ne segue che $n \mid r_2 - r_1$. Dalle limitazioni sui resti segue che $-n < r_2 - r_1 < n$. Allora necessariamente $r_2 - r_1 = 0$.

Vogliamo ora esaminare le classi di equivalenza modulo \equiv_n . Denoteremo con $[a]_n$ o più semplicemente con \bar{a} [ove non sia necessario evidenziare n] la classe di equivalenza di $a \in \mathbf{Z}$ modulo \equiv_n . Per definizione,

$$\bar{a} = [a]_n = \{x \in \mathbf{Z} \mid a \equiv_n x\}.$$

Poiché $a \equiv_n x \iff x = a + tn, \exists t \in \mathbf{Z}$, allora

$$\bar{a} = [a]_n = \{x \in \mathbf{Z} \mid x = a + tn, \exists t \in \mathbf{Z}\} = \{a + tn, \forall t \in \mathbf{Z}\}.$$

Pertanto denoteremo tale insieme anche nella forma $a + n\mathbf{Z}$ (dove $n\mathbf{Z} = \{nt, \forall t \in \mathbf{Z}\}$).

Quante sono le classi di equivalenza modulo \equiv_n ? Eseguiamo la divisione euclidea di a per n e sia $a = nq + r$, con $0 \leq r < n$. Poiché anche la divisione di r per n ha resto r [in quanto $r = n0 + r$], allora $a \equiv_n r$, cioè $\bar{a} = \bar{r}$. Se poi

$$r \equiv_n r_1 \text{ con } 0 \leq r < n, \quad 0 \leq r_1 < n, \quad \text{allora } r = r_1$$

[infatti, dalle due diseguaglianze segue che $-n < r_1 - r < n$; se quindi $r_1 - r = nt$, allora $t = 0$, cioè $r = r_1$]. Pertanto di classi di equivalenza distinte ne esistono esattamente n , tante quanti i possibili resti della divisione euclidea di un intero per n . Tali classi di equivalenza vengono chiamate *classi resto modulo n* .

Ad esempio, se $n = 2$, le classi resto modulo 2 sono due, cioè

$$\bar{0} = 0 + 2\mathbf{Z} = \{0, \pm 2, \pm 4, \dots\} \quad \text{e} \quad \bar{1} = 1 + 2\mathbf{Z} = \{\pm 1, \pm 3, \pm 5, \dots\}$$

$\bar{0}$ è l'insieme degli interi pari, mentre $\bar{1}$ è l'insieme degli interi dispari. Ovviamente $\bar{0}$ può essere rappresentata da un qualsiasi intero pari (ad esempio $\bar{0} = \bar{12}$), mentre $\bar{1}$ può essere rappresentata da un qualsiasi intero dispari (ad esempio $\bar{1} = \bar{-1}$).

Se invece $n = 3$, abbiamo tre classi resto modulo 3, cioè

$$\bar{0} = 3\mathbf{Z} = \{3k, \forall k \in \mathbf{Z}\}, \quad \bar{1} = 1 + 3\mathbf{Z} = \{1 + 3k, \forall k \in \mathbf{Z}\}, \quad \bar{2} = 2 + 3\mathbf{Z} = \{2 + 3k, \forall k \in \mathbf{Z}\}.$$

L'insieme quozione $\mathbf{Z}_{/\equiv_n}$ di \mathbf{Z} modulo la relazione \equiv_n viene indicato più semplicemente con \mathbf{Z}_n ed è chiamato *insieme delle classi resto modulo n* . L'insieme \mathbf{Z}_n ha cardinalità n ed è formato dalle classi resto

$$\overline{0}, \overline{1}, \overline{2}, \dots, \overline{n-1}.$$

Vogliamo ora assegnare a \mathbf{Z}_n una struttura algebrica. Per far questo bisogna definire un'operazione di "somma di classi resto" ed una di "prodotto di classi resto".

Come definiamo ad esempio $\overline{a} + \overline{b}$? Sarebbe naturale dire che $\overline{a} + \overline{b} = \overline{a+b}$, ma per poterlo fare occorre verificare che cambiando il rappresentante delle due classi resto, il risultato è lo stesso. Si tratta cioè di verificare che, se $\overline{a} = \overline{a_1}$ e $\overline{b} = \overline{b_1}$, allora $\overline{a+b} = \overline{a_1+b_1}$. Si dice che in tal caso la relazione di congruenza è *compatibile* con l'addizione di \mathbf{Z} .

Proposizione 2. *La relazione \equiv_n è compatibile con le operazioni di somma e prodotto in \mathbf{Z} . $(\mathbf{Z}_n, +, \cdot)$ è un anello commutativo unitario.*

Dim. Siano $a \equiv_n a_1$ e $b \equiv_n b_1$. Bisogna verificare che:

$$a + b \equiv_n a_1 + b_1 \quad \text{e} \quad a \cdot b \equiv_n a_1 \cdot b_1$$

Se infatti $a - a_1 = nt$, $b - b_1 = ns$, allora $a + b - (a_1 + b_1) = n(t + s)$ e dunque $a + b \equiv_n a_1 + b_1$. Inoltre:

$$a \cdot b - a_1 \cdot b_1 = a \cdot b - a \cdot b_1 + a \cdot b_1 - a_1 \cdot b_1 = a(b - b_1) + (a - a_1)b_1 = ans + nt b_1 = n(as + tb_1).$$

Dunque $a \cdot b \equiv_n a_1 \cdot b_1$.

Sono quindi ben definite in \mathbf{Z}_n le due operazioni:

$$\overline{a} + \overline{b} = \overline{a+b}, \quad \overline{a} \cdot \overline{b} = \overline{a \cdot b}, \quad \forall \overline{a}, \overline{b} \in \mathbf{Z}_n.$$

Verifichiamo che $(\mathbf{Z}_n, +)$ è un gruppo commutativo. Si ha:

- $(\overline{a} + \overline{b}) + \overline{c} = \overline{a} + (\overline{b} + \overline{c}), \quad \forall \overline{a}, \overline{b}, \overline{c} \in \mathbf{Z}_n;$
- $\overline{a} + \overline{0} = \overline{a} = \overline{0} + \overline{a}, \quad \forall \overline{a} \in \mathbf{Z}_n;$
- $\overline{a} + \overline{-a} = \overline{0} = \overline{-a} + \overline{a}, \quad \forall \overline{a} \in \mathbf{Z}_n;$
- $\overline{a} + \overline{b} = \overline{b} + \overline{a}, \quad \forall \overline{a}, \overline{b} \in \mathbf{Z}_n.$

[Le verifiche sono lasciate per esercizio]. Valgono inoltre le seguenti proprietà [anch'esse lasciate per esercizio]:

- $(\overline{a} \cdot \overline{b}) \cdot \overline{c} = \overline{a} \cdot (\overline{b} \cdot \overline{c}), \quad \forall \overline{a}, \overline{b}, \overline{c} \in \mathbf{Z}_n;$
- $\overline{a} \cdot (\overline{b} + \overline{c}) = \overline{a} \cdot \overline{b} + \overline{a} \cdot \overline{c}, \quad (\overline{a} + \overline{b}) \cdot \overline{c} = \overline{a} \cdot \overline{c} + \overline{b} \cdot \overline{c}, \quad \forall \overline{a}, \overline{b}, \overline{c} \in \mathbf{Z}_n;$
- $\overline{a} \cdot \overline{b} = \overline{b} \cdot \overline{a}, \quad \forall \overline{a}, \overline{b} \in \mathbf{Z}_n.$
- $\overline{a} \cdot \overline{1} = \overline{a} = \overline{1} \cdot \overline{a}, \quad \forall \overline{a} \in \mathbf{Z}_n.$

Si conclude che $(\mathbf{Z}_n, +, \cdot)$ è un anello commutativo unitario.

In generale \mathbf{Z}_n non è integro. Ad esempio, in \mathbf{Z}_4 , $\overline{2} \cdot \overline{2} = \overline{0}$ e in \mathbf{Z}_6 , $\overline{2} \cdot \overline{3} = \overline{0}$. In generale in \mathbf{Z}_n non vale la legge di cancellazione. Ad esempio, in \mathbf{Z}_4 risulta: $\overline{2} \cdot \overline{2} = \overline{2} \cdot \overline{0}$, ma $\overline{2}$ non può essere cancellato, perché $\overline{2} \neq \overline{0}$. Altro esempio, in \mathbf{Z}_{12} : $\overline{3} \cdot \overline{4} = \overline{3} \cdot \overline{8}$, ma $\overline{4} \neq \overline{8}$.

Per approfondire la struttura algebrica di \mathbf{Z}_n ci resta da studiare il gruppo degli elementi invertibili di \mathbf{Z}_n e verificare che \mathbf{Z}_n è un campo $\iff n$ è un numero primo. Useremo qui l'identità di Bézout illustrata nel **Coroll. 1**.

Proposizione 3. *$(\mathbf{Z}_n, +, \cdot)$ è un campo $\iff n$ è un numero primo.*

Dim. (\implies) Se per assurdo n non fosse primo, esisterebbero $a, b \in \mathbf{Z}$ tali che $n = ab$, con $1 < a < n$, $1 < b < n$. Passando in \mathbf{Z}_n si avrebbe:

$$\overline{0} = \overline{n} = \overline{a} \overline{b} = \overline{a} \overline{b}$$

e dunque \mathbf{Z}_n sarebbe un anello non integro, mentre è un campo (e quindi è integro).

(\impliedby) Sia n primo e sia $a \in \mathbf{Z}$ tale che $1 \leq a < n$. Ovviamente a, n sono coprimi, cioè $MCD(a, n) = 1$, e quindi, calcolando un'identità di Bézout, $1 = ar + ns$, per opportuni $r, s \in \mathbf{Z}$. Passando in \mathbf{Z}_n , si ottiene:

$$\bar{1} = \overline{a r + n s} = \bar{a} \bar{r} + \bar{n} \bar{s} = \bar{a} \bar{r} + \bar{0} = \bar{a} \bar{r}.$$

Abbiamo così ottenuto che \bar{a} è invertibile in \mathbf{Z}_n , con inverso \bar{r} . Ogni elemento non nullo di \mathbf{Z}_n è quindi invertibile e pertanto \mathbf{Z}_n è un campo.

Dal risultato precedente ricaviamo che gli anelli \mathbf{Z}_n o sono campi o sono anelli non integri. Ci chiediamo, in quest'ultimo caso, quali siano i loro elementi invertibili.

Proposizione 4. $\bar{a} \in \mathcal{U}(\mathbf{Z}_n) \iff a, n$ sono interi coprimi.

Dim. (\implies) Se $\bar{a} \in \mathcal{U}(\mathbf{Z}_n)$, esiste $\bar{b} \in \mathbf{Z}_n$ tale che $\bar{a} \bar{b} = \bar{1}$. Ne segue che $\overline{1 - a b} = \bar{0}$. Allora, passando in \mathbf{Z} :

$$1 - a b \equiv_n 0, \text{ cioè } 1 - a b = n t, \exists t \in \mathbf{Z}.$$

Poiché $1 = a b + n t$, considerato l'insieme S (della dimostrazione del **Teor. 2**) relativo agli interi a, n , risulta che $1 \in S$ e dunque (essendo 1 necessariamente il minimo di S) il *MCD* tra n ed a è 1, cioè a, n sono interi coprimi.

(\impliedby) Se a, n sono coprimi, $1 = a r + n s$, per opportuni $r, s \in \mathbf{Z}$. Dunque, ripetendo un ragionamento già fatto, \bar{r} è inverso di \bar{a} . Pertanto $\bar{a} \in \mathcal{U}(\mathbf{Z}_n)$.

Ad esempio:

$$\mathcal{U}(\mathbf{Z}_6) = \{\bar{1}, \bar{5}\}, \quad \mathcal{U}(\mathbf{Z}_9) = \{\bar{1}, \bar{2}, \bar{4}, \bar{5}, \bar{7}, \bar{8}\}, \quad \mathcal{U}(\mathbf{Z}_{12}) = \{\bar{1}, \bar{5}, \bar{7}, \bar{11}\}, \text{ ecc.}$$

e, per quanto abbiamo osservato nel primo paragrafo, tali insiemi sono gruppi rispetto al prodotto.

La cardinalità di $\mathcal{U}(\mathbf{Z}_n)$ è data del numero degli interi positivi coprimi con n e minori di n . Esiste una funzione aritmetica importante, la *funzione di Eulero*, che esprime tale numero.

Definizione 3. Si chiama *funzione di Eulero* la funzione $\varphi : \mathbf{N}^* \rightarrow \mathbf{N}^*$ tale che

$$\varphi(n) = |\{k \in \mathbf{Z} : 1 \leq k \leq n \text{ e } k, n \text{ sono coprimi}\}|.$$

Vale il seguente risultato, che fornisce una formula diretta per il calcolo di φ , in funzione della fattorizzazione di un intero come prodotto di fattori primi.

Proposizione 5. Se $n = p_1^{r_1} \dots p_s^{r_s}$, con p_1, \dots, p_s primi, risulta:

$$\varphi(n) = (p_1^{r_1} - p_1^{r_1-1}) \cdot \dots \cdot (p_s^{r_s} - p_s^{r_s-1}).$$

Non dimostreremo tale proposizione [rinviamo a [www.mat.uniroma1.it/people/campanella, Appunti di Algebra 1](http://www.mat.uniroma1.it/people/campanella/Appunti_di_Algebra_1.pdf), Cap. 2.6, pag. 91]. Vogliamo soltanto osservare che la dimostrazione consegue subito dai due seguenti risultati:

(A) Se r, s sono naturali coprimi, $\varphi(rs) = \varphi(r) \cdot \varphi(s)$.

(B) Se p è primo, $\varphi(p^r) = p^r - p^{r-1}$ ($\forall r \geq 1$).

Ad esempio,

$$\varphi(144) = \varphi(12^2) = \varphi(2^4 3^2) = \varphi(2^4) \varphi(3^2) = (2^4 - 2^3)(3^2 - 3) = 48.$$

Ci poniamo ora il problema di risolvere equazioni di primo grado a coefficienti in \mathbf{Z}_n , cioè equazioni del tipo

$$(*) \quad \bar{a} X = \bar{b}, \text{ con } \bar{a}, \bar{b} \in \mathbf{Z}_n, \bar{a} \neq \bar{0}.$$

L'equazione (*) è detta *risolubile su \mathbf{Z}_n* se $\exists \bar{x} \in \mathbf{Z}_n$ tale che $\bar{a} \bar{x} = \bar{b}$. Accanto all'equazione (*) consideriamo la cosiddetta equazione "congruenziale"

$$(**) \quad a X \equiv b \pmod{n}$$

Si osserva subito che $(*)$ è risolubile su \mathbf{Z}_n se e solo se $(**)$ è risolubile su \mathbf{Z} [cioè $\exists x \in \mathbf{Z}$ tale che $ax \equiv b \pmod{n}$]. Infatti \bar{x} è soluzione di $(*) \iff x$ è soluzione di $(**)$.

Si noti che, se $(**)$ è risolubile e x ne è una soluzione, ogni $x + nc$ ($\forall c \in \mathbf{Z}$) è ancora soluzione di $(**)$. Ma tali soluzioni corrispondono ad un'unica soluzione di $(*)$ [infatti $\bar{x} = \overline{x+nc}$]. Invece le soluzioni di $(**)$ comprese tra 0 e $n-1$ corrispondono biunivocamente a quelle di $(*)$.

Osserviamo subito che se $\bar{a} \in \mathcal{U}(\mathbf{Z}_n)$ [cioè se $MCD(a, n) = 1$] l'equazione $(*)$ ha un'unica soluzione, che è $\bar{a}^{-1}\bar{b}$. Se invece $\bar{a} \notin \mathcal{U}(\mathbf{Z}_n)$ la situazione è più complicata. Cerchiamo di chiarirla con due esempi.

(A) Risolvere l'equazione

$$\overline{15}X = \overline{10} \text{ in } \mathbf{Z}_{21}.$$

Si tratta di risolvere l'equazione congruenziale

$$15X \equiv 10 \pmod{21}.$$

Assumiamo che $x \in \mathbf{Z}$ sia una soluzione di tale equazione. Allora $15x \equiv 10 \pmod{21}$ e quindi, per un opportuno $t \in \mathbf{Z}$, $15x - 10 = 21t$. Ne segue che $3(5x - 7t) = 10$ e dunque $3 \mid 10$: assurdo. Ne segue che l'equazione assegnata è priva di soluzioni, cioè è incompatibile.

(B) Risolvere l'equazione

$$\overline{15}X = \overline{12} \text{ in } \mathbf{Z}_{21}.$$

Si tratta di risolvere l'equazione congruenziale

$$15X \equiv 12 \pmod{21}.$$

Si può osservare che tale equazione ha le stesse soluzioni dell'equazione congruenziale

$$5X \equiv 4 \pmod{7}$$

[ottenuta dividendo a, b, n per il loro MCD 3]. Se infatti $15x \equiv 12 \pmod{21}$, allora, per un opportuno $s \in \mathbf{Z}$, $15x - 12 = 21s$ e quindi, dividendo per 3, $5x - 4 = 7s$, cioè $5x \equiv 4 \pmod{7}$. Viceversa, se $5y \equiv 4 \pmod{7}$, allora $5y - 4 = 7t$, per un opportuno $t \in \mathbf{Z}$. Moltiplicando per 3, $15y - 12 = 21t$ e quindi $15y \equiv 12 \pmod{21}$.

Ora risolviamo l'equazione $5X \equiv 4 \pmod{7}$. Essendo 5, 7 coprimi e $\overline{5} \cdot \overline{3} = \overline{1}$, allora, in \mathbf{Z}_7 : $X = \overline{3} \cdot \overline{4} = \overline{12} = \overline{5}$. L'equazione assegnata $\overline{15}X = \overline{12}$ ammette quindi soluzione $\overline{5}$ (in \mathbf{Z}_{21}). Ma, com facilmente si verifica, ammette anche altre due soluzioni: $\overline{12} = \overline{5} + \overline{7}$ e $\overline{19} = \overline{5} + \overline{7} \cdot \overline{2}$.

Come si giustifica tutto questo? Ci limitiamo ad enunciare il seguente risultato, che riassume l'analisi della risolubilità di equazioni di primo grado a coefficienti in \mathbf{Z}_n .

Proposizione 6. Assegnata l'equazione $\overline{a}X = \overline{b}$ su \mathbf{Z}_n , con $\overline{a} \neq \overline{0}$, risulta, posto $d = MCD(a, n)$:

$$\text{l'equazione } \overline{a}X = \overline{b} \text{ è risolubile } \iff d \mid b.$$

Se poi tale equazione è risolubile, ammette in \mathbf{Z}_n d soluzioni distinte, così ottenute: se $\overline{x_0}$ è una di tali soluzioni, le altre sono date da $\overline{x_0 + \frac{n}{d}h}$, per $h = 1, 2, \dots, d-1$.

Concludiamo il paragrafo enunciando un importante risultato, noto come *Teorema di Eulero-Fermat*, che ha estrema importanza in Crittografia.

Teorema 4. (*Teorema di Eulero-Fermat*). Sia $n \geq 2$ e sia a un intero coprimo con n . Risulta:

$$a^{\varphi(n)} \equiv 1 \pmod{n}.$$

[Dunque $\overline{a}^{\varphi(n)} = \overline{1}$, $\forall \overline{a} \in \mathcal{U}(\mathbf{Z}_n)$].

Un caso particolare di tale teorema, che è relativo al caso in cui il modulo n è primo, è il seguente risultato noto col nome di *Piccolo Teorema di Fermat*.

Teorema 5. (*Piccolo Teorema di Fermat*). Siano a, p interi coprimi. Se p è primo, risulta:

$$a^{p-1} \equiv 1 \pmod{p}.$$

[Dunque $\bar{a}^{p-1} = \bar{1}$, $\forall \bar{a} \in \mathbf{Z}_p$.]

Il teorema di Eulero-Fermat è un utile strumento per risolvere problemi aritmetici, come negli esempi che seguono.

Esempio 2. Usando il teorema di Eulero-Fermat, calcolare le ultime due cifre di $n = 81^{82}$.

Si osservi che le ultime due cifre di un naturale n sono date dal resto della divisione euclidea di n per 100. In altri termini, si ottengono risolvendo la congruenza $n \equiv X \pmod{100}$. Nel caso in esame la congruenza da risolvere è $81^{82} \equiv X \pmod{100}$.

Poiché $MCD(81, 100) = 1$ e $\varphi(100) = \varphi(4) \cdot \varphi(25) = 2 \cdot 20 = 40$, in base al teorema di Eulero-Fermat, $81^{40} = 81^{\varphi(100)} \equiv 1 \pmod{100}$. Dunque

$$81^{82} = (81^{40})^2 \cdot 81^2 \equiv 1^2 \cdot 81^2 = 6561 \equiv 61 \pmod{100}.$$

Le ultime due cifre di 81^{82} sono 6, 1.

Esempio 3. Usando il teorema di Eulero-Fermat, calcolare le ultime tre cifre di $n = 7^{827}$.

Si tratta di risolvere la congruenza $7^{827} \equiv X \pmod{1000}$.

Si ha: $MCD(7, 1000) = 1$ e $\varphi(1000) = \varphi(8) \cdot \varphi(125) = 4 \cdot 100 = 400$. Allora, in base al teorema di Eulero-Fermat, $7^{400} \equiv 1 \pmod{1000}$. Allora $7^{827} = (7^{400})^2 \cdot 7^{27} \equiv 1^2 \cdot 7^{27} \pmod{1000}$. Si tratta quindi di calcolare $7^{27} \pmod{1000}$.

Essendo $27 = 2^4 + 2^3 + 2 + 1$, allora $7^{27} \equiv 7^{16} \cdot 7^8 \cdot 7^2 \cdot 7 \pmod{1000}$. Si ha:

$$\begin{aligned} 7 &\equiv 7 \pmod{1000}, \\ 7^2 &\equiv 49 \pmod{1000}, \\ 7^4 &\equiv 401 \pmod{1000}, \\ 7^8 &\equiv (401)^2 \equiv 801 \pmod{1000}, \\ 7^{16} &\equiv (801)^2 \equiv 601 \pmod{1000}. \end{aligned}$$

Si verifica subito che $7^{16} \cdot 7^8 \equiv 601 \cdot 801 \equiv 401 \pmod{1000}$ e che $7^2 \cdot 7 \equiv 49 \cdot 7 \equiv 343 \pmod{1000}$. Allora

$$7^{27} = (7^{16} \cdot 7^8)(7^2 \cdot 7) \equiv 401 \cdot 343 \equiv 543 \pmod{1000}.$$

Si conclude che le ultime tre cifre di 7^{827} sono 5, 4, 3.

ESERCIZI PROPOSTI

2.3.1. Sia p un intero ≥ 2 . Dimostrare il seguente risultato:

p è primo \iff se p divide un prodotto, divide almeno un fattore [cioè $p | ab \implies p | a$ o $p | b$].

Per dimostrare tale risultato si proceda come segue:

(1) Usando l'identità di Bézout provare il seguente *Lemma di Euclide*:

siano $a, b, c \in \mathbf{Z}$. Se $a | bc$ e $MCD(a, b) = 1$, allora $a | c$.

(2) Usando il lemma di Euclide, dimostrare (\implies): se $p | ab$ e $p \nmid a$, allora $p | b$.

(3) Dimostrare (\iff): p ha solo fattori banali [cioè: $p = xy \implies x = \pm 1$ o $y = \pm 1$].

2.3.2. Scrivere la tavola moltiplicativa di \mathbf{Z}_6 . Verificare se $\{0, 2, 4\}$ è un sottoanello di \mathbf{Z}_6 .

2.3.3. Scrivere la tavola additiva di \mathbf{Z}_5 e la tavola moltiplicativa \mathbf{Z}_5 .

2.3.4. Dimostrare il Piccolo Teorema di Fermat, procedendo come segue:

(1) Facendo uso del lemma di Euclide verificare che, se p è primo ed a è coprimo con p , gli interi $a, 2a, 3a, \dots, (p-1)a$ sono a due a due non congruenti \pmod{p} .

(2) Usando (1) ed il lemma di Euclide, dimostrare il Piccolo teorema di Fermat, cioè:

se p è primo ed a è coprimo con p , $a^{p-1} \equiv 1 \pmod{p}$.

2.3.5. Calcolare le ultime due cifre del numero naturale 82^{81} .

2.3.6. Risolvere l'equazione $\overline{39}X = \overline{12}$ in \mathbf{Z}_{603} .

2.3.7. Risolvere l'equazione congruenziale lineare $14X \equiv 10 \pmod{120}$.

4. Omomorfismi tra strutture algebriche

Se A e B sono due insiemi (senza alcuna struttura algebrica), per spostarci da A a B utilizziamo semplicemente le applicazioni da A a B .

Se invece A e B sono due strutture algebriche dello stesso tipo (ad esempio sono due gruppi, due anelli o due K -spazi vettoriali) le applicazioni "significative" che ci fanno di passare da A a B sono quelle che "conservano" l'operazione (o le operazioni) delle due strutture algebriche. Cosa significa? Considerati in A due elementi, possiamo procedere in due modi: o eseguire l'operazione tra essi e considerare poi l'immagine in B del risultato, oppure considerare l'immagine in B dei due elementi ed eseguirne poi l'operazione in B . Diremo che l'applicazione "conserva l'operazione" se otterremo lo stesso risultato. Tali applicazioni sono chiamate *omomorfismi*. Eccone la definizione formale, a partire dai gruppi.

Definizione 1. Siano $(G, *)$ e (H, \cdot) due gruppi e sia $f : G \rightarrow H$ un'applicazione. f è detta *omomorfismo di gruppi* se risulta verificata la seguente condizione:

$$f(a_1) \cdot f(a_2) = f(a_1 * a_2), \quad \forall a_1, a_2 \in G.$$

Nella definizione precedente abbiamo messo in evidenza (usando simboli diversi) il fatto che i due gruppi hanno operazioni diverse [ciò che è ovvio, essendo i due gruppi diversi tra loro]. Nel seguito, per non appesantire le notazioni, indicheremo nello stesso modo le due operazioni, pur continuando ad assumere che possano essere diverse.

Facciamo subito un esempio, forse inatteso, di omomorfismo. Consideriamo i due gruppi (\mathbf{R}^+, \cdot) e $(\mathbf{R}, +)$ e l'applicazione $\log : \mathbf{R}^+ \rightarrow \mathbf{R}$ che associa ad ogni reale positivo x il suo logaritmo naturale $\log(x)$. Poichè, come noto,

$$\log(xy) = \log(x) + \log(y), \quad \forall x, y \in \mathbf{R}^+,$$

allora \log è un omomorfismo tra i due gruppi. Si tratta poi di un omomorfismo biettivo. Infatti l'applicazione \log è invertibile, con inversa la funzione esponenziale:

$$\exp : \mathbf{R} \rightarrow \mathbf{R}^+ \text{ tale che } \exp(x) = e^x, \quad \forall x \in \mathbf{R}.$$

Si noti che anche \exp è un omomorfismo. Infatti $e^{x+y} = e^x e^y$, cioè

$$\exp(x+y) = \exp(x) \exp(y), \quad \forall x, y \in \mathbf{R}.$$

Osservazione 1. Un omomorfismo di gruppi $f : (G, *) \rightarrow (H, \cdot)$ trasforma l'elemento neutro e di G nell'elemento neutro e' di H . Infatti si ha:

$$f(e) = f(e * e) = f(e) \cdot f(e), \quad f(e) = f(e) \cdot e',$$

da cui $f(e) \cdot f(e) = f(e) \cdot e'$. Cancellando $f(e)$, si ottiene $f(e) = e'$.

Inoltre f trasforma il reciproco di un elemento nel reciproco della sua immagine, cioè

$$f(a') = f(a)',$$

dove a' è il reciproco di a e $f(a)'$ è il reciproco di $f(a)$. Infatti si ha:

$$e' = f(e) = f(a * a') = f(a) \cdot f(a')$$

e, ovviamente, $f(a) \cdot f(a)' = e'$. Pertanto $f(a) \cdot f(a') = f(a) \cdot f(a)'$ e quindi, cancellando $f(a)$, $f(a') = f(a)'$.

Con riferimento all'omomorfismo \log sopra considerato, si noti che $\log(1) = 0$ e che, $\forall x > 0$, $\log(x^{-1}) = -\log(x)$.

Diamo ora la definizione di omomorfismo tra anelli e di omomorfismo tra spazi vettoriali. Qui, per non appesantire le notazioni abbiamo denotato con lo stesso simbolo le analoghe operazioni delle due strutture algebriche.

Definizione 2. Siano $(A, +, \cdot)$ e $(B, +, \cdot)$ due anelli e sia $f : A \rightarrow B$ un'applicazione. f è detta

omomorfismo di anelli se risultano verificate le seguenti condizioni:

$$f(a_1) + f(a_2) = f(a_1 + a_2), \quad f(a_1) f(a_2) = f(a_1 a_2), \quad \forall a_1, a_2 \in A.$$

Definizione 3. Siano V e W due K -spazi vettoriali e sia $f : V \rightarrow W$ un'applicazione. f è detta *omomorfismo di K -spazi vettoriali* se risultano verificate le seguenti condizioni:

$$f(\underline{v}_1) + f(\underline{v}_2) = f(\underline{v}_1 + \underline{v}_2), \quad c f(\underline{v}) = f(c \underline{v}), \quad \forall \underline{v}_1, \underline{v}_2, \underline{v} \in V, \quad \forall c \in K.$$

Gli omomorfismi tra K -spazi vettoriali vengono tradizionalmente chiamati *applicazioni lineari* [e ci atterremo, nei capitoli successivi, a questa abitudine].

Definizione 4. Ogni omomorfismo biettivo tra due gruppi (o due anelli o due K -spazi vettoriali) viene chiamato *isomorfismo*. Due gruppi G, H (o due anelli o due spazi vettoriali) sono detti *isomorfi* se esiste tra essi un isomorfismo; in tal caso si scrive $G \cong H$.

Se i due gruppi (anelli o spazi vettoriali) coincidono tra loro, l'isomorfismo prende nome di *automorfismo*.

Un omomorfismo iniettivo prende il nome di *monomorfismo* mentre un omomorfismo suriettivo prende il nome di *epimorfismo*. Infine, un omomorfismo di un gruppo (anello o spazio vettoriale) in sé prende il nome di *endomorfismo*. Un endomorfismo tra spazi vettoriali viene anche chiamato *operatore lineare*.

Ad esempio, per quanto sopra osservato, $(\mathbf{R}, +) \cong (\mathbf{R}^+, \cdot)$. Si noti poi che l'applicazione identica stabilisce sempre un automorfismo di ogni gruppo (anello o spazio vettoriale) in sé. L'inclusione canonica di un sottogruppo in un gruppo è un monomorfismo. Lo stesso è vero per l'inclusione canonica di sottoanelli o sottospazi vettoriali.

Verifichiamo ora che la composizione di omomorfismi è un omomorfismo e che l'applicazione inversa di un isomorfismo è un isomorfismo.

Proposizione 1. (i) Se $f : G \rightarrow G'$ e $g : G' \rightarrow G''$ sono omomorfismi di gruppi, anche l'applicazione $g \circ f : G \rightarrow G''$ è un omomorfismo di gruppi.

(ii) Se $f : G \rightarrow G'$ è un isomorfismo di gruppi, l'applicazione inversa $f^{-1} : G' \rightarrow G$ è un omomorfismo (e quindi un isomorfismo).

Verificare che gli stessi risultati valgono per anelli e spazi vettoriali.

Dim. (i) Utilizzeremo per i tre gruppi la notazione moltiplicativa [ma si intende che le operazioni dei tre gruppi non sono necessariamente le stesse]. Per ogni $x, y \in G$, va verificato che

$$(g \circ f)(x y) = (g \circ f)(x) (g \circ f)(y).$$

Infatti: $(g \circ f)(x y) = g(f(x y)) = g(f(x) f(y)) = g(f(x)) g(f(y)) = (g \circ f)(x) (g \circ f)(y)$.

(ii) Bisogna verificare che

$$f^{-1}(a' b') = f^{-1}(a') \cdot f^{-1}(b'), \quad \forall a', b' \in G'.$$

Sia $a = f^{-1}(a')$, $b = f^{-1}(b')$, $c = f^{-1}(a' b')$. Allora

$$f(a) = a', \quad f(b) = b', \quad f(c) = a' b'.$$

Poiché f è un omomorfismo iniettivo e $f(c) = a' b' = f(a) f(b) = f(ab)$, allora $c = ab$, cioè, come richiesto, $f^{-1}(a' b') = f^{-1}(a') \cdot f^{-1}(b')$.

Se $f : A \rightarrow A'$ è un isomorfismo di anelli, da quanto precede si ha che $f^{-1} : A' \rightarrow A$ è un isomorfismo tra i gruppi additivi $(A', +)$ e $(V, +)$. Con le stesse considerazioni svolte sopra si verifica poi che

$$f^{-1}(a' b') = f^{-1}(a') \cdot f^{-1}(b') \quad \forall a', b' \in A'.$$

Infine, se $f : V \rightarrow V'$ è un isomorfismo di K -spazi vettoriali, f^{-1} è un isomorfismo tra i gruppi $(V', +)$ e $(A', +)$ e va solo verificato che

$$f^{-1}(c \underline{v}') = c f^{-1}(\underline{v}'), \quad \forall c \in K, \quad \forall \underline{v}' \in V'.$$

Se infatti $f^{-1}(\underline{v}') = \underline{v}$ e $f^{-1}(c \underline{v}') = \underline{w}$, bisogna verificare che $c \underline{v} = \underline{w}$. Si ha:

$$f(c \underline{v}) = c f(\underline{v}) = c \underline{v}' = f(\underline{w})$$

e quindi, essendo f iniettiva, $c\underline{v} = \underline{w}$, cioè $cf^{-1}(\underline{v}') = f^{-1}(c\underline{v}')$.

Si noti che la proposizione precedente ci permette di concludere che l'*essere isomorfi* è una relazione di equivalenza nell'insieme dei gruppi (o degli anelli o degli spazi vettoriali).

La *teoria dei gruppi* si occupa di studiare quelle proprietà dei gruppi che sono *invarianti per isomorfismo*, cioè che, valendo per un gruppo, valgono per ogni gruppo ad esso isomorfo. Lo stesso vale per la *teoria degli anelli* e la *teoria degli spazi vettoriali*.

Ogni omomorfismo tra due strutture algebriche definisce due importanti sottostrutture algebriche dello stesso tipo: *nucleo* ed *immagine* dell'omomorfismo. Prima di introdurle ci serve un criterio per riconoscere quando un sottoinsieme di una data struttura algebrica è una sua sottostruttura (dello stesso tipo). Cominciamo dai gruppi [per i quali utilizziamo la notazione astratta (cfr. **Osserv.1.1**)].

Proposizione 2. *Sia (G, \cdot) un gruppo e sia H un sottoinsieme non vuoto di G . Risulta:*

$$H \text{ è un sottogruppo di } G \iff \begin{cases} h_1 h_2 \in H, \forall h_1, h_2 \in H; \\ 1_G \in H; \\ h^{-1} \in H, \forall h \in H. \end{cases}$$

Dim. L'implicazione (\implies) è evidente, perché H è un gruppo, rispetto all'operazione \cdot (ristretta ad H).

Proviamo l'implicazione (\iff). Dalla prima delle tre condizioni segue che in H è definita la stessa operazione di G . Inoltre tale operazione ammette elemento neutro ed inverso di ogni elemento, in base alle altre due condizioni. L'associatività vale in H , in quanto vale in G . Pertanto H è un sottogruppo di G .

Osservazione 2. Possiamo riscrivere le tre condizioni di **Prop. 2** nella seguente "forma compatta"

$$HH \subseteq H; \quad 1_G \in H; \quad H^{-1} \subseteq H.$$

Si osservi poi che se l'operazione del gruppo G è indicata in forma additiva (cioè con $+$), le tre condizioni della precedente proposizione diventano:

$$h_1 + h_2 \in H, \quad \forall h_1, h_2 \in H; \quad 0_G \in H; \quad -h \in H, \quad \forall h \in H.$$

In forma compatta quindi si scrivono nella forma:

$$H + H \subseteq H; \quad 0_G \in H; \quad -H \subseteq H.$$

Le tre condizioni di **Prop. 2** possono essere riunificate in un'unica condizione.

Proposizione 3. *Sia (G, \cdot) un gruppo e sia H un sottoinsieme non vuoto di G . Risulta:*

$$H \text{ è un sottogruppo di } G \iff h_1 h_2^{-1} \in H, \quad \forall h_1, h_2 \in H.$$

[ovvero, in forma compatta, $HH^{-1} \subseteq H$].

Dim. L'implicazione (\implies) è evidente, perché H è un gruppo. Proviamo l'implicazione (\iff). Scelto $h \in H$, allora $hh^{-1} \in H$, cioè $1_G \in H$. Per ogni $h \in H$, $1_G h^{-1} \in H$, cioè $h^{-1} \in H$. Infine, siano $h_1, h_2 \in H$: allora $h_2^{-1} \in H$ e dunque $h_1(h_2^{-1})^{-1} \in H$, cioè $h_1 h_2 \in H$.

Osservazione 3. Per un gruppo $(G, +)$ (espresso cioè in notazione additiva) la **Prop. 3** si enuncia nella forma:

$$H \text{ è un sottogruppo di } (G, +) \iff h_1 - h_2 \in H, \quad \forall h_1, h_2 \in H \iff H - H \subseteq H.$$

Proposizione 4. *Sia $(A, +, \cdot)$ un anello e sia B un sottoinsieme non vuoto di A . Risulta:*

$$B \text{ è un sottoanello di } A \iff b_1 - b_2 \in B, \quad b_1 b_2 \in B, \quad \forall b_1, b_2 \in B.$$

[ovvero, in forma compatta, $\iff B - B \in B$ e $BB \subseteq B$].

Dim. (\implies) L'implicazione è evidente. Verifichiamo l'implicazione (\impliedby). La prima condizione equivale a dire che $(B, +)$ è un sottogruppo di $(A, +)$; la seconda che in B è definita l'operazione di moltiplicazione (indotta da quella in A). La proprietà associativa e le distributive valgono in B in quanto già valevano in A .

Proposizione 5. Sia V un K -spazio vettoriale e sia W un sottoinsieme non vuoto di V . Risulta:

$$W \text{ è un } K\text{-sottospazio vettoriale di } V \iff \underline{w}_1 - \underline{w}_2 \in W, a\underline{w} \in W, \forall \underline{w}_1, \underline{w}_2, \underline{w} \in W, \forall a \in K.$$

Dim. La prima condizione equivale a dire che $(W, +)$ è un sottogruppo di $(V, +)$; la seconda che in W è definita la moltiplicazione per uno scalare (indotta da quella in V). Gli assiomi di spazio vettoriale valgono in W in quanto valgono in V .

Osservazione 4. Si osserva subito che le due condizioni della proposizione precedente possono essere riunite in un'unica condizione:

$$W \text{ è un } K\text{-sottospazio vettoriale di } V \iff a_1\underline{w}_1 + a_2\underline{w}_2 \in W, \forall \underline{w}_1, \underline{w}_2 \in W, \forall a_1, a_2 \in K.$$

Ora veniamo alla definizione di *nucleo* ed *immagine* di un omomorfismo. Al solito, cominciamo dai gruppi.

Definizione 5. Siano (G, \cdot) e (G', \cdot) due gruppi e sia $f : G \rightarrow G'$ un omomorfismo. Si chiama *nucleo* di f la controimmagine (in G) dell'elemento neutro di G' , cioè l'insieme

$$f^{-1}(1_{G'}) = \{a \in G \mid f(a) = 1_{G'}\},$$

che viene tradizionalmente denotato $\text{Ker}(f)$ (o $\text{Ker } f$) [da "kernel" = nucleo].

[Si noti che, se G' è espresso in notazione additiva, $\text{Ker}(f) = f^{-1}(0_{G'}) = \{a \in G \mid f(a) = 0_{G'}\}$].

Si chiama *immagine* di f l'immagine insiemistica $\text{Im}(f)$ (o $\text{Im } f$), cioè l'insieme

$$\text{Im}(f) = \{f(a), \forall a \in G\}.$$

Proveremo ora che $\text{Ker}(f)$ è un sottogruppo di G mentre $\text{Im}(f)$ è un sottogruppo di G' . Poi verificheremo che il nucleo è collegato all'iniettività dell'omomorfismo.

Proposizione 6. Se $f : (G, \cdot) \rightarrow (G', \cdot)$ è un omomorfismo di gruppi, $\text{Ker}(f)$ è un sottogruppo di (G, \cdot) e $\text{Im}(f)$ è un sottogruppo di (G', \cdot) .

Dim. Dalla **Prop. 2**, per dimostrare che $\text{Ker}(f)$ è un sottogruppo di G basta verificare che

$$ab^{-1} \in \text{Ker}(f), \forall a, b \in \text{Ker}(f).$$

Infatti, se $f(a) = f(b) = 1_{G'}$, allora $f(a b^{-1}) = f(a) f(b^{-1}) = f(a) f(b)^{-1} = 1_{G'} (1_{G'})^{-1} = 1_{G'}$.

Analogamente, per provare che $\text{Im}(f)$ è un sottogruppo di G' basta verificare che,

$$f(a) f(b)^{-1} \in \text{Im}(f), \forall f(a), f(b) \in \text{Im}(f).$$

Infatti $f(a) f(b)^{-1} = f(a) f(b^{-1}) = f(a b^{-1})$ e ovviamente $f(a b^{-1}) \in \text{Im}(f)$.

Proposizione 7. Sia $f : (G, \cdot) \rightarrow (G', \cdot)$ un omomorfismo di gruppi. Risulta:

$$f \text{ è iniettivo (cioè un monomorfismo)} \iff \text{Ker}(f) = \{1_G\}.$$

$$f \text{ è suriettivo (cioè un epimorfismo)} \iff \text{Im}(f) = G'.$$

Dim. La seconda affermazione è una definizione insiemistica. Dimostriamo la prima.

Se f è iniettivo e $a \in \text{Ker}(f)$, $f(a) = 1_{G'} = f(1_G)$. Ne segue che $a = 1_G$, cioè $\text{Ker}(f) = \{1_G\}$. Viceversa, sia $\text{Ker}(f) = \{1_G\}$ e sia $f(a) = f(b)$. Allora $1_{G'} = f(a) f(b)^{-1} = f(a) f(b^{-1}) = f(ab^{-1})$.

Dunque $ab^{-1} \in \text{Ker}(f)$, da cui $ab^{-1} = 1_G$ e quindi $a = b$.

Per anelli e spazi vettoriali, le due proposizioni precedenti continuano a sussistere, con analoghe dimostrazioni. Esplicitiamo la definizione di nucleo e immagine nei due casi.

Sia $f : (A, +, \cdot) \rightarrow (B, +, \cdot)$ un omomorfismo di anelli. Si chiamano rispettivamente *nucleo di f* ed *immagine di f* gli insiemi

$$\text{Ker}(f) = \{a \in A \mid f(a) = 0_B\}, \quad \text{Im}(f) = \{f(a), \forall a \in A\}.$$

Lasciamo allo studente la verifica dei seguenti fatti:

- $\text{Ker}(f)$ è un sottoanello di A ; $\text{Im}(f)$ è un sottoanello di B ;
- f è un omomorfismo iniettivo $\iff \text{Ker}(f) = \{0_A\}$;
- f è un omomorfismo suriettivo $\iff \text{Im}(f) = B$.

Analogamente, sia $f : V \rightarrow W$ un omomorfismo di K -spazi vettoriali. Si chiamano rispettivamente *nucleo di f* ed *immagine di f* gli insiemi

$$\text{Ker}(f) = \{\underline{v} \in V \mid f(\underline{v}) = 0_W\}, \quad \text{Im}(f) = \{f(\underline{v}), \forall \underline{v} \in V\}.$$

Si verifichi che:

- $\text{Ker}(f)$ è un K -sottospazio vettoriale di V ;
- $\text{Im}(f)$ è un K -sottospazio vettoriale di W ;
- f è un omomorfismo iniettivo $\iff \text{Ker}(f) = \{0_V\}$;
- f è un omomorfismo suriettivo $\iff \text{Im}(f) = W$.

ESERCIZI PROPOSTI

2.4.1. Sia $\mathfrak{L} = \{A \in \mathfrak{M}_3(\mathbf{R}) \mid a_{12} = a_{13} = a_{21} = a_{31} = 0\}$. Verificare se \mathfrak{L} è un \mathbf{R} -sottospazio vettoriale di $\mathfrak{M}_3(\mathbf{R})$ e se è un sottoanello di $\mathfrak{M}_3(\mathbf{R})$.

2.4.2. Sia $n \geq 2$ e sia $f : \mathfrak{M}_n(K) \rightarrow \mathfrak{M}_{n-1}(K)$ l'applicazione che ad ogni matrice $A \in \mathfrak{M}_n(K)$ associa la matrice ottenuta da A privandola degli elementi dell'ultima riga e dell'ultima colonna.

Verificare che f è un omomorfismo di K -spazi vettoriali e determinarne nucleo ed immagine. È vero che f è un omomorfismo di anelli?

2.4.3. Sia τ una fissata permutazione di S_n . Sia $f : S_n \rightarrow S_n$ l'applicazione così definita:

$$f(\sigma) = \tau \sigma \tau^{-1}, \quad \forall \sigma \in S_n.$$

Verificare che f è un automorfismo di S_n [è detto *automorfismo di coniugio* (relativo a τ)].

2.4.4. Sia $n \geq 2$ e sia $\pi : \mathbf{Z} \rightarrow \mathbf{Z}_n$ l'applicazione così definita:

$$\pi(a) = \bar{a}, \quad \forall a \in \mathbf{Z}.$$

Verificare che π è un omomorfismo di anelli e determinare $\text{Ker}(\pi)$. π è detta *proiezione canonica* di \mathbf{Z} sull'anello quoziante \mathbf{Z}_n .

2.4.5. L'applicazione $f : \mathbf{R} \rightarrow \mathfrak{M}_2(\mathbf{R})$ tale che

$$\varphi(t) = \begin{pmatrix} t & 0 \\ 0 & t \end{pmatrix}, \quad \forall t \in \mathbf{R},$$

è un omomorfismo di spazi vettoriali o di anelli? In caso affermativo, calcolarne nucleo ed immagine.

2.4.6. Sia $f : V \rightarrow W$ un omomorfismo di K -spazi vettoriali.

- (i) Se W_1 è un sottospazio vettoriale di W , verificare che $f^{-1}(W_1)$ è un sottospazio vettoriale di V .
- (ii) Se V_1 è un sottospazio vettoriale di V , verificare che $f(V_1)$ è un sottospazio vettoriale di W .

2.4.7. Sia $\partial : K[X] \rightarrow K[X]$ l'applicazione di derivazione, così definita: $\forall P = \sum_{i=0}^n a_i X^i \in K[X]$,

$$\partial P = a_1 + 2a_2 X + 3a_3 X^2 + \dots + n a_n X^{n-1}$$

[se $K = \mathbf{R}$, ∂P è l'usuale derivata prima del polinomio P].

- (i) Verificare che ∂ è un omomorfismo di K -spazi vettoriali ma non un omomorfismo di anelli.

(ii) Posto $K = \mathbf{R}$, calcolare $\text{Im } \partial$.

(iii) Posto $K = \mathbf{R}$, calcolare $\text{Ker } \partial$.

2.4.8. In $\mathfrak{M}_2(\mathbf{R})$ si consideri il sottoinsieme

$$\mathcal{H} = \{A \in \mathfrak{M}_2(\mathbf{R}) \mid (A)_{11} = 0\}$$

[si ricorda che $(A)_{11}$ denota l'elemento a_{11} della matrice A]. Verificare che \mathcal{H} è un \mathbf{R} -sottospazio vettoriale di $\mathfrak{M}_2(\mathbf{R})$ e che è nucleo di un opportuno omomorfismo $\varphi : \mathfrak{M}_2(\mathbf{R}) \rightarrow \mathbf{R}$.

5. Generalità sugli spazi vettoriali

Nei precedenti paragrafi di questo capitolo abbiamo definito il concetto di spazio vettoriale, di sottospazio vettoriale e di omomorfismo tra spazi vettoriali. Abbiamo inoltre presentato alcuni esempi di spazi vettoriali. Ora presenteremo altri importanti concetti relativi ad uno spazio vettoriale, cominciando con alcune semplici regole di calcolo all'interno di uno spazio vettoriale, che discendono con facilità dalla definizione.

Proposizione 1. *Sia V un K -spazio vettoriale. Risulta:*

- (i) $0\underline{v} = \underline{0}$, $\forall \underline{v} \in V$.
- (ii) $(-c)\underline{v} = -(c\underline{v})$, $\forall c \in K$, $\forall \underline{v} \in V$.
- (iii) $c\underline{0} = \underline{0}$, $\forall c \in K$.
- (iv) se $c\underline{v} = \underline{0}$, allora $c = 0$ oppure $\underline{v} = \underline{0}$.

Dim. (i) $0\underline{v} = (0+0)\underline{v} = 0\underline{v} + 0\underline{v}$. Ne segue che $\underline{0} + 0\underline{v} = 0\underline{v} + 0\underline{v}$. Cancellando, $\underline{0} = 0\underline{v}$.

(ii) $(-c)\underline{v} + c\underline{v} = (-c+c)\underline{v} = 0\underline{v} = \underline{0}$. Allora $(-c)\underline{v}$ è l'opposto di $c\underline{v}$.

(iii) $c\underline{0} = c(\underline{0} + \underline{0}) = c\underline{0} + c\underline{0}$. Ne segue che $\underline{0} + c\underline{0} = c\underline{0} + c\underline{0}$. Cancellando, $\underline{0} = c\underline{0}$.

(iv) Se $c \neq 0$, allora $c^{-1} \in K$. Quindi $\underline{v} = 1\underline{v} = (c^{-1}c)\underline{v} = c^{-1}(c\underline{v}) = c^{-1}\underline{0} = \underline{0}$.

Sappiamo che è possibile eseguire in un K -spazio vettoriale V somme e moltiplicazioni per scalari. Se "combiniamo" tra loro tali operazioni, otteniamo la definizione di *combinazione lineare di vettori*.

Definizione 1. Assegnati in V i vettori $\underline{v}_1, \underline{v}_2, \dots, \underline{v}_n$ e in K gli scalari c_1, c_2, \dots, c_n , si chiama *combinazione lineare dei vettori $\underline{v}_1, \underline{v}_2, \dots, \underline{v}_n$, a coefficienti c_1, c_2, \dots, c_n* , il vettore

$$c_1\underline{v}_1 + c_2\underline{v}_2 + \dots + c_n\underline{v}_n \in V.$$

Assegnato un vettore $\underline{v} \in V$, diremo che \underline{v} è *combinazione lineare dei vettori $\underline{v}_1, \underline{v}_2, \dots, \underline{v}_n$* se esistono $c_1, c_2, \dots, c_n \in K$ tali che $\underline{v} = c_1\underline{v}_1 + c_2\underline{v}_2 + \dots + c_n\underline{v}_n$.

Il vettore nullo $\underline{0}$ è sempre combinazione lineare dei vettori $\underline{v}_1, \underline{v}_2, \dots, \underline{v}_n$ (qualunque essi siano). Infatti $0\underline{v}_1 + 0\underline{v}_2 + \dots + 0\underline{v}_n = \underline{0}$. La combinazione lineare $0\underline{v}_1 + 0\underline{v}_2 + \dots + 0\underline{v}_n$ è detta *combinazione lineare banale di $\underline{v}_1, \underline{v}_2, \dots, \underline{v}_n$* .

Può talvolta avvenire che il vettore nullo $\underline{0}$ sia ottenibile anche come combinazione lineare **non banale** [cioè a coefficienti non tutti nulli] di $\underline{v}_1, \underline{v}_2, \dots, \underline{v}_n$. In tal caso tali vettori sono detti *linearmente dipendenti*.

Ad esempio, consideriamo in \mathbf{R}^2 i due vettori $\underline{x} = (1, -2)$, $\underline{y} = (-2, 4)$. Si ha:

$$2\underline{x} + \underline{y} = 2(1, -2) + (-2, 4) = (2, -4) + (-2, 4) = (0, 0) = \underline{0}.$$

Dunque $\underline{x}, \underline{y}$ sono linearmente dipendenti. Formalizziamo queste considerazioni in una definizione.

Definizione 2. Assegnati i vettori $\underline{v}_1, \underline{v}_2, \dots, \underline{v}_n \in V$, diciamo che essi sono *linearmente dipendenti* se esistono $c_1, c_2, \dots, c_n \in K$, non tutti nulli, tali che $c_1\underline{v}_1 + c_2\underline{v}_2 + \dots + c_n\underline{v}_n = \underline{0}$.

In caso contrario i vettori $\underline{v}_1, \underline{v}_2, \dots, \underline{v}_n \in V$, sono detti *linearmente indipendenti*. Si ha quindi: $\underline{v}_1, \underline{v}_2, \dots, \underline{v}_n \in V$ sono linearmente indipendenti se risulta:

$$c_1\underline{v}_1 + c_2\underline{v}_2 + \dots + c_n\underline{v}_n = \underline{0} \implies c_1 = c_2 = \dots = c_n = 0$$

[cioè: l'unica loro combinazione lineare che fornisce il vettore nullo è quella banale].

Ad esempio i due vettori $\underline{x} = (1, -3)$, $\underline{y} = (-2, 0) \in \mathbf{R}^2$ sono linearmente indipendenti. Come possiamo verificarlo?

Supponiamo che sia $a\underline{x} + b\underline{y} = \underline{0}$, con $a, b \in \mathbf{R}$. Allora

$$(a, -3a) + (-2b, 0) = (a - 2b, -3a) = (0, 0) \text{ e dunque } a - 2b = 0, -3a = 0.$$

Dobbiamo quindi verificare se esistono o meno due elementi $a, b \in \mathbf{R}$, non entrambi nulli, che risolvono il seguente *sistema di equazioni lineari* (cioè di primo grado) nelle variabili a, b :

$$\begin{cases} a - 2b = 0 \\ -3a = 0. \end{cases}$$

Dalla seconda equazione, $a = 0$. Sostituendo $a = 0$ nella prima, segue che anche $b = 0$. Si conclude che $\underline{x}, \underline{y}$ sono linearmente indipendenti.

Dalle considerazioni appena svolte si deduce che, per stabilire se vettori di \mathbf{R}^n sono linearmente dipendenti o indipendenti, bisogna saper risolvere sistemi di equazioni lineari (in più variabili).

Proposizione 2. Assegnati i vettori $\underline{v}_1, \underline{v}_2, \dots, \underline{v}_n \in V$, risulta:

$$\underline{v}_1, \underline{v}_2, \dots, \underline{v}_n \text{ sono linearmente dipendenti} \iff \text{almeno uno di essi è combinazione lineare degli altri.}$$

Dim. (\Rightarrow) Sia $c_1 \underline{v}_1 + c_2 \underline{v}_2 + \dots + c_n \underline{v}_n = \underline{0}$, con c_1, c_2, \dots, c_n non tutti nulli. Se ad esempio $c_1 \neq 0$, allora $\underline{v}_1 = -\frac{c_2}{c_1} \underline{v}_2 - \dots - \frac{c_n}{c_1} \underline{v}_n$ e dunque \underline{v}_1 è combinazione lineare di $\underline{v}_2, \dots, \underline{v}_n$.

(\Leftarrow) Sia ad esempio \underline{v}_1 combinazione lineare di $\underline{v}_2, \dots, \underline{v}_n$, cioè $\underline{v}_1 = c_2 \underline{v}_2 + \dots + c_n \underline{v}_n$. Allora

$$\underline{v}_1 - c_2 \underline{v}_2 - \dots - c_n \underline{v}_n = \underline{0}.$$

Tale combinazione lineare di $\underline{0}$ è non banale e dunque $\underline{v}_1, \underline{v}_2, \dots, \underline{v}_n$ sono linearmente dipendenti.

Osservazione 1. (i) Un singolo vettore \underline{v} è linearmente dipendente \iff è il vettore nullo [infatti $c\underline{v} = \underline{0}$, con $c \neq 0$ implica $\underline{v} = \underline{0}$].

(ii) Se n vettori sono linearmente dipendenti non è detto che *ciascuno di essi* sia combinazione lineare degli altri. Ad esempio, per ogni vettore $\underline{v} \neq \underline{0}$, i due vettori $\underline{0}, \underline{v}$ sono linearmente dipendenti [infatti $1\underline{0} + 0\underline{v} = \underline{0}$], ma \underline{v} non è combinazione lineare di $\underline{0}$ [in quanto $\underline{v} \neq c\underline{0}$, $\forall c \in \mathbf{R}$], mentre $\underline{0}$ lo è di \underline{v} [infatti $\underline{0} = 0\underline{v}$].

(iii) La **Prop. 2**, letta in forma contrappositiva, diventa: *i vettori $\underline{v}_1, \underline{v}_2, \dots, \underline{v}_n$ sono linearmente indipendenti \iff nessuno di essi è combinazione lineare degli altri.*

(iv) È evidente che, se $\underline{v}_1, \dots, \underline{v}_n$ sono linearmente indipendenti, ogni sottoinsieme non vuoto di $\{\underline{v}_1, \dots, \underline{v}_n\}$ è formato da vettori linearmente indipendenti. Viceversa, se un sottoinsieme proprio di $\{\underline{v}_1, \dots, \underline{v}_n\}$ è formato da vettori linearmente dipendenti, i vettori $\underline{v}_1, \dots, \underline{v}_n$ sono linearmente dipendenti. È altresì evidente che, se W è un sottospazio vettoriale di V e $\underline{w}_1, \dots, \underline{w}_n \in W$, risulta:

$$\underline{w}_1, \dots, \underline{w}_n \text{ sono linearmente indipendenti in } W \iff \text{lo sono in } V.$$

Definizione 3. Assegnati in un *K-spazio vettoriale* V i vettori $\underline{u}_1, \underline{u}_2, \dots, \underline{u}_t$, denoteremo con $\langle \underline{u}_1, \underline{u}_2, \dots, \underline{u}_t \rangle$ l'insieme di tutte le possibili loro combinazioni lineari, cioè

$$\langle \underline{u}_1, \underline{u}_2, \dots, \underline{u}_t \rangle = \left\{ \sum_{i=1}^t a_i \underline{u}_i, \quad \forall a_i \in K \right\}.$$

Tale insieme è detto *sottospazio generato* da $\underline{u}_1, \underline{u}_2, \dots, \underline{u}_t$.

Si verifica facilmente, usando l'**Osserv. 4** del paragrafo precedente, che $\langle \underline{u}_1, \underline{u}_2, \dots, \underline{u}_t \rangle$ è un sottospazio vettoriale di V e che è il più piccolo sottospazio vettoriale di V contenente $\underline{u}_1, \underline{u}_2, \dots, \underline{u}_t$.

Ad esempio consideriamo i tre vettori $\underline{x} = (1, -3, 0)$, $\underline{y} = (2, 1, 1)$, $\underline{z} = (1, 4, 1) \in \mathbf{R}^3$ e denotiamo con W il sottospazio vettoriale da essi generato. Dunque

$$W = \langle \underline{x}, \underline{y}, \underline{z} \rangle = \{a \underline{x} + b \underline{y} + c \underline{z}, \quad \forall a, b, c \in \mathbf{R}\}.$$

Consideriamo ora un vettore $\underline{v} \in \mathbf{R}^3$, ad esempio il vettore $\underline{v} = (3, 5, 2)$. Ci chiediamo se tale vettore appartiene a W . Risulta:

$$\underline{v} \in W \iff \exists a, b, c \in \mathbf{R} \text{ tali che } a\underline{x} + b\underline{y} + c\underline{z} = \underline{v}.$$

Uguagliando le tre componenti dei due vettori, si ottiene:

$$\begin{cases} a + 2b + c = 3 \\ -3a + b + 4c = 5 \\ b + c = 2. \end{cases}$$

Dunque $\underline{v} \in W \iff$ il precedente sistema di tre equazioni lineari (nelle incognite a, b, c) ammette soluzioni. Lo studente probabilmente sa come risolvere tale sistema (o lo imparerà in questo corso). Per il momento può verificare che ad esempio la terna $(a, b, c) = (3, -2, 4)$ è soluzione di tale sistema e dunque che $\underline{v} = 3\underline{x} - 2\underline{y} + 4\underline{z}$. Pertanto $\underline{v} \in W$.

Dalle considerazioni appena svolte si deduce che anche per stabilire se un vettore appartiene al sottospazio generato da vettori di \mathbf{R}^n bisogna saper risolvere sistemi di equazioni lineari (in più variabili).

La risoluzione di sistemi di equazioni lineari (di cui ci occuperemo nel prossimo capitolo) è lo strumento tecnico fondamentale nello studio degli spazi vettoriali \mathbf{R}^n (o più generalmente dei K -spazi vettoriali di "dimensione finita").

Definizione 4. Sia V un K -spazio vettoriale. I vettori $\underline{u}_1, \underline{u}_2, \dots, \underline{u}_t \in V$ sono detti *sistema di generatori di V* se il sottospazio vettoriale da essi generato coincide con V , cioè se

$$\langle \underline{u}_1, \underline{u}_2, \dots, \underline{u}_t \rangle = V.$$

[Ovviamente i vettori $\underline{u}_1, \underline{u}_2, \dots, \underline{u}_t$ sono un sistema di generatori del sottospazio vettoriale $\langle \underline{u}_1, \underline{u}_2, \dots, \underline{u}_t \rangle$ da essi generato].

Osservazione 2. (i) Ad esempio, i due vettori $\underline{e}_1 = (1, 0), \underline{e}_2 = (0, 1) \in K^2$ sono un sistema di generatori di K^2 . Infatti, $\forall (a, b) \in K^2$, risulta: $(a, b) = a(1, 0) + b(0, 1) = a\underline{e}_1 + b\underline{e}_2$. Dunque $\langle \underline{e}_1, \underline{e}_2 \rangle = K^2$. Più in generale, gli n vettori

$$\underline{e}_1 = (1, 0, 0, \dots, 0), \underline{e}_2 = (0, 1, 0, \dots, 0), \dots, \underline{e}_n = (0, 0, \dots, 0, 1) \in K^n$$

formano un sistema di generatori di K^n .

(ii) Esistono spazi vettoriali che **non** ammettono sistemi di generatori formati da un numero **finito** di vettori. Di essi non ci occuperemo ed i nostri spazi vettoriali saranno quindi, come si dice, *finitamente generati*. Vogliamo però indicare due esempi di spazi vettoriali non finitamente generati.

(1) Nel precedente paragrafo 1 abbiamo considerato l'anello dei polinomi $K[X]$, a valori su un campo K , ed osservato che è anche un K -spazio vettoriale.

Se fosse finitamente generato, esisterebbero $P_1, \dots, P_n \in K[X]$ tali che $\langle P_1, \dots, P_n \rangle = K[X]$. Per ogni $i = 1, \dots, n$, denotiamo con d_i il grado di P_i e poniamo $d := \max\{d_1, \dots, d_n\}$. Ogni polinomio di $\langle P_1, \dots, P_n \rangle$ è della forma $\sum_{i=1}^n c_i P_i$. Poiché il suo grado è $\leq d$, ad esempio $X^{d+1} \notin \langle P_1, \dots, P_n \rangle$. Ne segue che $\langle P_1, \dots, P_n \rangle \neq K[X]$.

(2) Sia I un sottoinsieme non vuoto di \mathbf{R} e sia \mathfrak{F}_I l'insieme di tutte le funzioni da I a \mathbf{R} . Lasciamo come esercizio la verifica che \mathfrak{F}_I è un \mathbf{R} -spazio vettoriale rispetto alla seguente operazione di somma:

$$(f + g)(x) = f(x) + g(x), \quad \forall x \in I, \quad \forall f, g \in \mathfrak{F}_I,$$

ed alla seguente moltiplicazione per uno scalare:

$$(cf)(x) = cf(x), \quad \forall x \in I, \quad \forall f \in \mathfrak{F}_I, \quad \forall c \in \mathbf{R}.$$

Si potrebbe verificare che, se I è un insieme infinito, \mathfrak{F}_I non è finitamente generato.

(iii) I vettori *generatori* di uno spazio vettoriale (finitamente generato) non sono ovviamente unici, ma possono facilmente essere sostituiti da altri, che generino lo stesso spazio vettoriale. Ad esempio possiamo considerare come sistema di generatori di \mathbf{R}^2 anche i due vettori $\underline{x} = (3, 5), \underline{y} = (-1, 2)$. Va verificato che ogni vettore $(a, b) \in \mathbf{R}^2$ è combinazione lineare di $\underline{x}, \underline{y}$ e cioè che esistono $r, s \in \mathbf{R}$ tali che $r\underline{x} + s\underline{y} = (a, b)$, ovvero che il sistema di equazioni lineari

$$\begin{cases} 3r - s = a \\ 5r + 2s = b \end{cases}$$

(nelle incognite r, s) ammette soluzione. Lo studente può verificare che tale sistema ha come (unica) soluzione la coppia $(r, s) = \frac{1}{11}(2a + b, 3b - 5a)$.

(iv) Si noti che, se aggiungiamo ad un sistema di generatori di V altri vettori, otteniamo ancora un sistema di generatori. Ma ovviamente è più utile andare nella direzione opposta, cioè "rimpicciolire" un sistema di generatori, sfrondandolo da vettori superflui.

Ad esempio riconsideriamo i tre vettori $\underline{x} = (1, -3, 0), \underline{y} = (2, 1, 1), \underline{z} = (1, 4, 1) \in \mathbf{R}^3$ (già considerati in un esempio precedente). Proviamo a porre $a\underline{x} + b\underline{y} = \underline{z}$ e a risolvere il sistema di equazioni lineari che se ne deduce (uguagliando le tre componenti dei vettori)

$$\begin{cases} a + 2b = 1 \\ -3a + b = 4 \\ b = 1. \end{cases}$$

Osserviamo che tale sistema ammette soluzione $(a, b) = (-1, 1)$ e dunque che $\underline{z} = -\underline{x} + \underline{y}$. Allora è evidente che \underline{z} è superfluo come generatore di $\langle \underline{x}, \underline{y}, \underline{z} \rangle$, in quanto ogni vettore $a\underline{x} + b\underline{y} + c\underline{z}$ si riscrive nella forma $a\underline{x} + b\underline{y} + c(-\underline{x} + \underline{y}) = (a - c)\underline{x} + (b + c)\underline{y}$. Dunque $\langle \underline{x}, \underline{y}, \underline{z} \rangle = \langle \underline{x}, \underline{y} \rangle$.

Non è possibile rimpicciolire ulteriormente il sistema di generatori. Infatti si verifica che $\underline{x} \notin \langle \underline{y} \rangle$ e $\underline{y} \notin \langle \underline{x} \rangle$ (cioè, in base a **Osserv. 1(iii)**, che i due vettori $\underline{x}, \underline{y}$ sono linearmente indipendenti).

Le considerazioni svolte nel punto (iv) dell'osservazione precedente ci hanno portato a determinare sistemi di generatori formati da vettori linearmente indipendenti. Siamo così arrivati alla definizione di *base* di uno spazio vettoriale.

Definizione 5. Sia V un K -spazio vettoriale. I vettori $\underline{u}_1, \underline{u}_2, \dots, \underline{u}_n \in V$ sono detti *base di V* se sono linearmente indipendenti e formano un sistema di generatori di V .

È evidente che i vettori

$$\underline{e}_1 = (1, 0, 0, \dots, 0), \underline{e}_2 = (0, 1, 0, \dots, 0), \dots, \underline{e}_n = (0, 0, \dots, 0, 1) \in K^n$$

formano una base di K^n . Infatti già abbiamo osservato che sono un sistema di generatori. Verifichiamo quindi che sono linearmente indipendenti.

Sia infatti $a_1\underline{e}_1 + a_2\underline{e}_2 + \dots + a_n\underline{e}_n = \underline{0}$. Poiché $a_1\underline{e}_1 + a_2\underline{e}_2 + \dots + a_n\underline{e}_n = (a_1, a_2, \dots, a_n)$, allora $a_1 = a_2 = \dots = a_n = 0$.

La base $\{\underline{e}_1, \underline{e}_2, \dots, \underline{e}_n\}$ è detta *base canonica di K^n* . L'aggettivo "canonica" significa che tale base è la più "naturale" che si possa considerare in K^n [per un motivo che sarà chiarito nel prossimo esempio].

Proposizione 3. Sia V un K -spazio vettoriale e siano $\underline{u}_1, \dots, \underline{u}_n \in V$. Risulta:

$\{\underline{u}_1, \dots, \underline{u}_n\}$ è una base di $V \iff$ ogni vettore $\underline{v} \in V$ si scrive **in modo unico** come combinazione lineare dei vettori $\underline{u}_1, \dots, \underline{u}_n$ [cioè $\exists! (c_1, \dots, c_n) \in K^n$ tale che $\underline{v} = \sum_{i=1}^n c_i \underline{u}_i$].

Dim. (\implies) Poiché $\{\underline{u}_1, \dots, \underline{u}_n\}$ è un sistema di generatori di V , risulta: $\underline{v} = \sum_{i=1}^n c_i \underline{u}_i$ (per opportuni $c_1, \dots, c_n \in K$). Dimostriamo che tali coefficienti sono unici. Se risultasse anche $\underline{v} = \sum_{i=1}^n d_i \underline{u}_i$, con $d_1, \dots, d_n \in K$, si avrebbe

$$\sum_{i=1}^n (c_i - d_i) \underline{u}_i = \underline{0}.$$

Poiché $\underline{u}_1, \dots, \underline{u}_n$ sono linearmente indipendenti, $c_i - d_i = 0, \forall i = 1, \dots, n$. Dunque la scrittura è unica.

(\Leftarrow) Per ipotesi ogni vettore $\underline{v} \in V$ può essere scritto nella forma: $\underline{v} = \sum_{i=1}^n c_i \underline{u}_i$. Dunque $\{\underline{u}_1, \dots, \underline{u}_n\}$ è un sistema di generatori di V . Se poi $\sum_{i=1}^n c_i \underline{u}_i = \underline{0}$, allora $\sum_{i=1}^n c_i \underline{u}_i = \sum_{i=1}^n 0 \underline{u}_i$ e dunque, per l'unicità della combinazione lineare, $c_1 = \dots = c_n = 0$. Pertanto i vettori $\underline{u}_1, \dots, \underline{u}_n$ sono linearmente indipendenti.

Definizione 6. Sia V un K -spazio vettoriale e sia $\{\underline{u}_1, \dots, \underline{u}_n\}$ una sua base. Per ogni vettore $\underline{v} \in V$, se $\underline{v} = \sum_{i=1}^n c_i \underline{u}_i$, la n -pla (c_1, \dots, c_n) è detta n -pla delle coordinate di \underline{v} rispetto alla base $\{\underline{u}_1, \dots, \underline{u}_n\}$.

Un esempio. Cosideriamo in \mathbf{R}^3 i tre vettori $\underline{u}_1 = (1, 0, 0)$, $\underline{u}_2 = (1, 1, 0)$, $\underline{u}_3 = (1, 1, 1)$. Supponiamo per il momento di aver dimostrato che tali vettori formano una base di \mathbf{R}^3 . Ci chiediamo quale sia la terna delle coordinate del vettore $\underline{v} = (1, 2, 3)$ rispetto a tale base. Si tratta di determinare la terna $(c_1, c_2, c_3) \in \mathbf{R}^3$ tale che $c_1 \underline{u}_1 + c_2 \underline{u}_2 + c_3 \underline{u}_3 = \underline{v}$. Si ottiene il sistema di equazioni lineari

$$\begin{cases} c_1 + c_2 + c_3 = 1 \\ c_2 + c_3 = 2 \\ c_3 = 3. \end{cases}$$

Probabilmente il lettore sa risolvere tale sistema ("a scala") o lo imparerà nel prossimo capitolo. Si può verificare comunque che la terna cercata è $(-1, -1, 3)$. Dunque le coordinate del vettore \underline{v} in base $\{\underline{u}_1, \underline{u}_2, \underline{u}_3\}$ sono $(-1, -1, 3)$ [ben diverse quindi dalle tre componenti 1, 2, 3 di \underline{v}].

Se invece consideriamo la base canonica $\{\underline{e}_1, \underline{e}_2, \underline{e}_3\}$ di \mathbf{R}^3 , si verifica subito che le coordinate di $\underline{v} = (1, 2, 3)$ sono proprio $(1, 2, 3)$. Dunque, rispetto alla base canonica, le coordinate di un vettore di \mathbf{R}^3 sono esattamente le sue componenti. Per questo motivo $\{\underline{e}_1, \underline{e}_2, \underline{e}_3\}$ si chiama "base canonica".

Ci resta il problema di verificare che $\{\underline{u}_1, \underline{u}_2, \underline{u}_3\}$ è una base di \mathbf{R}^3 . Per verificare che i tre vettori sono linearmente indipendenti, bisogna verificare che il sistema di equazioni lineari

$$\begin{cases} c_1 + c_2 + c_3 = 0 \\ c_2 + c_3 = 0 \\ c_3 = 0 \end{cases}$$

non ammette soluzioni diverse dalla terna $(0, 0, 0)$ (e ciò è praticamente immediato). Per verificare che i tre vettori formano un sistema di generatori bisogna verificare che, per ogni terna $(a, b, c) \in \mathbf{R}^3$, il sistema di equazioni lineari (nelle incognite c_1, c_2, c_3)

$$\begin{cases} c_1 + c_2 + c_3 = a \\ c_2 + c_3 = b \\ c_3 = c \end{cases}$$

ammette soluzioni. Si verifica facilmente che una soluzione è la terna $(a - b, b - c, c)$.

Due commenti. 1. La risoluzione dei precedenti sistemi di equazioni lineari risulta probabilmente piuttosto agevole, in quanto i dati numerici assegnati sono semplici e il numero delle equazioni e delle variabili è molto basso. E' evidente che, se i dati sono più complicati, per risolvere tali sistemi occorre un po' di teoria.

2. Per verificare che i tre vettori $\underline{u}_1, \underline{u}_2, \underline{u}_3$ formano una base abbiamo eseguito due calcoli diversi (uno per verificare l'indipendenza lineare ed uno per verificare che i vettori sono un sistema di generatori). I prossimi risultati ci diranno che abbiamo fatto un lavoro parzialmente inutile: in \mathbf{R}^n , se n vettori sono linearmente indipendenti, sono automaticamente un sistema di generatori e viceversa.

Vogliamo introdurre il concetto di *dimensione* di uno spazio vettoriale e dimostrare l'affermazione fatta qui sopra, in conclusione del commento 2. Cominciamo con il seguente risultato, che è il cuore di tutto.

Proposizione 4. Sia $\{\underline{u}_1, \dots, \underline{u}_n\}$ un sistema di generatori di V . Siano $\underline{w}_1, \dots, \underline{w}_m$ m vettori di

V. Se $m > n$, tali vettori sono linearmente dipendenti.

Dim. Dei vettori assegnati $\underline{w}_1, \dots, \underline{w}_m$, consideriamo i primi n , cioè $\underline{w}_1, \dots, \underline{w}_n$. Se tali vettori fossero linearmente dipendenti, anche $\underline{w}_1, \dots, \underline{w}_m$ lo sarebbero (in base all'**Osserv. 1(iv)**) e dunque non ci sarebbe altro da dimostrare. Assumiamo invece che $\underline{w}_1, \dots, \underline{w}_n$ siano linearmente indipendenti. Basterà provare in tal caso che risulta:

$$(\star) \quad \langle \underline{w}_1, \dots, \underline{w}_n \rangle = V$$

[se infatti (\star) è verificato, $\underline{w}_{n+1} \in \langle \underline{w}_1, \dots, \underline{w}_n \rangle$ e quindi $\underline{w}_1, \dots, \underline{w}_n, \underline{w}_{n+1}$ sono linearmente dipendenti. Allora $\underline{w}_1, \dots, \underline{w}_m$ sono linearmente dipendenti].

Per provare (\star) faremo vedere che, partendo dal sistema di generatori $\{\underline{u}_1, \dots, \underline{u}_n\}$, è possibile sostituire uno dei generatori \underline{u}_i (ad esempio \underline{u}_1) con \underline{w}_1 ed ottenere un nuovo sistema di generatori $\{\underline{w}_1, \underline{u}_2, \dots, \underline{u}_n\}$. Partendo da tale sistema di generatori vedremo che è possibile sostituire \underline{w}_2 con uno dei vettori \underline{u}_i (ad esempio \underline{u}_2), ottenendo un sistema di generatori $\{\underline{w}_1, \underline{w}_2, \underline{u}_3, \dots, \underline{u}_n\}$. Passo passo, sostituiremo tutti i \underline{w}_i ai vecchi generatori \underline{u}_i , ed otterremo (\star) .

Poiché $V = \langle \underline{u}_1, \dots, \underline{u}_n \rangle$,

$$\underline{u}_1 = a_1 \underline{u}_1 + \dots + a_n \underline{u}_n, \text{ per opportuni } a_1, \dots, a_n \in K.$$

Poiché $\underline{u}_1 \neq \underline{0}$, qualche coefficiente a_i è non nullo. Se ad esempio $a_1 \neq 0$ si ha:

$$\underline{u}_1 = \frac{1}{a_1} \underline{u}_1 - \frac{a_2}{a_1} \underline{u}_2 - \dots - \frac{a_n}{a_1} \underline{u}_n.$$

Dunque $\underline{u}_1 \in \langle \underline{w}_1, \underline{u}_2, \dots, \underline{u}_n \rangle$. Ma allora $V = \langle \underline{u}_1, \underline{u}_2, \dots, \underline{u}_n \rangle \subseteq \langle \underline{w}_1, \underline{u}_2, \dots, \underline{u}_n \rangle$ e quindi

$$V = \langle \underline{w}_1, \underline{u}_2, \dots, \underline{u}_n \rangle.$$

Se $n = 1$, (\star) è provata. Assumiamo $n > 1$. Si ha allora

$$\underline{u}_2 = b_1 \underline{u}_1 + b_2 \underline{u}_2 + \dots + b_n \underline{u}_n.$$

I coefficienti b_2, \dots, b_n non sono tutti nulli [altrimenti $\underline{u}_2 = b_1 \underline{u}_1$ e dunque $\underline{u}_1, \underline{u}_2$ sarebbero linearmente dipendenti, contro l'ipotesi]. Se, se ad esempio $b_2 \neq 0$, si ottiene (operando come sopra) che $\underline{u}_2 \in \langle \underline{w}_1, \underline{u}_2, \underline{u}_3, \dots, \underline{u}_n \rangle$, da cui

$$V = \langle \underline{w}_1, \underline{u}_2, \underline{u}_3, \dots, \underline{u}_n \rangle.$$

Se $n = 2$, (\star) è provata. Altrimenti con la stessa procedura si ottiene $V = \langle \underline{w}_1, \underline{u}_2, \underline{u}_3, \underline{u}_4, \dots, \underline{u}_n \rangle$ e così via, sino a verificare (\star) .

Teorema 1. (*Teorema della dimensione*) Se un K -spazio vettoriale V ha una base formata da n vettori, ogni altra base di V è formata da n vettori.

Dim. Sia $\{\underline{u}_1, \dots, \underline{u}_n\}$ una base di V e sia $\{\underline{w}_1, \dots, \underline{w}_m\}$ un'altra base di V . I vettori $\underline{w}_1, \dots, \underline{w}_m$ sono linearmente indipendenti e quindi, in base alla **Prop. 4**, $m \leq n$. D'altra parte $\{\underline{w}_1, \dots, \underline{w}_m\}$ è un sistema di generatori di V e $\underline{u}_1, \dots, \underline{u}_n$ sono linearmente indipendenti. Sempre dalla **Prop. 4**, $n \leq m$. Pertanto $n = m$.

Definizione 7. Sia V un K -spazio vettoriale che ammette una base finita. Il numero di vettori di tale base [e quindi di ogni altra base di V] è detto *dimensione* di V e sarà denotato $\dim_K(V)$ oppure $\dim(V)$ o anche $\dim V$. Si noti che lo spazio vettoriale nullo $\{\underline{0}\}$ non ha basi; gli attribuiremo comunque dimensione 0.

Scriveremo talvolta $V = V_K^n$ per indicare che V è un K -spazio vettoriale di dimensione n .

È evidente che $\dim(K^n) = n$. Vogliamo ora calcolare la dimensione dello spazio vettoriale $\mathfrak{M}_{m,n}(K)$ delle matrici ad m righe ed n colonne, a valori in K (introdotto nel paragrafo **2.2**).

Considerate le mn matrici elementari E^{h^k} (cfr. **Osserv. 2.2**), abbiamo già notato che tali matrici formano un sistema di generatori di $\mathfrak{M}_{m,n}(K)$. Verifichiamo che tali matrici sono anche linearmente indipendenti (e dunque una base dello spazio vettoriale). Risulta:

$$\sum_{h=1}^m \sum_{k=1}^n a_{hk} E^{hk} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}.$$

Dunque, se $\sum_{h=1}^m \sum_{k=1}^n a_{hk} E^{hk} = \mathbf{0}$ (matrice nulla), allora $a_{hk} = 0$, $\forall h = 1, \dots, m$, $\forall k = 1, \dots, n$.

Concludiamo quindi che $\dim(\mathfrak{M}_{m,n}(K)) = mn$. La base formata dalle matrici elementari è detta *base canonica* di $\mathfrak{M}_{m,n}(K)$.

Si noti che $\mathfrak{M}_{m,n}(K)$ si identifica in modo naturale con K^{mn} . Basta scrivere le righe di ogni matrice una di seguito all'altra e si ottiene una mn -pla.

Veniamo ora ad un'altra conseguenza della **Prop. 4**, preannunciata nel precedente commento 2.

Proposizione 5. *Sia $\dim_K(V) = n$. Risulta:*

- (i) *n vettori linearmente indipendenti formano una base.*
- (ii) *Un sistema di generatori di V formato da n vettori è una base.*

Dim. (i) Siano $\underline{u}_1, \dots, \underline{u}_n \in V$ vettori linearmente indipendenti. Vogliamo dimostrare che ogni $\underline{v} \in V$ si esprime come loro combinazione lineare. Consideriamo gli $n+1$ vettori $\underline{v}, \underline{u}_1, \dots, \underline{u}_n$. In base alla **Prop. 4**, tali vettori sono linearmente dipendenti. Sia

$$a\underline{v} + b_1\underline{u}_1 + \dots + b_n\underline{u}_n = \underline{0}$$

una loro relazione di dipendenza lineare non banale. Se per assurdo fosse $a = 0$, risulterebbe $b_1\underline{u}_1 + \dots + b_n\underline{u}_n = \underline{0}$, con b_1, \dots, b_n non tutti nulli. Dunque $\underline{u}_1, \dots, \underline{u}_n$ sarebbero linearmente dipendenti, contro l'ipotesi. Pertanto $a \neq 0$. Ma allora

$$\underline{v} = -\frac{b_1}{a}\underline{u}_1 - \frac{b_2}{a}\underline{u}_2 - \dots - \frac{b_n}{a}\underline{u}_n$$

e dunque \underline{v} è combinazione lineare di $\underline{u}_1, \dots, \underline{u}_n$, come richiesto.

(ii) Sia $\langle \underline{u}_1, \dots, \underline{u}_n \rangle = V$. Vogliamo dimostrare che tali vettori sono linearmente indipendenti, cioè che nessuno di essi è combinazione lineare dei rimanenti. Per assurdo, sia ad esempio \underline{u}_1 combinazione lineare di $\underline{u}_2, \dots, \underline{u}_n$. Ne segue che $\langle \underline{u}_2, \dots, \underline{u}_n \rangle = \langle \underline{u}_1, \underline{u}_2, \dots, \underline{u}_n \rangle = V$. Ma allora, poiché V ha un sistema di generatori formato da $n-1$ vettori, in base alla **Prop. 4** non può ammettere basi formate da n vettori. Da ciò segue un assurdo. Quindi $\underline{u}_1, \dots, \underline{u}_n$ sono linearmente indipendenti.

Osservazione 3. Sia V un K -spazio vettoriale di dimensione finita e sia W un suo sottospazio vettoriale. Risulta subito che $\dim(W) \leq \dim(V)$. Se poi $\dim(W) = \dim(V)$, allora $W = V$.

Dimostriamo la prima affermazione. Sia $\dim(V) = n$ e, per assurdo, $\dim(W) > n$. Esisterebbero in W $n+1$ vettori linearmente indipendenti. Tali vettori sarebbero linearmente indipendenti anche in V e ciò è assurdo in base alla **Prop. 4**. Dunque $\dim(W) \leq \dim(V)$.

Se poi $\dim(W) = \dim(V) = n$, in W esiste una base formata da n vettori. Tali vettori sono linearmente indipendenti anche in V . Dunque, dalla **Prop. 5**, sono anche una base di V . Pertanto $V = W$.

Ora ci poniamo due problemi in qualche modo opposti tra loro:

- (1) Assegnati in uno spazio vettoriale V di dimensione finita un certo numero di vettori linearmente indipendenti, è possibile "completarli", aggiungendone altri in modo da ottenere una base di V ?
- (2) Assegnati in uno spazio vettoriale V di dimensione finita un sistema di generatori, è possibile "estrarre" una base di V ?

La risposta è positiva in entrambi i casi.

Teorema 2. (Teorema del completamento) *Sia $\dim_K(V) = n$ e siano $\underline{u}_1, \dots, \underline{u}_t \in V$ t vettori linearmente indipendenti, con $t < n$.*

Esistono in V $n-t$ vettori $\underline{u}_{t+1}, \dots, \underline{u}_n$ tali che $\{\underline{u}_1, \dots, \underline{u}_t, \underline{u}_{t+1}, \dots, \underline{u}_n\}$ è una base di V .

Dim. Poiché $t < n$, $\langle \underline{u}_1, \dots, \underline{u}_t \rangle \subset V$ (in base a **Prop. 4**). Si scelga arbitrariamente un vettore $\underline{u}_{t+1} \in V - \langle \underline{u}_1, \dots, \underline{u}_t \rangle$. Verifichiamo che i vettori $\underline{u}_1, \dots, \underline{u}_t, \underline{u}_{t+1}$ sono linearmente indipendenti. Sia

$$a_1 \underline{u}_1 + \dots + a_t \underline{u}_t + a_{t+1} \underline{u}_{t+1} = \underline{0}.$$

Se fosse $a_{t+1} \neq 0$, allora $\underline{u}_{t+1} \in \langle \underline{u}_1, \dots, \underline{u}_t \rangle$, contro l'ipotesi. Dunque $a_{t+1} = 0$. Ma allora, essendo $\underline{u}_1, \dots, \underline{u}_t$ linearmente indipendenti, anche $a_1 = \dots = a_t = 0$. Dunque $\underline{u}_1, \dots, \underline{u}_t, \underline{u}_{t+1}$ sono linearmente indipendenti.

Se ora $t+1 < n$ risulta $\langle \underline{u}_1, \dots, \underline{u}_t, \underline{u}_{t+1} \rangle \subset V$ e quindi si può scegliere arbitrariamente un vettore $\underline{u}_{t+2} \in V - \langle \underline{u}_1, \dots, \underline{u}_{t+1} \rangle$. Come fatto sopra si verifica che i vettori $\underline{u}_1, \dots, \underline{u}_t, \underline{u}_{t+1}, \underline{u}_{t+2}$ sono linearmente indipendenti. Procedendo in questo modo, dopo un numero finito di passi si perviene ad una base $\{\underline{u}_1, \dots, \underline{u}_t, \underline{u}_{t+1}, \dots, \underline{u}_n\}$ di V .

Teorema 3. (*Teorema dell'estrazione di una base*) Sia $\dim_K(V) = n$ e sia $\{\underline{u}_1, \dots, \underline{u}_m\}$ un sistema di generatori di V [si noti che in ogni caso $m \geq n$, in base a **Prop. 4**].

Esistono n vettori distinti $\underline{u}_{i_1}, \dots, \underline{u}_{i_n} \in \{\underline{u}_1, \dots, \underline{u}_m\}$ formanti una base di V .

Dim. Se $n = 1$, il teorema è evidente [ogni vettore non nullo di $\{\underline{u}_1, \dots, \underline{u}_m\}$ è una base]. Sia dunque $n \geq 2$. Possiamo ovviamente assumere $\underline{u}_1, \dots, \underline{u}_m$ non nulli. Poniamo $\underline{u}_{i_1} = \underline{u}_1$. Tra $\underline{u}_2, \dots, \underline{u}_m$ scartiamo via via tutti i vettori che sono combinazioni lineari dei precedenti. I vettori superstiti sono un sistema di generatori $\{\underline{u}_{i_1}, \dots, \underline{u}_{i_s}\}$ di V , tali che

$$\langle \underline{u}_{i_1} \rangle \subset \langle \underline{u}_{i_1}, \underline{u}_{i_2} \rangle \subset \dots \subset \langle \underline{u}_{i_1}, \underline{u}_{i_2}, \dots, \underline{u}_{i_s} \rangle.$$

Proveremo che tali vettori formano una base di V [e quindi $s = n$], dimostrando che sono linearmente indipendenti.

Per assurdo siano linearmente dipendenti e sia $\sum_{k=1}^s a_k \underline{u}_{i_k} = \underline{0}$ una loro relazione di dipendenza lineare non banale. È evidente che almeno due coefficienti a_k sono non nulli. Supponiamo che a_r sia l'ultimo di essi. Il vettore \underline{u}_{i_r} può essere allora scritto come combinazione lineare dei vettori precedenti, cioè $\underline{u}_{i_r} \in \langle \underline{u}_{i_1}, \underline{u}_{i_2}, \dots, \underline{u}_{i_{r-1}} \rangle$. Ciò contrasta con la catena di inclusioni stretta riportata sopra. Si conclude che $\underline{u}_{i_1}, \underline{u}_{i_2}, \dots, \underline{u}_{i_s}$ sono linearmente indipendenti.

Osservazione 4. Il precedente teorema si applica ad ogni sottospazio vettoriale W di V . Se W ha un sistema di generatori $\{\underline{u}_1, \dots, \underline{u}_t\}$, una sua base è ottenibile scegliendo tra essi il massimo numero di vettori linearmente indipendenti. Ovviamente $\dim(W) \leq t$.

Presenteremo ora una formula che collega tra loro le dimensioni di certi sottospazi vettoriali di uno spazio vettoriale. Tale formula è nota come *formula di Grassmann*. Premettiamo un'osservazione.

Osservazione 5. Siano W_1, W_2 due K -sottospazi vettoriali di V . Risulta:

- (i) $W_1 \cap W_2$ è un sottospazio vettoriale di V , detto ovviamente *sottospazio intersezione di W_1, W_2* .
- (ii) $W_1 + W_2 := \{w_1 + w_2, \forall w_1 \in W_1, \forall w_2 \in W_2\}$ è un sottospazio vettoriale di V , detto *sottospazio somma di W_1, W_2* .

Entrambe le affermazioni si dimostrano utilizzando la **Prop. 5** del precedente paragrafo. Si osservi che $W_1 + W_2$ è il più piccolo sottospazio vettoriale contenente sia W_1 che W_2 . Invece $W_1 \cap W_2$ è il più grande sottospazio vettoriale contenuto sia in W_1 che in W_2 .

Si noti infine che l'unione insiemistica di due sottospazi vettoriali non è (se non in casi particolari) un sottospazio vettoriale. Infatti non è *chiusa* rispetto alla somma di vettori (cioè la somma di due vettori in $W_1 \cup W_2$ può non essere in $W_1 \cup W_2$).

Teorema 4. (*Formula di Grassmann*) Siano W_1, W_2 due sottospazi vettoriali di un K -spazio vettoriale V (di dimensione finita). Risulta:

$$\dim_K(W_1) + \dim_K(W_2) = \dim_K(W_1 + W_2) + \dim_K(W_1 \cap W_2).$$

Dim. Sia $n_1 = \dim(W_1)$, $n_2 = \dim(W_2)$ e $i = \dim(W_1 \cap W_2)$ [in base all'**Osserv. 3**, $i \leq n_1$, $i \leq n_2$]. Si fissi in $W_1 \cap W_2$ una base $\{\underline{z}_1, \dots, \underline{z}_i\}$.

Poiché $W_1 \cap W_2$ è un sottospazio di W_1 , in base al **Teor. 2** possiamo completare tali vettori (linearmente indipendenti) sino ad ottenere una base di W_1 . Sia $\{\underline{z}_1, \dots, \underline{z}_i, \underline{u}_1, \dots, \underline{u}_{n_1-i}\}$ tale base. Analogamente, completiamo la base $\{\underline{z}_1, \dots, \underline{z}_i\}$, sino ad ottenere una base $\{\underline{z}_1, \dots, \underline{z}_i, \underline{v}_1, \dots, \underline{v}_{n_2-i}\}$ di W_2 . Raduniamo ora tutti i vettori sopra considerati

$$\underline{z}_1, \dots, \underline{z}_i, \underline{u}_1, \dots, \underline{u}_{n_1-i}, \underline{v}_1, \dots, \underline{v}_{n_2-i}.$$

[Sono $i + (n_1 - i) + (n_2 - i) = n_1 + n_2 - i$]. Se dimostriamo che tali vettori formano una base di $W_1 + W_2$, allora $\dim(W_1 + W_2) = \dim(W_1) + \dim(W_2) - \dim(W_1 \cap W_2)$, da cui la formula cercata.

È evidente che quei vettori formano un sistema di generatori di $W_1 + W_2$. Infatti per ogni vettore $\underline{w}_1 + \underline{w}_2$, il vettore \underline{w}_1 è combinazione lineare di $\underline{z}_1, \dots, \underline{z}_i, \underline{u}_1, \dots, \underline{u}_{n_1-i}$, mentre il vettore \underline{w}_2 è combinazione lineare di $\underline{z}_1, \dots, \underline{z}_i, \underline{v}_1, \dots, \underline{v}_{n_2-i}$.

Sia ora

$$a_1 \underline{z}_1 + \dots + a_i \underline{z}_i + b_1 \underline{u}_1 + \dots + b_{n_1-i} \underline{u}_{n_1-i} + c_1 \underline{v}_1 + \dots + c_{n_2-i} \underline{v}_{n_2-i} = \underline{0}.$$

Per abbreviare la scrittura, poniamo

$$\underline{z} := a_1 \underline{z}_1 + \dots + a_i \underline{z}_i, \quad \underline{u} := b_1 \underline{u}_1 + \dots + b_{n_1-i} \underline{u}_{n_1-i}, \quad \underline{v} := c_1 \underline{v}_1 + \dots + c_{n_2-i} \underline{v}_{n_2-i}.$$

Dunque $\underline{z} + \underline{u} + \underline{v} = \underline{0}$. Risulta:

$$\underline{v} = -(\underline{z} + \underline{u}) \in W_1 \cap W_2$$

[infatti $\underline{v} \in W_2$ e $-(\underline{z} + \underline{u}) \in W_1$]. Ne segue che tale vettore può essere espresso in base $\{\underline{z}_1, \dots, \underline{z}_i\}$, per cui si ha:

$$\underline{v} = -(\underline{z} + \underline{u}) = d_1 \underline{z}_1 + \dots + d_i \underline{z}_i.$$

Ne segue che

$$\underline{z} + \underline{u} + d_1 \underline{z}_1 + \dots + d_i \underline{z}_i = \underline{0},$$

cioè

$$a_1 \underline{z}_1 + \dots + a_i \underline{z}_i + b_1 \underline{u}_1 + \dots + b_{n_1-i} \underline{u}_{n_1-i} + d_1 \underline{z}_1 + \dots + d_i \underline{z}_i = \underline{0},$$

ovvero

$$(a_1 + d_1) \underline{z}_1 + \dots + (a_i + d_i) \underline{z}_i + b_1 \underline{u}_1 + \dots + b_{n_1-i} \underline{u}_{n_1-i} = \underline{0}.$$

Poiché $\underline{z}_1, \dots, \underline{z}_i, \underline{u}_1, \dots, \underline{u}_{n_1-i}$ sono linearmente indipendenti, i coefficienti di tale relazione sono tutti nulli. In particolare $b_1 = \dots = b_{n_1-i} = 0$. Ma allora

$$a_1 \underline{z}_1 + \dots + a_i \underline{z}_i + c_1 \underline{v}_1 + \dots + c_{n_2-i} \underline{v}_{n_2-i} = \underline{0}.$$

Ma anche $\underline{z}_1, \dots, \underline{z}_i, \underline{v}_1, \dots, \underline{v}_{n_2-i}$ sono linearmente indipendenti e quindi $a_1 = \dots = a_i = c_1 = \dots = c_{n_2-i} = 0$. Pertanto i vettori considerati sono linearmente indipendenti.

Nel seguente esempio utilizzeremo la formula di Grassmann per determinare dimensione e base di sottospazi somma e intersezione di due sottospazi assegnati.

Siano $\underline{u}_1, \underline{u}_2, \underline{u}_3$ tre vettori linearmente indipendenti di uno spazio vettoriale V . Assegnati i due sottospazi vettoriali $W_1 = \langle \underline{u}_1, \underline{u}_3 \rangle$ e $W_2 = \langle \underline{u}_1 - \underline{u}_2, \underline{u}_1 - \underline{u}_3 \rangle$, vogliamo determinare la dimensione ed una base di $W_1 + W_2$ e di $W_1 \cap W_2$.

Osserviamo subito che $\dim(W_1) = 2$. Anche $\dim(W_2) = 2$ [infatti, se $a(\underline{u}_1 - \underline{u}_2) + b(\underline{u}_1 - \underline{u}_3) = \underline{0}$, allora $(a+b)\underline{u}_1 - a\underline{u}_2 - b\underline{u}_3 = \underline{0}$ e quindi $a+b = -a = -b = 0$, da cui $a = b = 0$; i due generatori di W_2 sono linearmente indipendenti].

Il sottospazio $W_1 + W_2$ è generato da $\{\underline{u}_1, \underline{u}_3, \underline{u}_1 - \underline{u}_2, \underline{u}_1 - \underline{u}_3\}$. Poiché $\underline{u}_1 - \underline{u}_3 \in \langle \underline{u}_1, \underline{u}_3 \rangle$, allora $W_1 + W_2 = \langle \underline{u}_1, \underline{u}_3, \underline{u}_1 - \underline{u}_2 \rangle$. Questi tre generatori sono linearmente indipendenti [infatti, sia $a\underline{u}_1 + b\underline{u}_3 + c(\underline{u}_1 - \underline{u}_2) = \underline{0}$; segue che $a+c = -c = b = 0$ e dunque $a = b = c = 0$]. Pertanto

$\dim(W_1 + W_2) = 3$ ed i tre generatori formano una base. Si noti che anche $\underline{u}_2 \in W_1 + W_2$ e dunque una base è anche $\{\underline{u}_1, \underline{u}_2, \underline{u}_3\}$.

Dalla formula di Grassmann, $\dim(W_1 \cap W_2) = 2+2-3 = 1$. Si osserva subito che $\underline{u}_1 - \underline{u}_3 \in W_1 \cap W_2$. Dunque tale vettore forma una base di $W_1 \cap W_2$.

Dalla formula di Grassmann segue che

$$\dim(W_1 + W_2) = \dim W_1 + \dim W_2 \iff \dim(W_1 \cap W_2) = 0 \iff W_1 \cap W_2 = \{\underline{0}\}.$$

In tal caso, una base di $W_1 + W_2$ è ottenuta unendo una base di W_1 con una di W_2 .

Diremo poi che due sottospazi W_1, W_2 di V sono *supplementari* se $W_1 + W_2 = V$ e $W_1 \cap W_2 = \{\underline{0}\}$. In tal caso V è detto *somma diretta* di W_1, W_2 e si scrive $W_1 \oplus W_2 = V$.

Concludiamo il paragrafo dimostrando che gli spazi vettoriali di una data dimensione (finita) sono tutti isomorfi tra loro. La teoria degli spazi vettoriali è quindi decisamente scarna: a meno di isomorfismi esiste un unico modello di spazio vettoriale n -dimensionale: K^n .

Teorema 5. Sia V un K -spazio vettoriale di dimensione n e sia $\{\underline{e}_1, \underline{e}_2, \dots, \underline{e}_n\}$ una sua base. L'applicazione $f : V \rightarrow K^n$ tale che, $\forall \underline{v} \in V$:

$$f(\underline{v}) = (c_1, c_2, \dots, c_n), \text{ se } \underline{v} = \sum_{i=1}^n c_i \underline{e}_i,$$

è un isomorfismo di spazi vettoriali. Ne segue che due spazi vettoriali n -dimensionali sono isomorfi.

Dim. L'applicazione f associa ad ogni vettore le sue coordinate in base $\{\underline{e}_1, \underline{e}_2, \dots, \underline{e}_n\}$. In base alla **Prop. 3**, f è biettiva. Verifichiamo che è un omomorfismo [e quindi un isomorfismo]. Infatti, presi comunque $\underline{v} = \sum_{i=1}^n c_i \underline{e}_i$, $\underline{w} = \sum_{i=1}^n d_i \underline{e}_i$ in V ed a in K , si ha:

$$f(\underline{v} + \underline{w}) = f\left(\sum_{i=1}^n (c_i + d_i) \underline{e}_i\right) = (c_1 + d_1, \dots, c_n + d_n) = (c_1, \dots, c_n) + (d_1, \dots, d_n) = f(\underline{v}) + f(\underline{w}),$$

$$f(a\underline{v}) = f\left(\sum_{i=1}^n (a c_i) \underline{e}_i\right) = (a c_1, \dots, a c_n) = a(c_1, \dots, c_n) = a f(\underline{v}).$$

Veniamo all'ultima affermazione. Siano V, W due K -spazi vettoriali n -dimensionali. Scelta una base in V ed una in W , sono definiti i due isomorfismi $f : V \rightarrow K^n$ e $g : W \rightarrow K^n$ che associano ad ogni vettore le rispettive coordinate.

Tenuto conto del fatto che la composizione di isomorfismi è un isomorfismo e che l'inverso di un isomorfismo è un isomorfismo, l'applicazione $g^{-1} \circ f : V \rightarrow W$ è un isomorfismo.

N.B. Un'osservazione sulle notazioni impiegate. Abbiamo indicato i vettori della base scelta in V con $\underline{e}_1, \underline{e}_2, \dots, \underline{e}_n$. In questo stesso modo avevamo indicato anche i vettori della base canonica di K^n .

Ma non c'è alcun rapporto tra le due cose e lo studente non deve fare confusione: non ha senso infatti parlare di "base canonica" in uno spazio vettoriale "astratto" V [dove non ci sono vettori più semplici o più naturali di altri, da scegliere per formare questa fantomatica "base canonica"]. Soltanto in alcuni spazi vettoriali "concreti" [come K^n o $\mathfrak{M}_{m,n}(K)$] troviamo una "base canonica" [cioè una base i cui vettori sono più adatti di altri a fungere da base].

ESERCIZI PROPOSTI

2.5.1. Assegnati in \mathbf{R}^3 i due vettori $\underline{x} = (1, 0, -1)$ ed $\underline{y} = (0, 1, 1)$, verificare che sono linearmente indipendenti e determinare un vettore $\underline{z} \in \mathbf{R}^3$ tale che $\{\underline{x}, \underline{y}, \underline{z}\}$ sia una base di \mathbf{R}^3 .

2.5.2. Assegnati in \mathbf{R}^3 i vettori

$$\underline{x} = (1, 0, -1), \underline{y} = (0, 1, 1), \underline{z} = (1, 1, 0), \underline{w} = (-1, 0, 2),$$

verificare se è possibile estrarre da essi una base di \mathbf{R}^3 .

2.5.3. Sono assegnate in $\mathfrak{M}_2(\mathbf{R})$ le seguenti quattro matrici

$$A_1 = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, A_2 = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}, A_3 = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}, A_4 = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}.$$

(i) Verificare che tali matrici sono linearmente dipendenti ed esprimere l'ultima in funzione delle prime tre.

(ii) Posto $W = \langle A_1, A_2, A_3 \rangle$, determinare $\dim(W)$ e indicare una base di W .

(iii) Verificare se la matrice elementare $E^{11} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ appartiene a W e, in caso affermativo, scriverne le coordinate rispetto alla base di W ottenuta in (ii).

2.5.4. Sia $\mathcal{L} = \{A \in \mathfrak{M}_3(\mathbf{R}) \mid a_{12} = a_{13} = a_{21} = a_{31} = 0\}$. Si tratta di un \mathbf{R} -sottospazio vettoriale di $\mathfrak{M}_3(\mathbf{R})$ (cfr. Eserc. 2.4.3). Determinare una base e la dimensione di \mathcal{L} .

2.5.5. Sia $\{\underline{e}_1, \underline{e}_2\}$ una base di un \mathbf{R} -spazio vettoriale V di dimensione 2. Siano

$$\underline{f}_1 = \underline{e}_1 - 2\underline{e}_2, \quad \underline{f}_2 = 2\underline{e}_1 + \underline{e}_2 \in V.$$

Verificare che $\{\underline{f}_1, \underline{f}_2\}$ è una base di V . Assegnato in V il vettore $\underline{v} = \underline{e}_1 - \underline{e}_2$, esprimere rispetto alla base $\{\underline{f}_1, \underline{f}_2\}$.

2.5.6. Assegnata una base $\{\underline{e}_1, \underline{e}_2, \underline{e}_3, \underline{e}_4\}$ di un \mathbf{R} -spazio vettoriale V avente dimensione 4, si considerino i due sottospazi vettoriali

$$V_1 = \langle \underline{e}_1 - \underline{e}_2, \underline{e}_1 - \underline{e}_3 \rangle, \quad V_2 = \langle \underline{e}_1 - \underline{e}_4, \underline{e}_3 - \underline{e}_4 \rangle.$$

Verificare se i due sottospazi vettoriali sono supplementari.

2.5.7. Una matrice $A \in \mathfrak{M}_n(K)$ è detta *antisimmetrica* se risulta $a_{ij} = -a_{ji}$, $\forall i, j = 1, \dots, n$ [cioè se $A = -A^T$]. Verificare che le matrici antisimmetriche di $\mathfrak{M}_3(\mathbf{R})$ formano un sottospazio vettoriale di dimensione 3 in $\mathfrak{M}_3(\mathbf{R})$ e indicarne una base.

2.5.8. Scelto $n \in \mathbf{N}$, si consideri in $K[X]$ il sottoinsieme

$$W = W_n = \{P \in K[X] \mid \deg(P) \leq n\}.$$

Verificare che W è un K -sottospazio vettoriale di $K[X]$. Indicarne la dimensione ed una base.

Capitolo 3

SISTEMI DI EQUAZIONI LINEARI

1. Generalità e algoritmo di Gauss

Nel capitolo precedente abbiamo visto come per risolvere problemi legati allo studio degli spazi vettoriali lo strumento tecnico fondamentale sia la risoluzione di sistemi di equazioni lineari.

In questo capitolo ci occuperemo di studiare come si risolve un sistema lineare. Nel presente paragrafo introdurremo alcune definizioni e descriveremo un algoritmo che permette di risolvere tali sistemi, noto come *algoritmo di Gauss*.

Definizione 1. Un sistema di m equazioni lineari in n indeterminate (o incognite) x_1, x_2, \dots, x_n , a valori in un campo K [abbreviato SL oppure $SL(m, n, K)$] è un insieme di m equazioni del tipo:

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

o (in forma abbreviata):

$$\left\{ \sum_{j=1}^n a_{ij}x_j = b_i \quad [i = 1, \dots, m], \right.$$

dove gli elementi $a_{ij}, b_i \in K$ sono detti rispettivamente *coefficienti* e *termini noti* delle equazioni del SL . Si chiama soluzione di tale SL ogni n -pla $\underline{z} = (z_1, z_2, \dots, z_n) \in K^n$ tale che:

$$\left\{ \sum_{j=1}^n a_{ij}z_j = b_i \quad [i = 1, \dots, m] \right.$$

[cioè che trasformi le m equazioni del sistema in m uguaglianze in K]. Un SL privo di soluzioni è detto *incompatibile*; altrimenti è detto *compatibile o risolubile*.

Un $SL(m, n, K)$ è detto *omogeneo* [abbreviato $SLO(m, n, K)$] se i suoi termini noti sono tutti nulli. Ovviamente un SLO è sempre compatibile: infatti ammette come soluzione la n -pla nulla $\underline{0} = (0, 0, \dots, 0)$, detta *soluzione banale*. Le altre sue (eventuali) soluzioni sono dette *autosoluzioni*.

Infine, assegnato un $SL(m, n, K)$, sostituendo i termini noti con 0, si ottiene un $SLO(m, n, K)$, detto *sistema lineare omogeneo associato al SL* dato.

Osservazione 1. Con il prodotto righe per colonne è possibile ottenere una scrittura "matriciale" dei sistemi di equazioni lineari. Assegnato infatti il $SL(m, n, K)$:

$$\left\{ \sum_{j=1}^n a_{ij}x_j = b_i \quad [i = 1, \dots, m], \right.$$

indichiamo con X la colonna $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ delle indeterminate, con $\mathbf{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \in \mathfrak{M}_{m,1}(K)$ la colonna dei termini noti e con $A = (a_{ij}) \in \mathfrak{M}_{m,n}(K)$ la matrice dei coefficienti del SL . Eseguendo il prodotto righe per colonne delle matrici A ed X , si ha:

$$AX = \begin{pmatrix} a_{11}x_1 + \dots + a_{1n}x_n \\ a_{21}x_1 + \dots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + \dots + a_{mn}x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} = \mathbf{b}.$$

Pertanto il SL assegnato può essere scritto nella forma:

$$AX = \mathbf{b}.$$

La matrice $(A \ b) \in \mathfrak{M}_{m,n+1}(K)$ è detta *matrice completa* (o *matrice orlata*) del SL $AX = \mathbf{b}$ e lo individua completamente. Esiste un'ovvia corrispondenza biunivoca tra $\mathfrak{M}_{m,n+1}(K)$ e l'insieme dei $SL(m, n, K)$. Quanto alle incognite, è chiaro che il loro nome è arbitrario e può essere modificato; quel che conta è il numero delle incognite ed il loro ordine.

È evidente che le eventuali soluzioni $\underline{z} = (z_1, \dots, z_n) \in K^n$ del $SL(m, n, K)$ $AX = \mathbf{b}$ corrispondono biunivocamente alle matrici colonne $\mathbf{z} = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} \in \mathfrak{M}_{n,1}(K)$ tali che $A\mathbf{z} = \mathbf{b}$. [Per tale motivo è opportuno identificare \underline{z} con \mathbf{z} , cioè identificare K^n con $\mathfrak{M}_{n,1}(K)$].

Si osservi infine che ogni soluzione $\underline{z} = (z_1, \dots, z_n)$ del SL $AX = \mathbf{b}$ esprime la colonna \mathbf{b} come combinazione lineare delle colonne di A . Infatti risulta:

$$\mathbf{b} = A\mathbf{z} = (A_{(1)} \dots A_{(n)}) \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} = \sum_{i=1}^n A_{(i)} z_i.$$

Due esempi. (1) È assegnato il $SL(1, 3, \mathbf{R})$ $\{x_1 = 1\}$. Tale sistema ha tre incognite (di cui x_1 è ragionevolmente la prima). In forma matriciale si scrive

$$(1 \ 0 \ 0) \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = (1).$$

Tale SL è compatibile e le sue soluzioni sono le terne $(1, b, c) \in \mathbf{R}^3$, $\forall b, c \in \mathbf{R}$.

(2) È assegnato un SL con matrice completa $\begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & 2 \end{pmatrix}$. Si tratta del $SL(2, 2, \mathbf{R})$

$$\begin{cases} x_1 = 1 \\ x_1 = 2, \end{cases} \quad [\text{ovvero, scritto in forma matriciale, } \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}].$$

Tale sistema è manifestamente incompatibile (in quanto implica $1 = 2$).

Il $SLO(2, 2, \mathbf{R})$ ad esso associato è $\begin{cases} x_1 = 0 \\ x_1 = 0. \end{cases}$ Le soluzioni di tale SLO formano il sottoinsieme $\Sigma_0 = \{(0, t), \forall t \in \mathbf{R}\}$ di \mathbf{R}^2 . Si tratta del sottospazio vettoriale $\langle(0, 1)\rangle$ di \mathbf{R}^2 . Che l'insieme delle soluzioni di tale SLO sia uno spazio vettoriale non è un fatto casuale, come ora vedremo.

Proposizione 1. Sia $AX = \mathbf{0}$ un $SLO(m, n, K)$. L'insieme Σ_0 delle sue soluzioni è un sottospazio vettoriale di K^n .

Dim. Basta verificare che, $\forall \underline{y}, \underline{z} \in \Sigma_0$ e $\forall a, b \in K$, risulta $a\underline{y} + b\underline{z} \in \Sigma_0$, ovvero [indicate con \mathbf{y} e \mathbf{z} le colonne corrispondenti a \underline{y} , $\underline{z}\]$ $A(a\mathbf{y} + b\mathbf{z}) = \mathbf{0}$. Infatti si ha (utilizzando le proprietà di Prop. 1 di Cap. 2.2):

$$A(a\mathbf{y} + b\mathbf{z}) = A(a\mathbf{y}) + A(b\mathbf{z}) = a(A\mathbf{y}) + b(A\mathbf{z}) = a\mathbf{0} + b\mathbf{0} = \mathbf{0}.$$

Sia $AX = \mathbf{b}$ un SL compatibile e non omogeneo. In tal caso l'insieme Σ delle sue soluzioni non può mai essere un sottospazio vettoriale di K^n [infatti $\mathbf{0} \notin \Sigma$]. Tuttavia, come ora vedremo, Σ è in corrispondenza biunivoca con il sottospazio vettoriale Σ_0 delle soluzioni del suo SLO associato.

Proposizione 2. Sia $AX = \mathbf{b}$ un $SL(m, n, K)$ compatibile e sia \underline{z}_0 una sua soluzione. Denotato con Σ l'insieme delle sue soluzioni e con Σ_0 il sottospazio vettoriale delle soluzioni del SLO associato $AX = \mathbf{0}$, risulta:

$$\Sigma = \underline{z}_0 + \Sigma_0 \quad [= \{\underline{z}_0 + \underline{y}, \forall \underline{y} \in \Sigma_0\}].$$

Ne segue che Σ e Σ_0 sono in corrispondenza biunivoca.

Dim. (\subseteq). Sia $\underline{z} \in \Sigma$ [cioè $A\underline{z} = \mathbf{b}$]. Poiché, per ipotesi, $A\underline{z}_0 = \mathbf{b}$, allora:

$$A(\underline{z} - \underline{z}_0) = A\underline{z} - A\underline{z}_0 = \mathbf{b} - \mathbf{b} = \mathbf{0}$$

e dunque $\underline{z} - \underline{z}_0 \in \Sigma_0$. Pertanto $\underline{z} = \underline{z}_0 + (\underline{z} - \underline{z}_0) \in \underline{z}_0 + \Sigma_0$.

(\supseteq). Verifichiamo che, $\forall \underline{y} \in \Sigma_0$, $\underline{z}_0 + \underline{y} \in \Sigma$. Infatti:

$$A(\underline{z}_0 + \underline{y}) = A\underline{z}_0 + A\underline{y} = \mathbf{b} + \mathbf{0} = \mathbf{b}.$$

L'ultima affermazione è ovvia: l'applicazione $\underline{y} \rightarrow \underline{z}_0 + \underline{y}$ stabilisce una biiezione da Σ_0 a Σ .

La precedente proposizione suggerisce come ottenere tutte le soluzioni di un SL compatibile non omogeneo. Una volta ottenutane una soluzione, per ottenere le altre basta risolvere il SLO associato e sommarne le soluzioni a quella del SL già ottenuta.

Poiché Σ non è uno spazio vettoriale, "non ha diritto" ad una dimensione; tuttavia possiamo in qualche modo attribuirgli la dimensione di Σ_0 . Precisamente, se $\dim(\Sigma_0) = t$, diremo che il SL $AX = \mathbf{b}$ (se compatibile) *ha ∞^t soluzioni*. Questa terminologia proviene dal fatto che, come vedremo, le soluzioni dipendono da t parametri in K . Poiché K è tradizionalmente il campo \mathbf{R} , che ha cardinalità infinita, i t parametri variano ciascuno in infiniti modi e le soluzioni sono quindi ∞^t . Si noti infine che, se $\Sigma_0 = \{\underline{0}\}$, la terminologia introdotta ci dice che il SL (se compatibile) ha ∞^0 soluzioni; poiché in questo caso il SL ha una sola soluzione, conveniamo di porre: $\infty^0 = 1$.

Vogliamo ora risolvere un SL , cioè determinarne le soluzioni. Cominciamo dai SL più semplici: i sistemi di equazioni lineari *a scala* (o *a gradini*). Premettiamo una definizione.

Definizione 2. Siano $AX = \mathbf{b}$ e $A'X = \mathbf{b}'$ due SL aventi lo stesso numero di incognite. Diciamo che tali SL sono equivalenti se hanno le stesse soluzioni (cioè se $\Sigma = \Sigma'$).

Definizione 3. Un $SL(m, n, K)$ $AX = \mathbf{b}$ è detto *a scala* (o *a gradini*) se verifica le seguenti tre condizioni:

$$m \leq n, \quad a_{ij} = 0 \text{ se } i > j, \quad a_{ii} \neq 0, \quad \forall i = 1, \dots, m.$$

La matrice di un SL a scala è dunque del tipo

$$A = \begin{pmatrix} a_{11} & \dots & \dots & \dots & \dots \\ 0 & a_{22} & \dots & \dots & \dots \\ \vdots & \ddots & \ddots & & \vdots \\ 0 & 0 & 0 & a_{mm} & \dots \end{pmatrix}, \quad \text{con } a_{11} \cdot a_{22} \cdots \cdot a_{mm} \neq 0.$$

Risolvere un SL a scala è piuttosto semplice e intuitivo. Si comincia dall'ultima equazione, cioè

$$a_{mm}x_m + \dots + a_{mn}x_n = b_m.$$

Se $n = m$, si ottiene $x_n = \frac{b_n}{a_{nn}}$ e si sostituisce tale valore nella penultima equazione, ottenendo così un'unica espressione per x_{n-1} . Sostituendo ciascuno dei valori via via trovati nell'equazione precedente, si perviene ad un'unica soluzione del sistema.

Se invece $n > m$, si attribuiscono alle incognite x_{m+1}, \dots, x_n valori parametrici arbitrari (in K), ad esempio t_1, \dots, t_{n-m} e si ottiene un valore per x_m . Si procede poi come nel caso precedente e si otterrà un'unica soluzione del SL , dipendente da $n - m$ parametri indipendenti, cioè ∞^{n-m} soluzioni. Con la convenzione fatta sopra (cioè $\infty^0 = 1$) abbiamo quindi ottenuto il seguente risultato.

Proposizione 3. Ogni $SL(m, n, K)$ a scala è compatibile ed ha ∞^{n-m} soluzioni.

Ad esempio vogliamo risolvere il $SL(3, 4, \mathbf{R})$ a scala

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

Poniamo $x_4 = t$. Risolvendo l'ultima equazione, $x_3 = t$. Sostituendo i valori di x_3 (e x_4) nella seconda equazione si ottiene $x_2 = 1 - \frac{t}{2}$. Sostituendo i valori di x_2 , x_3 e x_4 nella prima equazione, si ottiene $x_1 = \frac{t}{2}$. Pertanto il SL assegnato ha le ∞^1 soluzioni

$$\left(\frac{t}{2}, 1 - \frac{t}{2}, t, t\right), \quad \forall t \in \mathbf{R}.$$

Utilizzando la **Prop. 2**, possiamo ottenere subito le soluzioni del SLO associato. Infatti, una soluzione del SL dato è $\underline{z}_0 = (0, 1, 0, 0)$ (ottenuta ponendo $t = 0$). Pertanto

$$\Sigma_0 = \Sigma - \underline{z}_0 = \left\{ \left(\frac{t}{2}, -\frac{t}{2}, t, t\right), \quad \forall t \in \mathbf{R} \right\} = \langle \left(\frac{1}{2}, -\frac{1}{2}, 1, 1\right) \rangle = \langle (1, -1, 2, 2) \rangle.$$

Veniamo ora all'*algoritmo di Gauss* (o *di Gauss-Jordan*) per la risoluzione di un sistema di equazioni lineari. Tale procedimento consiste nel *trasformare (se possibile) un assegnato $SL(m, n, K)$ $AX = \mathbf{b}$ in un SL a gradini ad esso equivalente* [che verrà poi risolto come visto sopra].

Per trasformare il SL si fa ricorso a tre tipi di operazioni sulle equazioni del SL , dette *operazioni elementari (sulle equazioni)*, e cioè:

I operazione elementare: scambiare di posizione due equazioni del SL ;

II operazione elementare: sostituire un'equazione con un multiplo non nullo della stessa equazione;

III operazione elementare: sostituire un'equazione con la stessa equazione sommata ad un multiplo di un'altra.

Tali operazioni [che del resto abbiamo già eseguito nella risoluzione "empirica" dei SL incontrati nel capitolo precedente] evidentemente non cambiano le soluzioni del sistema, e dunque trasformano il SL in un altro ad esso equivalente.

Se al sistema assegnato $AX = \mathbf{b}$ si sostituisce la sua matrice completa $M = (A \quad \mathbf{b})$, le tre operazioni elementari (sulle equazioni) si trasformano nelle corrispondenti *operazioni elementari di riga*, che indicheremo schematicamente come segue:

$$\mathbf{I}[M^{(i)} \leftrightarrow M^{(j)}];$$

$$\mathbf{II}[M^{(i)} \rightarrow cM^{(i)}], \quad \text{con } c \neq 0;$$

$$\mathbf{III}[M^{(i)} \rightarrow M^{(i)} + cM^{(j)}], \quad \text{con } i \neq j \text{ e } c \in K.$$

Ad esempio, l'operazione **III** sopra considerata sostituisce alla riga $M^{(i)}$ la riga $M^{(i)} + cM^{(j)}$ [ovvero sostituisce alla i -esima equazione la somma della i -esima con la j -esima moltiplicata per c].

Per abbreviare le notazioni, converremo di denotare tale operazione con $\mathbf{III}[(i^a) \rightarrow (i^a) + c(j^a)]$ ed usare analoghe notazioni per le altre due operazioni.

Oltre alle tre operazioni elementari di riga, nell'algoritmo di Gauss può talvolta essere necessario eseguire uno scambio di colonne della matrice A . Ciò evidentemente corrisponde ad uno scambio delle variabili del SL . In tal caso sarà poi necessario, al termine dell'algoritmo, procedere allo scambio di variabili opposto.

Descriviamo ora l'algoritmo di Gauss, suddividendolo in "blocchi" di quattro passi, da ripetere un numero finito di volte. Consideriamo un $SL(m, n, K)$ $AX = \mathbf{b}$, con matrice completa $M = (A \quad \mathbf{b})$.

1° passo. Si elimina ogni eventuale riga nulla di M [corrispondente all'equazione banale $0 = 0$].

2° passo. Si fa in modo che risulti $M_{(1)} \neq \mathbf{0}$ (cioè $A_{(1)} \neq \mathbf{0}$). Ciò può essere ottenuto scambiando ad esempio la colonna nulla $A_{(1)}$ con una successiva colonna $A_{(i)}$ non nulla [ed è meglio scegliere l'ultima colonna non nulla di A , per evitare di dover ripetere tale operazione].

3° passo. Si fa in modo che risulti $a_{11} = 1$. A tale scopo, se $a_{11} = 0$ e ad esempio $a_{i1} \neq 0$, si esegue l'operazione $\mathbf{I}[M^{(1)} \leftrightarrow M^{(i)}]$ e si ottiene quindi una nuova matrice con $a_{11} \neq 0$; successivamente, se $a_{11} \neq 1$, si esegue l'operazione $\mathbf{II}[M^{(1)} \rightarrow \frac{1}{a_{11}} M^{(1)}]$ e si ottiene $a_{11} = 1$.

4° passo. Si fa in modo che risulti: $a_{21} = a_{31} = \dots = a_{m1} = 0$. A tale scopo, basta eseguire (per $i = 2, \dots, m$) le operazioni $\mathbf{III}[M^{(i)} \rightarrow M^{(i)} - a_{i1} M^{(1)}]$.

A questo punto la matrice completa M del SL assegnato si è trasformata in una matrice del tipo

[ma si noti che in tale matrice gli elementi a_{12}, \dots della seconda e delle successive colonne non sono ovviamente gli stessi della matrice A ; abbiamo mantenuto le stesse lettere soltanto per semplificare le notazioni]:

$$\begin{pmatrix} 1 & a_{12} & \dots & \dots \\ 0 & a_{22} & \dots & \dots \\ 0 & \dots & \dots & \dots \\ 0 & a_{m'2} & \dots & \dots \end{pmatrix},$$

con $m' \leq m$ (e $m' < m$ a seguito di cancellazione di righe nulle).

Si ripete ora il blocco dei quattro passi sopra descritto, a partire dall'elemento che si trova sulla seconda riga e seconda colonna di tale matrice. Si eliminano quindi eventuali righe nulle e si fa in modo (operando relativamente alla seconda colonna) che la matrice completa del SL diventi del tipo:

$$\begin{pmatrix} 1 & a_{12} & \dots & \dots & \dots \\ 0 & 1 & \dots & \dots & \dots \\ 0 & 0 & a_{33} & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & a_{m''3} & \dots & \dots \end{pmatrix}$$

(con $m'' \leq m'$). Si ripete quindi il blocco dei quattro passi a partire dalla terza riga e terza colonna, dalla quarta ... e così via.

Se nel corso del procedimento si ottiene una riga della forma $(0 \ 0 \dots 0 \ b)$, con $b \neq 0$, il SL è incompatibile [ed il procedimento di Gauss ovviamente si interrompe]. In caso contrario il SL si riduce ad un SL a scala, come richiesto. Si osservi infine che, se nel procedimento sono stati necessari scambi di variabili, sarà necessario ripristinare le variabili iniziali, procedendo agli scambi opposti.

Illustreremo l'algoritmo con un paio di esempi.

Esempio 1. Risolviamo il seguente $SL(4, 4, \mathbf{R})$:

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

La matrice completa del SL è

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

Per completare il primo blocco dell'algoritmo è sufficiente procedere con **III**[(3^a) \rightarrow (3^a) $-$ (1^a)]. Si ottiene:

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

Iniziamo il secondo blocco dell'algoritmo. Poiché la seconda colonna è nulla [a partire dall'elemento di posto (2, 2)], si procede (ad esempio) allo scambio di variabili $x_2 \leftrightarrow x_4$. Si pone quindi:

$$y_1 = x_1, \quad y_2 = x_4, \quad y_3 = x_3, \quad y_4 = x_2$$

e si ottiene un SL avente matrice completa:

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

Con **II**[(2^a) \rightarrow $\frac{1}{4}(2^a)$] si ottiene:

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

Poi, con $\text{III}[(3^a) \rightarrow (3^a) + 2(2^a)]$ e con $\text{III}[(4^a) \rightarrow (4^a) - (2^a)]$ si ottiene:

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

Iniziamo ora il terzo blocco dell'algoritmo. Con $\text{II}[(3^a) \rightarrow -\frac{2}{7}(3^a)]$ si ottiene:

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

Infine, con $\text{III}[(4^a) \rightarrow (4^a) - \frac{5}{4}(3^a)]$, si ottiene:

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

Veniamo ora al quarto ed ultimo blocco dell'algoritmo: basta eliminare la quarta riga (che è nulla) ed otteniamo una matrice a gradini. Si è così ottenuto il $SL(3, 4, \mathbf{R})$:

$$\begin{cases} y_1 + 4y_2 + 3y_3 + 2y_4 = 1 \\ y_2 - \frac{1}{4}y_3 = 0 \\ y_3 = 0. \end{cases}$$

Posto $y_4 = t$, si ha: $y_3 = 0$, $y_2 = 0$, $y_1 = 1 - 2t$. Ripristiniamo le variabili di partenza:

$$x_1 = y_1 = 1 - 2t, \quad x_2 = y_4 = t, \quad x_3 = y_3 = 0, \quad x_4 = y_2 = 0.$$

Pertanto l'insieme Σ delle soluzioni del SL assegnato è:

$$\Sigma = \{(1 - 2t, t, 0, 0), \forall t \in \mathbf{R}\}.$$

Esempio 2. Risolviamo il seguente $SL(3, 4, \mathbf{R})$ ed il corrispondente SLO associato:

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

La matrice completa del SL è

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

Per completare il primo blocco dell'algoritmo è sufficiente eseguire $\text{III}[(2^a) \rightarrow (2^a) + (1^a)]$. Si ottiene

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

Eseguiamo il secondo blocco. Con $\text{II}[(2^a) \rightarrow \frac{1}{3}(2^a)]$ otteniamo

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

e con $\text{III}[(3^a) \rightarrow (3^a) - 3(2^a)]$ otteniamo

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

Esaminando ora l'ultima riga concludiamo che il SL è incompatibile.

Consideriamo ora il $SLO(3, 4, \mathbf{R})$ asssociato. La sua matrice è

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

[si osservi che è inutile considerare la matrice completa del *SLO*, in quanto la colonna dei termini noti è nulla e non viene alterata nel corso dell'algoritmo].

Ripetiamo i passi dell'algoritmo appena eseguito. Con $\text{III}[(2^a) \rightarrow (2^a) + (1^a)]$ otteniamo

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

Con $\text{II}[(2^a) \rightarrow \frac{1}{3}(2^a)]$ otteniamo

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

e con $\text{III}[(3^a) \rightarrow (3^a) - 3(2^a)]$ otteniamo

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

Eliminiamo la terza riga (nulla) ed otteniamo il *SLO*(2, 4, \mathbf{R}) a scala

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

Poniamo $x_3 = t$, $x_4 = s$ ed otteniamo

$$x_2 = -\frac{1}{3}(t+s), \quad x_1 = \frac{2}{3}t - \frac{7}{3}s.$$

Ne segue che l'insieme delle soluzioni di tale *SLO* è

$$\Sigma_0 = \left\{ \left(\frac{2}{3}t - \frac{7}{3}s, -\frac{1}{3}(t+s), t, s \right), \quad \forall t, s \in \mathbf{R} \right\}$$

ovvero

$$\Sigma_0 = \langle \left(\frac{2}{3}, -\frac{1}{3}, 1, 0 \right), \left(-\frac{7}{3}, -\frac{1}{3}, 0, 1 \right) \rangle.$$

Risolvere un *SL* con l'algoritmo di Gauss può essere molto vantaggioso se si è dotati della velocità di un computer o se si dispone di molta pazienza (e molta carta!). Quando poi il *SL* non ha coefficienti tutti numerici, ma alcuni di essi dipendono da parametri, l'algoritmo di Gauss si complica notevolmente, in quanto si deve tener conto del possibile annullamento delle espressioni dipendenti da parametri, che vengono via via a determinarsi tramite le operazioni fondamentali di riga.

Per questo è utile saper risolvere un *SL* anche con mezzi più teorici, cioè utilizzando i classici teoremi di Cramer e di Rouché-Capelli. Quest'ultimo teorema poi è centrale nell'Algebra Lineare, in quanto ci permette di rappresentare, usando sistemi di equazioni lineari, i sottospazi di uno spazio vettoriale.

Gli strumenti tecnici che ci serviranno per descrivere tali risultati e pervenire quindi ad un altro metodo per la risoluzione di un *SL* sono due: il *determinante di una matrice quadrata* ed il *rango di una matrice*. Di essi ci occuperemo nei due paragrafi successivi.

ESERCIZI PROPOSTI

3.1.1. Risolvere con l'algoritmo di Gauss il seguente *SL*(3, 5, \mathbf{R})

$$\begin{cases} x_3 + 2x_5 = 2 \\ x_4 - x_5 = 3 \\ 2x_1 + x_3 = 1. \end{cases}$$

Dedurne una base del sottospazio vettoriale Σ_0 delle soluzioni del *SLO* associato.

3.1.2. Al variare del parametro $a \in \mathbf{R}$, risolvere il seguente *SLO*(2, 2, \mathbf{R})

$$\begin{cases} 2x + (a+2)y = 0 \\ (a+1)x + (a^2+2)y = 0. \end{cases}$$

3.1.3. Al variare del parametro $a \in \mathbf{R}$, risolvere il seguente *SLO*(3, 2, \mathbf{R})

$$\begin{cases} x + ay = 0 \\ 2x + 2y = 0 \\ ax = 0. \end{cases}$$

3.1.4. Al variare dei parametri non nulli $a, b \in \mathbf{R}$, risolvere il seguente $SLO(3, 3, \mathbf{R})$

$$\begin{cases} a y + b z = 0 \\ -a x + z = 0 \\ -b x - y = 0. \end{cases}$$

3.1.5 Risolvere con l'algoritmo di Gauss il seguente $SL(2, 2, \mathbf{Z}_5)$

$$\begin{cases} \bar{3}x + y = \bar{1} \\ x - y = \bar{0}. \end{cases}$$

2. Determinante di una matrice quadrata

Se $A \in \mathfrak{M}_n(K)$, il *determinante di A* è un elemento di K , denotato $\det(A)$ (o $\det A$ o anche $|A|$), che è somma di $n!$ addendi, ciascuno dei quali è prodotto di n fattori. Questi n fattori sono elementi di A e vanno scelti uno in ciascuna riga ed uno in ciascuna colonna di A . Ad ogni addendo va poi attribuito un segno.

Prima di completare e formalizzare la definizione di determinante, scriviamo esplicitamente il determinante di matrici quadrate di ordini 1, 2 e 3. Poniamo:

$$\det(a_{11}) = a_{11},$$

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}a_{22} - a_{12}a_{21},$$

$$\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32}.$$

Come si vede il determinante di una matrice di ordine 3 è la somma di $3! = 6$ addendi e ciascuno di essi è prodotto di 3 fattori. Tali fattori sono scelti su ciascuna delle tre righe e su ciascuna delle tre colonne della matrice. Gli indici di colonna di ogni addendo definiscono una permutazione di S_3 . Ad esempio, l'ultimo addendo scritto sopra corrisponde alla permutazione $\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} = (2\ 3)$. Si può osservare che il segno positivo è stato attribuito a quegli addendi che corrispondono a permutazioni di classe pari, mentre il segno negativo agli addendi corrispondenti a permutazioni di classe dispari.

Stesse considerazioni valgono per il determinante di una matrice di ordine 2 e di ordine 1. Siamo ormai pronti alla definizione generale di determinante.

Definizione 1. Sia $A \in \mathfrak{M}_n(K)$. Si chiama *determinante di A* l'elemento di K , denotato $\det(A)$ (o $\det A$ o $|A|$), così definito:

$$\det(A) = \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{1\sigma_1} a_{2\sigma_2} \dots a_{n\sigma_n},$$

dove S_n è il gruppo delle permutazioni su n elementi ed $\varepsilon(\sigma)$ è il segno della permutazione $\sigma \in S_n$, che è così definito:

$$\varepsilon(\sigma) = +1, \text{ se } \sigma \text{ è di classe pari}; \quad \varepsilon(\sigma) = -1, \text{ se } \sigma \text{ è di classe dispari}.$$

Per matrici quadrate di ordine ≥ 4 il calcolo del determinante diventa sempre più impegnativo (per l'alto numero degli addendi). Ma, come vedremo, esiste un risultato che riduce il calcolo di un tale determinante al calcolo di più determinanti di matrici di ordine inferiore: si tratta del *teorema di Laplace*.

Per il momento elenchiamo (senza dimostrazione) alcuni risultati che discendono, più o meno immediatamente, dalla definizione di determinante.

Proposizione 1. Sia $A = (a_{ij}) \in \mathfrak{M}_n(K)$. Risulta:

- (i) Se A è una matrice diagonale, $\det(A) = \prod_{i=1}^n a_{ii}$. In particolare $\det(I_n) = 1$.
- (ii) Se A ha una riga o una colonna nulla, risulta: $\det(A) = 0$.
- (iii) $\det(A) = \det({}^t A)$.
- (iv) Sia $A^{(i)} = bU + cV$, con $U, V \in \mathfrak{M}_{1,n}(K)$ e $b, c \in K$. Siano $B, C \in \mathfrak{M}_n(K)$ le matrici ottenute da A sostituendo la riga $A^{(i)}$ rispettivamente con U e con V . Risulta:

$$\det(A) = b \det(B) + c \det(C).$$

[Analogo risultato vale per le colonne di A].

(v) Sia $B \in \mathfrak{M}_n(K)$ la matrice ottenuta da A scambiando tra loro due righe oppure due colonne. Risulta: $\det(B) = -\det(A)$.

(vi) Se A ha due righe o due colonne proporzionali (o, in particolare, uguali), risulta: $\det(A) = 0$.

(vii) (Teorema di Binet). Risulta:

$$\det(AB) = \det(A)\det(B), \quad \forall A, B \in \mathfrak{M}_n(K).$$

(viii) Per ogni $A \in \mathbf{GL}_n(K)$, risulta: $\det(A^{-1}) = \frac{1}{\det(A)}$. Ne segue in particolare che $\det(A) \neq 0$.

Dei precedenti risultati, solo la dimostrazione di (vii) (Teorema di Binet) presenta qualche difficoltà. La (viii) ne è invece un'immmediata conseguenza. Tutte le altre affermazioni discendono facilmente dalla definizione di determinante e dalle proprietà delle permutazioni.

Combinando le proprietà (iv) e (vi), si ottiene che se una riga (o colonna) di A è combinazione lineare di altre due righe (o colonne) di A , $\det(A) = 0$. Dalla proprietà (iv) segue poi il seguente risultato:

$$\det(cA) = c^n \det(A), \quad \forall c \in K, \quad \forall A \in \mathfrak{M}_n(K).$$

Infatti:

$$\det(cA) = \begin{vmatrix} cA^{(1)} \\ cA^{(2)} \\ \vdots \\ cA^{(n)} \end{vmatrix} = c \begin{vmatrix} A^{(1)} \\ cA^{(2)} \\ \vdots \\ cA^{(n)} \end{vmatrix} = c^2 \begin{vmatrix} A^{(1)} \\ A^{(2)} \\ \vdots \\ cA^{(n)} \end{vmatrix} = \dots = c^n \begin{vmatrix} A^{(1)} \\ A^{(2)} \\ \vdots \\ A^{(n)} \end{vmatrix} = c^n \det(A).$$

Sempre la proprietà (iv) può essere utile per semplificare il calcolo del determinante. Ad esempio consideriamo la matrice

$$A = \begin{pmatrix} 1 & 40 & 1 \\ 20 & 200 & 20 \\ 0 & 60 & 2 \end{pmatrix} \in \mathfrak{M}_3(\mathbf{R}).$$

Per calcolare $\det(A)$ osserviamo che $A^{(2)} = 20 \begin{pmatrix} 1 & 10 & 1 \end{pmatrix}$. Dunque

$$\det(A) = 20 \det(B), \quad \text{con } B = \begin{pmatrix} 1 & 40 & 1 \\ 1 & 10 & 1 \\ 0 & 60 & 2 \end{pmatrix}.$$

Poiché poi $B_{(2)} = 10 \begin{pmatrix} 4 \\ 1 \\ 6 \end{pmatrix}$, allora $\det(B) = 10 \begin{vmatrix} 1 & 4 & 1 \\ 1 & 1 & 1 \\ 0 & 6 & 2 \end{vmatrix} = 10 \cdot (2 + 6 - 6 - 8) = -60$ e dunque $\det(A) = 20 \cdot (-60) = -1200$.

Per enunciare il teorema di Laplace abbiamo bisogno della definizione di *complemento algebrico* di un elemento di una matrice quadrata. Se ad una matrice quadrata A di ordine $n \geq 2$ togliamo gli elementi della riga i -sima e della colonna j -sima, otteniamo una matrice quadrata di ordine $n - 1$. Tale matrice ha il suo determinante e a tale determinante può essere attribuito un segno: quello di $(-1)^{i+j}$. Otteniamo un elemento di K , che viene detto *complemento algebrico dell'elemento a_{ij}* e che viene denotato α_{ij} .

Teorema 1. (Teorema di Laplace) Sia $A \in \mathfrak{M}_n(K)$, con $n \geq 2$. Risulta, per ogni $i, j = 1, \dots, n$:

$$\det(A) = \sum_{t=1}^n a_{it} \alpha_{it} \quad \text{e} \quad \det(A) = \sum_{t=1}^n a_{tj} \alpha_{tj}.$$

La prima sommatoria è detta *sviluppo di Laplace di $\det(A)$ rispetto alla riga i -sima*. È la somma degli elementi della riga $A^{(i)}$, ciascuno moltiplicato per il rispettivo complemento algebrico. Analogamente, la seconda sommatoria è lo *sviluppo di Laplace di $\det(A)$ rispetto alla colonna j -sima*.

Segue da questo risultato che, per calcolare il determinante di una matrice quadrata di ordine n , basta calcolare n determinanti di matrici quadrate di ordine $n - 1$. Naturalmente, nei casi concreti, sta alla furbizia dello studente eseguire lo sviluppo di Laplace rispetto alla riga o alla colonna che

presenti il massimo numero di zeri (perché è inutile calcolare i complementi algebrici degli elementi nulli, visto che poi vanno moltiplicati per 0). Ad esempio consideriamo la seguente matrice triangolare superiore:

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 2 & 3 & 4 \\ 0 & 0 & 3 & 4 \\ 0 & 0 & 0 & 4 \end{pmatrix} \in \mathfrak{M}_4(\mathbf{R}).$$

Possiamo subito affermare che il determinante di tale matrice è il prodotto degli elementi della diagonale. Infatti, con sviluppi successivi sempre rispetto alla prima colonna, otteniamo:

$$|A| = 1 \cdot \begin{vmatrix} 2 & 3 & 4 \\ 0 & 3 & 4 \\ 0 & 0 & 4 \end{vmatrix} = 1 \cdot 2 \cdot \begin{vmatrix} 3 & 4 \\ 0 & 4 \end{vmatrix} = 1 \cdot 2 \cdot 3 \cdot |4| = 1 \cdot 2 \cdot 3 \cdot 4 = 24.$$

Con lo stesso ragionamento si prova che ogni matrice triangolare superiore o inferiore ha come determinante il prodotto degli elementi della sua diagonale.

Assegnata una matrice $A \in \mathfrak{M}_n(K)$, possiamo considerare la matrice dei complementi algebrici dei suoi elementi. Si tratta della matrice

$$\begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ \alpha_{n1} & \alpha_{n2} & \cdots & \alpha_{nn} \end{pmatrix} \in \mathfrak{M}_n(K),$$

che è chiamata *matrice cofattore* (o *matrice dei complementi algebrici*) di A e che denoteremo \mathcal{C}_A .

Tale matrice ha un'importante proprietà, che ora illustriamo.

Data una matrice $A \in \mathfrak{M}_n(K)$, sia $B \in \mathfrak{M}_n(K)$ la matrice ottenuta da A sostituendo la colonna $A_{(1)}$ con la colonna $A_{(2)}$. La matrice B ha le prime due colonne uguali e quindi [in base alla Prop. 1(vi)] $\det(B) = 0$. Con lo sviluppo di Laplace di $\det(B)$ rispetto alla prima colonna, si ottiene:

$$0 = \det(B) = \sum_{t=1}^n b_{t1} \alpha_{t1}(B),$$

dove $\alpha_{t1}(B)$ è il complemento algebrico in B dell'elemento b_{t1} . Ma $b_{t1} = a_{t2}$ e $\alpha_{t1}(B) = \alpha_{t1}$ [complemento algebrico di a_{t1} in A], in quanto nel calcolo di $\alpha_{t1}(B)$ la prima colonna di B viene eliminata ed il resto della matrice B coincide con la corrispondente parte di A . Abbiamo quindi

$$\sum_{t=1}^n a_{t2} \alpha_{t1} = 0.$$

Alla stessa conclusione si perviene rimpiazzando la colonna j -sima con la i -sima (con $i \neq j$) ovvero rimpiazzando la riga j -sima con la i -sima (sempre con $i \neq j$) [e si usa in tal caso lo sviluppo di Laplace rispetto alle righe]. Otteniamo quindi

$$\sum_{t=1}^n a_{ti} \alpha_{tj} = 0 \quad \text{e} \quad \sum_{t=1}^n a_{it} \alpha_{jt} = 0, \quad \text{se } i \neq j.$$

In forma di prodotto righe per colonne le due precedenti uguaglianze si riscrivono rispettivamente:

$$({}^t\mathcal{C}_A)^{(j)} A_{(i)} = 0 \quad \text{e} \quad A^{(i)} ({}^t\mathcal{C}_A)_{(j)} = 0, \quad \text{se } i \neq j.$$

Infine i due sviluppi di Laplace riportati nell'enunciato del Teor. 1 si riscrivono rispettivamente nella forma

$$\det(A) = A^{(i)} ({}^t\mathcal{C}_A)_{(i)}, \quad \det(A) = ({}^t\mathcal{C}_A)^{(j)} A_{(j)}, \quad \forall i, j = 1, \dots, n.$$

Tenuto conto che $({}^t\mathcal{C}_A)^{(j)} A_{(i)} = ({}^t\mathcal{C}_A A)_{ji}$ e $A^{(i)} ({}^t\mathcal{C}_A)_{(j)} = (A {}^t\mathcal{C}_A)_{ij}$, abbiamo ottenuto il seguente risultato.

Proposizione 2. Sia $A \in \mathfrak{M}_n(K)$, con $n \geq 2$. Risulta:

$${}^t\mathcal{C}_A A = \det(A) I_n = A {}^t\mathcal{C}_A.$$

Avevamo osservato in Prop. 1(viii) che una matrice invertibile A ha determinante non nullo. Ora dalla Prop. 2 segue subito che, se $\det(A) \neq 0$, allora A è invertibile. Infatti, moltiplicando la

precedente uguaglianza per $\frac{1}{\det(A)}$ si ottiene

$$\left(\frac{1}{\det(A)} {}^t \mathcal{C}_A \right) A = I_n = A \left(\frac{1}{\det(A)} {}^t \mathcal{C}_A \right).$$

Pertanto abbiamo provato il seguente risultato.

Proposizione 3. *Sia $A \in \mathfrak{M}_n(K)$, con $n \geq 2$. Risulta:*

$$A \text{ è invertibile [cioè } A \in \mathbf{GL}_n(K)] \iff \det(A) \neq 0.$$

Se A è invertibile si ha:

$$A^{-1} = \frac{1}{\det(A)} {}^t \mathcal{C}_A.$$

La precedente proposizione ci dice che per calcolare l'inversa di una matrice occorre calcolare la matrice dei complementi algebrici, trasporla e poi dividere tutti gli elementi per $\det(A)$. Ad esempio, assegnata la matrice $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbf{GL}_2(K)$, risulta:

$$\mathcal{C}_A = \begin{pmatrix} d & -c \\ -b & a \end{pmatrix} \text{ e dunque } A^{-1} = \frac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \begin{pmatrix} \frac{d}{ad-bc} & \frac{-b}{ad-bc} \\ \frac{-c}{ad-bc} & \frac{a}{ad-bc} \end{pmatrix}.$$

Il Teorema di Laplace sopra enunciato è in realtà un caso particolare del risultato effettivamente dimostrato da Laplace. Invece di eseguire lo sviluppo del determinante rispetto ad una riga o ad una colonna, Laplace ha dimostrato che è possibile, più generalmente, eseguire lo sviluppo rispetto ad "un certo numero" di righe o di colonne. Abbiamo innanzitutto bisogno di una definizione.

Definizione 2. *Sia $M \in \mathfrak{M}_{m,n}(K)$. Siano p, q interi positivi tali che $1 \leq p \leq m$ e $1 \leq q \leq n$. Si scelgano p interi $\{i_1, i_2, \dots, i_p\}$ tali che*

$$1 \leq i_1 < i_2 < \dots < i_p \leq m;$$

si scelgano inoltre q interi $\{j_1, j_2, \dots, j_q\}$ tali che

$$1 \leq j_1 < j_2 < \dots < j_q \leq n.$$

Si chiama sottomatrice di M relativa alle righe i_1, \dots, i_p ed alle colonne j_1, \dots, j_q la matrice formata dagli elementi di M ottenuti intersecando le righe $M^{(i_1)}, \dots, M^{(i_p)}$ con le colonne $M_{(j_1)}, \dots, M_{(j_q)}$. Tale sottomatrice verrà indicata $M(i_1, \dots, i_p | j_1, \dots, j_q)$.

Ad esempio, se A è la matrice triangolare superiore di $\mathfrak{M}_4(\mathbf{R})$ considerata sopra, risulta ad esempio:

$$A(2, 3 | 2, 4) = \begin{pmatrix} 2 & 4 \\ 0 & 4 \end{pmatrix}, \quad A(2, 3, 4 | 1, 2, 4) = \begin{pmatrix} 0 & 2 & 4 \\ 0 & 0 & 4 \\ 0 & 0 & 4 \end{pmatrix}.$$

Data una sottomatrice quadrata B di una matrice quadrata A , si chiama *complemento algebrico di B* il determinante della sottomatrice ottenuta con le righe e le colonne residue rispetto a quelle di B , con un segno dato dalla parità delle somme degli indici delle righe e delle colonne di B . Ad esempio la sottomatrice $A(2, 3 | 2, 4)$ determina come suo complemento algebrico

$$(-1)^{2+3+2+4} \det(A(1, 4 | 1, 3)) = (-1)^{11} \begin{vmatrix} 1 & 3 \\ 0 & 0 \end{vmatrix} = 0,$$

mentre il complemento algebrico di $A(2, 3, 4 | 1, 2, 4)$ è

$$(-1)^{2+3+4+1+2+4} \det(A(1 | 3)) = (-1)^{16} |3| = 3.$$

Possiamo ora enunciare la versione completa del Teorema di Laplace

Teorema 2. (Teorema di Laplace - versione completa) *Sia $A \in \mathfrak{M}_n(K)$ e sia $t \in \mathbf{N}$, con $1 \leq t < n$. Si fissino t righe (o t colonne) di A e sia B la matrice da esse formata. Si considerino ora tutte le*

sottomatrici quadrate di B di ordine t . Risulta: $\det(A)$ coincide con la somma dei determinanti di tali sottomatrici, moltiplicati ciascuno per il rispettivo complemento algebrico.

[Si noti che il **Teor. 1** è il caso $t = 1$].

Consideriamo ancora la precedente matrice $A \in \mathfrak{M}_4(\mathbf{R})$ e fissiamone ad esempio le prime due righe. Otteniamo la sottomatrice $B = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 2 & 3 & 4 \end{pmatrix}$. Tale matrice possiede sei sottomatrici quadrate di ordine 2, i cui determinanti sono:

$$\begin{vmatrix} 1 & 2 \\ 0 & 2 \end{vmatrix} = 2, \quad \begin{vmatrix} 1 & 3 \\ 0 & 3 \end{vmatrix} = 3, \quad \begin{vmatrix} 1 & 4 \\ 0 & 4 \end{vmatrix} = 4, \quad \begin{vmatrix} 2 & 3 \\ 2 & 3 \end{vmatrix} = 0, \quad \begin{vmatrix} 2 & 4 \\ 2 & 4 \end{vmatrix} = 0, \quad \begin{vmatrix} 3 & 4 \\ 3 & 4 \end{vmatrix} = 0.$$

Val la pena di calcolare soltanto i complementi algebrici delle prime tre sottomatrici sopra considerate (perché le ultime tre hanno determinante nullo). Sono rispettivamente:

$$(-1)^{1+2+1+2} \begin{vmatrix} 3 & 4 \\ 0 & 4 \end{vmatrix} = 12, \quad (-1)^{1+2+1+3} \begin{vmatrix} 0 & 4 \\ 0 & 4 \end{vmatrix} = 0, \quad (-1)^{1+2+1+4} \begin{vmatrix} 0 & 3 \\ 0 & 0 \end{vmatrix} = 0.$$

Segue che $\det(A) = 2 \cdot 12 + 3 \cdot 0 + 4 \cdot 0 + 0 + 0 + 0 = 24$ (come già noto).

È evidente che la migliore scelta delle righe o delle colonne rispetto a cui eseguire lo sviluppo di Laplace è quella che diminuisce il più possibile i calcoli. Se nell'esempio precedente avessimo scelto come matrice B quella formata dalle ultime due righe di A , avremmo avuto una sola sottomatrice di ordine 2 a determinante non nullo, cioè $A(3, 4 | 3, 4) = \begin{pmatrix} 3 & 4 \\ 0 & 4 \end{pmatrix}$. Il suo complemento algebrico è $(-1)^{3+4+3+4} \begin{vmatrix} 1 & 2 \\ 0 & 2 \end{vmatrix} = 2$ e dunque $\det(A) = \begin{vmatrix} 3 & 4 \\ 0 & 4 \end{vmatrix} \cdot 2 = 24$. Analogo risultato avremmo ottenuto scegliendo come sottomatrice B quella formata dalle prime due colonne di A .

Il determinante di una matrice quadrata può essere calcolato anche utilizzando l'algoritmo di Gauss, ma con qualche variante. Sia $A \in \mathfrak{M}_n(K)$. Osserviamo subito che:

I[$A^{(i)} \leftrightarrow A^{(j)}$] cambia il segno al determinante [cfr. **Prop. 1(v)**];

II[$A^{(i)} \rightarrow c A^{(i)}$] moltiplica il determinante per c [cfr. **Prop. 1(iv)**];

III[$A^{(i)} \rightarrow A^{(i)} + c A^{(j)}$] lascia invariato il determinante [cfr. **Prop. 1(iv), (vi)**];

un eventuale scambio di colonne cambia il segno del determinante.

Conviene allora applicare alla matrice A l'algoritmo di Gauss senza però eliminare eventuali righe nulle e senza utilizzare operazioni di tipo **II**, cioè senza imporre che gli elementi di posto (i, i) siano ridotti ad 1. In questo modo ad ogni passo il determinante resta lo stesso o cambia al più di segno. Alla conclusione dell'algoritmo avremo una matrice triangolare superiore, il cui determinante, come già osservato, è il prodotto degli elementi della diagonale. Basta allora cambiare il segno di tale prodotto, tante volte quanti sono stati gli scambi di riga (o di colonna) effettuati e si ottiene $\det(A)$.

Calcoliamo con lo sviluppo di Laplace e con l'algoritmo di Gauss il determinante della matrice

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

Eseguiamo lo sviluppo di Laplace ad esempio rispetto alla prima colonna. Otteniamo:

$$\det(A) = 2 \cdot (-1)^{3+1} \begin{vmatrix} -1 & 2 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & -1 \end{vmatrix} + 2 \cdot (-1)^{4+1} \begin{vmatrix} -1 & 2 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = 2(3) - 2(-3) = 12.$$

Eseguiamo ora l'algoritmo di Gauss. Con **I**[(1^a) \leftrightarrow (3^a)] otteniamo:

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

Con $\text{III}[(4^a) \rightarrow (4^a) - (1^a)]$ otteniamo:

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

Con $\text{III}[(3^a) \rightarrow (3^a) + (2^a)]$ e $\text{III}[(4^a) \rightarrow (4^a) - (2^a)]$ otteniamo:

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

La matrice ottenuta è triangolare superiore ed il suo determinante è -12 . Poiché però abbiamo eseguito uno scambio di riga, cambiamo segno ed otteniamo, come sopra, $\det(A) = -(-12) = 12$.

ESERCIZI PROPOSTI

3.2.1. Eseguire il calcolo del determinante delle due seguenti matrici, sfruttando opportunamente le proprietà dei determinanti (per semplificare il calcolo)

$$A = \begin{pmatrix} 999 & 1000 & 1001 \\ 1 & 2 & 3 \\ 0 & -1 & 2 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 50 & 0 \\ 2 & 100 & -1 \\ 3 & 75 & 2 \end{pmatrix}.$$

3.2.2 Per ogni $\alpha \in \mathbf{R}$, si considerino le due matrici

$$A = \begin{pmatrix} \cos \alpha & \sin \alpha \\ \sin \alpha & -\cos \alpha \end{pmatrix}, \quad B = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \in \mathfrak{M}_2(\mathbf{R}).$$

Verificare che sono entrambe invertibili e che $A^{-1} = {}^t A$, $B^{-1} = {}^t B$.

3.2.3 Assegnata la matrice quadrata

$$A = \begin{pmatrix} 1 & 2 & -1 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix} \in \mathfrak{M}_3(\mathbf{R}),$$

(i) Calcolare la matrice dei complementi algebrici \mathcal{C}_A e verificare che $\det(\mathcal{C}_A) = (\det(A))^2$.

(ii) Dimostrare che, $\forall A \in \mathbf{GL}_n(K)$, risulta: $\det(\mathcal{C}_A) = (\det(A))^{n-1}$.

3.2.4. Consideriamo le matrici di $\mathfrak{M}_2(\mathbf{Z}_2)$. Quante sono? Quante e quali di esse sono quelle invertibili? Il gruppo che esse formano è commutativo?

3.2.5. Verificare che ogni matrice antisimmetrica reale di ordine dispari ha determinante nullo (cfr. Eserc. 2.5.7).

3.2.6. È assegnata in $\mathfrak{M}_4(\mathbf{Q})$ la matrice

$$A = \begin{pmatrix} a+1 & a-1 & 0 & a \\ 0 & 1 & 1 & 1 \\ b-1 & b & b+1 & 0 \\ 1 & -1 & 0 & 0 \end{pmatrix}, \text{ con } a, b \in \mathbf{Z}.$$

Verificare, senza eseguire il calcolo diretto del determinante ma usando solo le proprietà dei determinanti, che $\det(A)$ è un multiplo intero di ab .

3.2.7. Sia $\det : \mathbf{GL}_n(K) \rightarrow K^\times$ l'applicazione che associa ad ogni matrice quadrata invertibile A il rispettivo determinante $\det(A)$.

Verificare che \det è un epimorfismo del gruppo $(\mathbf{GL}_n(K), \cdot)$ sul gruppo (K^\times, \cdot) e calcolarne il nucleo.

3.2.8. Sia $H = \{A \in \mathbf{GL}_n(K) \mid \det(A) = \pm 1\}$. Verificare che H è nucleo di un omomorfismo di gruppi $\varphi : \mathbf{GL}_n(K) \rightarrow K^\times$. Determinare poi l'immagine $Im(\varphi)$.

3. Rango di una matrice

Una matrice non quadrata $A \in \mathfrak{M}_{m,n}(K)$ non possiede il determinante, ma possiede varie sottomatrici quadrate e ciascuna di esse ha un determinante. Se A possiede una sottomatrice quadrata di ordine t con determinante non nullo, mentre tutte le sottomatrici quadrate di A di ordine $> t$ (ammesso che ce ne siano) hanno determinante nullo, diremo che A ha *rango* t . Ma per il momento dimentichiamoci di questa definizione.

Il concetto di rango, come vedremo nel prossimo paragrafo, ha un'importanza centrale nella risoluzione dei SL ed è collegato all'indipendenza lineare delle righe e delle colonne di A . Cominciamo a dare la definizione di rango partendo proprio da questo aspetto.

Definizione 1. Sia $A \in \mathfrak{M}_{m,n}(K)$. Indichiamo con $A^{(1)}, A^{(2)}, \dots, A^{(m)}$ le sue m righe (sono elementi di $\mathfrak{M}_{1,n}(K)$) e con $A_{(1)}, A_{(2)}, \dots, A_{(n)}$ le sue n colonne (sono elementi di $\mathfrak{M}_{m,1}(K)$).

Chiameremo *rango per righe* (o *rango-righe*) di A , denotato \mathbf{r}_A , la dimensione del K -sottospazio vettoriale di $\mathfrak{M}_{1,n}(K)$ generato dalle righe di A , cioè

$$\mathbf{r}_A = \dim(\langle A^{(1)}, A^{(2)}, \dots, A^{(m)} \rangle).$$

[In altri termini, \mathbf{r}_A è il massimo numero di righe linearmente indipendenti di A].

Chiameremo analogamente *rango per colonne* (o *rango-colonne*) di A , denotato \mathbf{c}_A , la dimensione del K -sottospazio vettoriale di $\mathfrak{M}_{m,1}(K)$ generato dalle colonne di A , cioè

$$\mathbf{c}_A = \dim(\langle A_{(1)}, A_{(2)}, \dots, A_{(n)} \rangle).$$

[In altri termini, \mathbf{c}_A è il massimo numero di colonne linearmente indipendenti di A].

La prima cosa da osservare è che, essendo il sottospazio vettoriale $\langle A^{(1)}, A^{(2)}, \dots, A^{(m)} \rangle$ generato da m vettori ed essendo $\dim(\mathfrak{M}_{1,n}(K)) = n$, allora $\mathbf{r}_A \leq \min(m, n)$.

Per analoghi motivi, si osserva subito che anche $\mathbf{c}_A \leq \min(m, n)$.

Vale il seguente risultato, che non dimostreremo.

Teorema 1. Per ogni matrice $A \in \mathfrak{M}_{m,n}(K)$, risulta: $\mathbf{r}_A = \mathbf{c}_A$.

Visto che il rango per righe ed il rango per colonne coincidono, tanto vale chiamarlo semplicemente *rango*. Dunque introduciamo la seguente definizione.

Definizione 2. Data una matrice $A \in \mathfrak{M}_{m,n}(K)$, il numero naturale $\mathbf{r}_A [= \mathbf{c}_A]$ è detto *rango di A* ed è denotato $rg(A)$. Si tratta del massimo numero di righe linearmente indipendenti di A [e coincide con il massimo numero di colonne linearmente indipendenti di A].

A titolo di esempio calcoliamo il rango della matrice

$$A = \begin{pmatrix} 1 & 2 & -1 \\ 0 & 1 & 1 \\ 1 & 3 & 0 \end{pmatrix} \in \mathfrak{M}_3(\mathbf{R}).$$

Le prime due righe di A sono linearmente indipendenti. Infatti il $SLO(3, 2, \mathbf{R})$

$$a A^{(1)} + b A^{(2)} = \mathbf{0}, \text{ cioè } \begin{cases} a = 0 \\ 2a + b = 0 \\ -a + b = 0 \end{cases}$$

non ammette autosoluzioni. La terza riga è invece combinazione lineare delle prime due. Infatti il $SL(3, 2, \mathbf{R})$

$$a A^{(1)} + b A^{(2)} = A^{(3)}, \text{ cioè } \begin{cases} a = 1 \\ 2a + b = 3 \\ -a + b = 0, \end{cases}$$

è risolubile, con soluzione $(a, b) = (1, 1)$. Dunque $A^{(1)} + A^{(2)} = A^{(3)}$. Ne segue che $rg(A) = 2$.

Si noti che, dal **Teor. 1**, le tre colonne di A devono essere linearmente dipendenti. Poiché la prime due sono linearmente indipendenti [si verifichi infatti che il $SLO(3, 2, \mathbf{R})$ $a A_{(1)} + b A_{(2)} = \mathbf{0}$ non ha autosoluzioni], la terza colonna è necessariamente combinazione lineare delle prime due, ovvero il $SL(3, 2, \mathbf{R})$

$$a A_{(1)} + b A_{(2)} = A_{(3)}, \text{ cioè } \begin{cases} a + 2b = -1 \\ b = 1 \\ a + 3b = 0, \end{cases}$$

è risolubile. Infatti ammette soluzione $(a, b) = (-3, 1)$ e quindi $A_{(3)} = -3 A_{(1)} + A_{(2)}$.

Osservazione 1. (i) Si osserva subito che $rg(A) = rg({}^t A)$, $\forall A \in \mathfrak{M}_{m,n}(K)$.

Infatti le righe di ${}^t A$ sono le colonne di A . Dunque $\mathbf{r}_{{}^t A} = \mathbf{c}_A$ e quindi $rg({}^t A) = rg(A)$.

(ii) Si può verificare che con operazioni elementari di riga il rango di una matrice non cambia. Sia infatti $A \in \mathfrak{M}_{m,n}(K)$ e sia ad esempio B la matrice ottenuta da A con l'operazione elementare $III[A^{(i)} \rightarrow A^{(i)} + cA^{(j)}]$. Le righe di B sono

$$A^{(1)}, \dots, A^{(i)} + cA^{(j)}, \dots, A^{(j)}, \dots, A^{(m)}$$

e lo spazio vettoriale da esse generato coincide, come facilmente si può verificare, con quello generato da $A^{(1)}, \dots, A^{(m)}$. Dunque $rg(B) = rg(A)$.

Considerazioni analoghe si fanno per le operazioni elementari di riga di primo e secondo tipo.

(iii) Se B è una sottomatrice di A , verifichiamo che $rg(B) \leq rg(A)$.

Assumiamo che sia $B = A(i_1, \dots, i_p | j_1, \dots, j_q)$, con $1 \leq p \leq m$ e $1 \leq q \leq n$. Consideriamo allora la matrice $M = A(i_1, \dots, i_p | 1, \dots, n)$.

B è una sottomatrice di M formata da p colonne di M . Dunque $\mathbf{c}_B \leq \mathbf{c}_M$, cioè $rg(B) \leq rg(M)$. Analogamente, M è una sottomatrice di A formata da q righe di A . Dunque $\mathbf{r}_M \leq \mathbf{r}_A$, cioè $rg(M) \leq rg(A)$. Si conclude che $rg(B) \leq rg(M) \leq rg(A)$.

Proposizione 1. Sia $A \in \mathfrak{M}_n(K)$. Risulta:

$$A \in \mathbf{GL}_n(K) \iff \det(A) \neq 0 \iff rg(A) = n.$$

Dim. La prima equivalenza è già stata dimostrata (cfr. **Prop. 2.3**). È sufficiente allora provare che

$$(i) \det(A) \neq 0 \implies rg(A) = n; \quad (ii) \ rg(A) = n \implies A \in \mathbf{GL}_n(K).$$

Proviamo (i). Se per assurdo fosse $rg(A) < n$, una riga (ad esempio la prima) sarebbe combinazione lineare delle altre. Dunque

$$A^{(1)} = c_2 A^{(2)} + c_3 A^{(3)} + \dots + c_n A^{(n)}.$$

Dalla **Prop. 2.1(iv)**,

$$\det(A) = c_2 \det \begin{pmatrix} A^{(2)} \\ A^{(2)} \\ \vdots \\ A^{(n)} \end{pmatrix} + c_3 \det \begin{pmatrix} A^{(3)} \\ A^{(2)} \\ \vdots \\ A^{(n)} \end{pmatrix} + \dots + c_n \det \begin{pmatrix} A^{(n)} \\ A^{(2)} \\ \vdots \\ A^{(n)} \end{pmatrix}.$$

Le matrici a secondo membro hanno tutte due righe uguali e quindi hanno determinante nullo. Ne segue che $\det(A) = 0$, contro l'ipotesi.

Proviamo ora (ii). Per ipotesi, le righe $A^{(1)}, \dots, A^{(n)}$ sono una base di $\mathfrak{M}_{1,n}(K)$. Consideriamo le matrici riga elementari $E^1 = (1 \ 0 \ 0 \ \dots \ 0), \dots, E^n = (0 \ 0 \ \dots \ 0 \ 1)$. Possiamo esprimere ciascuna di esse come combinazione lineare della base $\{A^{(1)}, \dots, A^{(n)}\}$. Dunque, per $i = 1, \dots, n$:

$$E^i = \sum_{j=1}^n b_{ij} A^{(j)} = (b_{i1} \ b_{i2} \ \dots \ b_{in}) \begin{pmatrix} A^{(1)} \\ \vdots \\ A^{(n)} \end{pmatrix} = (b_{i1} \ b_{i2} \ \dots \ b_{in}) A.$$

I coefficienti b_{ij} formano una matrice $B \in \mathfrak{M}_n(K)$. Si ha quindi: $E^i = B^{(i)} A$ e pertanto

$$I_n = \begin{pmatrix} E^1 \\ \vdots \\ E^n \end{pmatrix} = B A.$$

Dunque $B A = I_n$, cioè B inverte "a sinistra" A .

Se ora consideriamo le colonne $A_{(1)}, \dots, A_{(n)}$ di A , anch'esse formano una base [di $\mathfrak{M}_{n,1}(K)$].

Possiamo esprimere le n matrici colonna elementari $E_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots, E_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$ come combinazione

lineare della base $\{A_{(1)}, \dots, A_{(n)}\}$. Otteniamo

$$E_j = \sum_{i=1}^n c_{ij} A_{(i)} = (A_{(1)} \ \dots \ A_{(n)}) \begin{pmatrix} c_{1j} \\ \vdots \\ c_{nj} \end{pmatrix} = A \begin{pmatrix} c_{1j} \\ \vdots \\ c_{nj} \end{pmatrix}.$$

Definita la matrice $C = (c_{ij}) \in \mathfrak{M}_n(K)$, si ha: $E_j = A C_{(j)}$ e dunque

$$I_n = (E_1 \ E_2 \ \dots \ E_n) = A C.$$

Pertanto $A C = I_n$, cioè C inverte "a destra" A .

Da $AC = I_n = BA$ segue: $B = BI_n = B(AC) = (BA)C = I_n C = C$, cioè $B = C$. Pertanto A è invertibile (con inversa B).

Proposizione 2. Se $A \in \mathfrak{M}_{m,n}(K)$ e $B \in \mathfrak{M}_{n,p}(K)$, risulta:

$$\operatorname{rg}(AB) \leq \min\{\operatorname{rg}(A), \operatorname{rg}(B)\}.$$

Se poi $A \in \mathfrak{M}_{m,n}(K)$ e $B \in \mathbf{GL}_n(K)$, allora $\operatorname{rg}(AB) = \operatorname{rg}(A)$. Analogamente, se $C \in \mathbf{GL}_m(K)$, allora $\operatorname{rg}(CA) = \operatorname{rg}(A)$. Dunque la moltiplicazione per una matrice invertibile non modifica il rango.

Dim. Per definizione,

$$\operatorname{rg}(AB) = \dim \langle (AB)^{(1)}, \dots, (AB)^{(m)} \rangle.$$

Si ha:

$$(AB)^{(1)} = (A^{(1)} B_{(1)} \ \dots \ A^{(1)} B_{(p)}) = (\sum_{t=1}^n a_{1t} b_{t1} \ \dots \ \sum_{t=1}^n a_{1t} b_{tp}) = \sum_{t=1}^n a_{1t} (b_{t1} \ \dots \ b_{tp}) = \sum_{t=1}^n a_{1t} B^{(t)}.$$

Dunque $(AB)^{(1)} \in \langle B^{(1)}, \dots, B^{(n)} \rangle$.

Allo stesso modo si verifica che $(AB)^{(2)}, \dots, (AB)^{(m)} \in \langle B^{(1)}, \dots, B^{(n)} \rangle$. Quindi

$$\langle (AB)^{(1)}, \dots, (AB)^{(m)} \rangle \subseteq \langle B^{(1)}, \dots, B^{(n)} \rangle$$

e pertanto $\operatorname{rg}(AB) \leq \operatorname{rg}(B)$ [cioè il rango del prodotto di due matrici è minore o uguale al rango della seconda matrice del prodotto].

Utilizzando tale disegualanza e la proprietà del rango di **Osserv. 1(i)** si ottiene

$$\operatorname{rg}(AB) = \operatorname{rg}({}^t(AB)) = \operatorname{rg}({}^t B {}^t A) \leq \operatorname{rg}({}^t A) = \operatorname{rg}(A),$$

cioè $\operatorname{rg}(AB) \leq \operatorname{rg}(A)$. È così provato che $\operatorname{rg}(AB) \leq \min\{\operatorname{rg}(A), \operatorname{rg}(B)\}$.

Proviamo ora che, se $A \in \mathfrak{M}_{m,n}(K)$ e $B \in \mathbf{GL}_n(K)$, allora $\operatorname{rg}(AB) = \operatorname{rg}(A)$. Infatti $A = (AB)B^{-1}$ e quindi $\operatorname{rg}(A) = \operatorname{rg}((AB)B^{-1}) \leq \operatorname{rg}(AB) \leq \operatorname{rg}(A)$.

Analogamente, se $C \in \mathbf{GL}_m(K)$, $A = C^{-1}C A$ e quindi $\operatorname{rg}(A) = \operatorname{rg}(C^{-1}C A) \leq \operatorname{rg}(CA) \leq \operatorname{rg}(A)$, cioè $\operatorname{rg}(CA) = \operatorname{rg}(A)$.

Come accennato all'inizio del paragrafo, il rango si collega all'annullamento dei determinanti delle

sottomatrici quadrate di A . Vediamo come, partendo da una definizione.

Definizione 3. Data una matrice $A \in \mathfrak{M}_{m,n}(K)$, sia $\rho = \rho(A)$ l'intero definito dalle due seguenti condizioni:

- esiste in A (almeno) una sottomatrice quadrata invertibile di ordine ρ ;
- le sottomatrici quadrate di A di ordine $> \rho$ (se ne esistono) hanno determinante nullo.

Chiameremo inoltre minore (di ordine t) di A il determinante di una sottomatrice quadrata di A di ordine t . Pertanto $\rho = \rho(A)$ è definito dalle due condizioni:

esiste in A un minore non nullo di ordine ρ ; i minori di ordine $> \rho$ (se esistono) sono nulli.

Possiamo quindi dire che ρ è l'ordine massimo dei minori non nulli di A .

Teorema 2. Per ogni matrice $A \in \mathfrak{M}_{m,n}(K)$, risulta: $rg(A) = \rho(A)$.

Dim. Verifichiamo che $\rho(A) \leq rg(A)$.

Poniamo $\rho := \rho(A)$. Esiste in A una sottomatrice quadrata invertibile B di ordine ρ . Dalla **Prop. 1** e dall'**Osserv. 1(iii)**, $\rho = rg(B) \leq rg(A)$.

Viceversa, verifichiamo che $rg(A) \leq \rho(A)$.

Poniamo $r := rg(A)$. Sceglio in A r righe linearmente indipendenti e sia M la sottomatrice di A formata da tali righe. Ovviamente $rg(M) = r$. In M esistono quindi r colonne linearmente indipendenti. La matrice B formata da tali colonne (di M) è una sottomatrice quadrata di A avente rango r e quindi invertibile. Pertanto $|B|$ è un minore non nullo di A e quindi $r = rg(A) \leq \rho(A)$.

Si noti che se una matrice $A \in \mathfrak{M}_{m,n}(K)$ ha nulli tutti i minori di un dato ordine t , sono nulli anche tutti gli eventuali minori di ordine superiore a t . Ciò segue subito dal teorema di Laplace.

Vogliamo calcolare il rango di una matrice $A \in \mathfrak{M}_{m,n}(K)$, utilizzando il **Teor. 2**. Procederemo come segue:

- si individua in A una sottomatrice quadrata invertibile di ordine t . Allora $rg(A) \geq t$.
- se $t = \min\{m, n\}$, non esistono in A minori di ordine $t+1$ e si conclude che $rg(A) = t$.
- se invece $t < \min\{m, n\}$, è necessario calcolare i minori di ordine $t+1$. Tali minori sono quanti le possibili scelte di $t+1$ righe [tra le m righe di A] per le possibili scelte di $t+1$ colonne [tra le n colonne di A]. Dunque sono complessivamente $\binom{m}{t+1} \binom{n}{t+1}$. Se tali minori sono tutti nulli, concludiamo che $rg(A) = t$; altrimenti possiamo solo dire che $rg(A) \geq t+1$ e dobbiamo procedere al calcolo dei minori di ordine $t+2$ [se ne esistono, cioè se $t+1 < \min\{m, n\}$].

In questo modo arriveremo dopo alcuni passi all'individuazione del rango. Ma i calcoli da fare sono decisamente troppi. Ad esempio, se $A \in \mathfrak{M}_{4,6}(K)$ ed abbiamo individuato una sottomatrice quadrata invertibile di ordine 2, i minori di ordine 3 sono $\binom{4}{3} \binom{6}{3} = 80$. Se uno di essi è non nullo, allora $rg(A) \geq 3$ e dovremo considerare i minori di ordine 4, che sono $\binom{4}{4} \binom{6}{4} = 15$. Se tutti sono nulli, allora $rg(A) = 3$; altrimenti $rg(A) = 4$.

Il risultato che segue ci offre un sensibile sconto sul numero dei calcoli da eseguire. È noto come *principio degli orlati* o *principio dei minori orlanti* ed è dovuto a Kronecker (per questo è anche chiamato *teorema di Kronecker*). Partiamo da una definizione.

Definizione 4. Sia B una sottomatrice quadrata di ordine r di una matrice $A \in \mathfrak{M}_{m,n}(K)$. Chiameremo *orlato* di B il determinante di ogni sottomatrice C di A , quadrata, di ordine $r+1$ ed avente B come sottomatrice. [Diremo che C è ottenuta ‘*orlando*’ B con un’ulteriore riga e colonna di A]. È evidente che, se $r = \min\{n, m\}$, B non ha orlati (e viceversa).

Consideriamo ad esempio la seguente matrice $A = \begin{pmatrix} 1 & \mathbf{2} & 3 & 4 \\ 2 & 3 & 4 & 1 \\ 3 & \mathbf{4} & 1 & \mathbf{2} \\ 4 & 1 & 2 & 3 \end{pmatrix} \in \mathfrak{M}_4(\mathbf{R})$ e fissiamone la

sottomatrice $B = A(1, 3 | 2, 4) = \begin{pmatrix} 2 & 4 \\ 4 & 2 \end{pmatrix}$. I suoi orlati sono quattro [infatti dobbiamo aggiungere a B una riga tra le due disponibili (la seconda o la quarta) ed una colonna tra le due disponibili (la prima o la terza)]. Dunque i quattro orlati di B sono i determinanti delle seguenti matrici:

$$A(\mathbf{1}, 2, \mathbf{3} | 1, \mathbf{2}, \mathbf{4}) = \begin{pmatrix} 1 & 2 & 4 \\ 2 & 3 & 1 \\ 3 & 4 & 2 \end{pmatrix}, \quad A(\mathbf{1}, 2, \mathbf{3} | \mathbf{2}, 3, 4) = \begin{pmatrix} 2 & 3 & 4 \\ 3 & 4 & 1 \\ 4 & 1 & \mathbf{2} \end{pmatrix},$$

$$A(\mathbf{1}, \mathbf{3}, 4 | 1, \mathbf{2}, \mathbf{4}) = \begin{pmatrix} 1 & 2 & 4 \\ 3 & 4 & 2 \\ 4 & 1 & 3 \end{pmatrix}, \quad A(\mathbf{1}, \mathbf{3}, 4 | \mathbf{2}, 3, 4) = \begin{pmatrix} 2 & 3 & 4 \\ 4 & 1 & \mathbf{2} \\ 1 & 2 & 3 \end{pmatrix}.$$

[Si può verificare che i quattro orlati di B sono rispettivamente $-4, -44, -44, -4$].

Una volta individuata in A una sottomatrice B quadrata, invertibile e di ordine t , il principio degli orlati ci consentirà di esaminare, anziché tutti i minori di ordine $t+1$ di A , soltanto gli orlati di B . Lo "sconto" sta nel fatto che invece di esaminare l'annullamento di $\binom{m}{t+1} \binom{n}{t+1}$ minori, è sufficiente esaminare l'annullamento soltanto di $(m-t)(n-t)$ minori.

Ad esempio, nel caso di una matrice $A \in \mathfrak{M}_{4,6}(K)$, gli orlati di una sottomatrice quadrata di ordine 2 sono soltanto $(4-2)(6-2) = 8$ (mentre i minori di ordine 3 sono 80, come sopra osservato).

Enunciamo finalmente (e senza dimostrazione) il principio degli orlati.

Teorema 3. (*Principio degli orlati*) Sia $A \in \mathfrak{M}_{m,n}(K)$. Risulta:

$rg(A) = r \iff$ valgono le due seguenti condizioni:

- esiste in A una sottomatrice quadrata B invertibile di ordine r ;
- gli orlati di B (se ne esistono) sono tutti nulli.

Utilizzando il principio degli orlati, vogliamo calcolare il rango della seguente matrice $A \in \mathfrak{M}_{3,4}(\mathbf{R})$, al variare dei due parametri reali a, b :

$$A = \begin{pmatrix} a & 0 & 1 & b \\ 1 & a & 0 & 1 \\ a & b & 0 & 1 \end{pmatrix}.$$

Conviene considerare un minore che sia sempre non nullo (indipendentemente dal valore dei parametri) e che sia di ordine più grande possibile. La scelta può cadere ad esempio sul minore corrispondente alla sottomatrice $B = A(1, 2 | 3, 4) = \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix}$ (avente determinante 1). Possiamo quindi già dire che $2 \leq rg(A) \leq 3, \forall a, b \in \mathbf{R}$.

Per stabilire quando il rango è 2 e quando è 3, consideriamo i due orlati di B . Sono

$$\begin{vmatrix} a & 1 & b \\ 1 & 0 & 1 \\ a & 0 & 1 \end{vmatrix} = a - 1, \quad \begin{vmatrix} 0 & 1 & b \\ a & 0 & 1 \\ b & 0 & 1 \end{vmatrix} = b - a.$$

Pertanto

$$rg(A) = 2 \iff \begin{cases} a - 1 = 0 \\ b - a = 0 \end{cases} \iff a = b = 1.$$

Se invece $(a, b) \neq (1, 1)$, si ha che $rg(A) = 3$.

Osservazione 2. L'algoritmo di Gauss consente di calcolare il rango di una matrice.

Assegnata infatti $A \in \mathfrak{M}_{m,n}(K)$, sia $A' \in \mathfrak{M}_{m',n}(K)$ la matrice ottenuta da A eseguendo l'algoritmo di Gauss. In particolare $m - m'$ è il numero complessivo delle righe che vengono eliminate nel corso del procedimento, in quanto nulle.

In base all'**Osserv. 1(ii)**, $rg(A) = rg(A')$. La matrice A' contiene una sottomatrice quadrata di ordine m' triangolare superiore ed avente tutti 1 sulla diagonale [si tratta della matrice formata dalle prime m' colonne di A']. Tale sottomatrice ha rango m' e dunque $rg(A') = m'$, cioè $rg(A) = m'$.

Illustriamo su un esempio la procedura di calcolo del rango con l'algoritmo di Gauss. Si consideri la matrice

$$A = \begin{pmatrix} 1 & 1 & 2 & 0 \\ 2 & -2 & 0 & 4 \\ 0 & 1 & 1 & -1 \end{pmatrix} \in \mathfrak{M}_{3,4}(\mathbf{R}).$$

Applicando l'algoritmo di Gauss ad A , si ottiene la matrice $A' = \begin{pmatrix} 1 & 1 & 2 & 0 \\ 0 & 1 & 1 & -1 \end{pmatrix}$ [verificare].

Dunque $rg(A) = rg(A') = 2$.

Sappiamo che $rg({}^t A) = rg(A)$. Se quindi applichiamo l'algoritmo di Gauss alla matrice ${}^t A$ (anziché ad A) otterremo ancora una matrice di rango 2. Infatti l'algoritmo di Gauss applicato ad ${}^t A$ fornisce la matrice $\begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & -\frac{1}{4} \end{pmatrix}$ [verificare].

ESERCIZI PROPOSTI

3.3.1. Sia $A = \begin{pmatrix} 1 & -1 & 2 \\ 0 & -2 & 1 \\ 0 & -2 & 1 \\ -1 & -5 & 1 \\ 0 & 4 & -2 \end{pmatrix} \in \mathfrak{M}_{5,3}(\mathbf{R})$. Calcolare $rg(A)$, come massimo numero di colonne linearmente indipendenti di A .

3.3.2 Calcolare il rango della seguente matrice A , interpretandolo sia come massimo numero di righe linearmente indipendenti, sia come ordine massimo dei minori non nulli:

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

3.3.3 Al variare di $a, b \in \mathbf{R}$, descrivere il rango della matrice

$$A = \begin{pmatrix} a & 1 & b \\ 1 & a & b \\ b & 0 & 1 \end{pmatrix}.$$

3.3.4. Sia $A = \begin{pmatrix} 0 & a & b \\ -a & 0 & c \\ -b & -c & 0 \end{pmatrix} \in \mathfrak{M}_3(\mathbf{R})$, con $a \neq 0$.

Verificare che $rg(A) = 2$ ed esprimere la terza riga come combinazione lineare delle prime due.

3.3.5. Sia $A \in \mathfrak{M}_2(\mathbf{R})$. Verificare che

$$rg(A^2) < rg(A) \iff A = \begin{pmatrix} a & b \\ c & -a \end{pmatrix}, \text{ con } a^2 + bc = 0 \text{ e } a, b, c \text{ non tutti nulli.}$$

3.3.6. Sono assegnate le tre matrici (a valori reali) dipendenti da un parametro $a \in \mathbf{R}$:

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

Sia $M = ABC$. Determinare $rg(M)$, al variare di $a \in \mathbf{R}$.

4. Risoluzione di sistemi di equazioni lineari

In questo paragrafo vedremo come risolvere sistemi di equazioni lineari, evitando il ricorso all'algoritmo di Gauss ed utilizzando invece le nozioni e le proprietà di determinante e rango di matrici, introdotte nei due paragrafi precedenti.

Avremo bisogno di due teoremi: il *teorema di Rouché-Capelli* ed il *teorema di Cramer*. Cominciamo da quest'ultimo, che risolve completamente il problema della risoluzione dei sistemi di equazione lineari "quadrati" [cioè $SL(n, n, K)$] aventi matrice dei coefficienti invertibile.

Sia $AX = \mathbf{b}$ un $SL(n, n, K)$, con $A \in \mathbf{GL}_n(K)$. Proveremo che tale SL ammette un'unica soluzione. Tale soluzione è data da una formula, detta *formula di Cramer*, che ora descriviamo.

Per $i = 1, \dots, n$, sia B_i la matrice quadrata ottenuta da A sostituendo la i -sima colonna di A con la colonna dei termini noti \mathbf{b} . L'unica soluzione del nostro SL sarà data dalla n -pla

$$\frac{1}{\det A} (\det B_1, \det B_2, \dots, \det B_n).$$

Verifichiamo la formula su un esempio numerico. È assegnato il $SL(3, 3, \mathbf{R})$

$$\begin{cases} -x + y = 1 \\ 2x + z = 0 \\ -x + 2y - 2z = -1. \end{cases}$$

La matrice di tale sistema è

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

Si tratta di una matrice invertibile (essendo $\det A = 5$). Le tre matrici B_i sono

$$B_1 = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & 2 & -2 \end{pmatrix}, \quad B_2 = \begin{pmatrix} -1 & 1 & 0 \\ 2 & 0 & 1 \\ -1 & -1 & -2 \end{pmatrix} \quad B_3 = \begin{pmatrix} -1 & 1 & 1 \\ 2 & 0 & 0 \\ -1 & 2 & -1 \end{pmatrix}.$$

Poiché $\det B_1 = -3$, $\det B_2 = 2$, $\det B_3 = 6$, l'unica soluzione del SL è la terna

$$\left(-\frac{3}{5}, \frac{2}{5}, \frac{6}{5}\right)$$

[come facilmente si verifica].

Teorema 1. (*Teorema di Cramer*) Sia $AX = \mathbf{b}$ un $SL(n, n, K)$, con $A \in \mathbf{GL}_n(K)$. Tale sistema ammette un'unica soluzione, data dalla n -pla

$$\frac{1}{\det A} (\det B_1, \det B_2, \dots, \det B_n)$$

dove, per $i = 1, \dots, n$, B_i è la matrice quadrata ottenuta da A sostituendone la i -sima colonna con la colonna \mathbf{b} dei termini noti.

Dim. Che il SL ammetta un'unica soluzione è molto semplice da verificare (ricordando di interpretare le soluzioni del SL come colonne, invece che come n -ple). Intanto si osserva subito che la colonna $A^{-1}\mathbf{b}$ è una soluzione del SL [in quanto $A(A^{-1}\mathbf{b}) = (AA^{-1})\mathbf{b} = \mathbf{b}$]. Se poi \mathbf{y} è un'altra soluzione del SL [cioè $A\mathbf{y} = \mathbf{b}$], allora $A\mathbf{y} = A(A^{-1}\mathbf{b})$. Cancellando la matrice A [che è invertibile], si ottiene $\mathbf{y} = A^{-1}\mathbf{b}$.

Per verificare la formula di Cramer calcoliamo $\det B_i$ usando lo sviluppo di Laplace rispetto alla i -sima colonna. Si ha:

$$\det B_i = b_1 \alpha_{1i} + \dots + b_n \alpha_{ni} = \sum_{t=1}^n b_t \alpha_{ti},$$

dove α_{ti} è il complemento algebrico dell'elemento di posto (t, i) della matrice A [si noti che A e B_i coincidono fuori della i -sima colonna].

Da un risultato sui determinanti, $A^{-1} = \frac{1}{\det A} {}^t \mathcal{C}_A$ e quindi

$$A^{-1}\mathbf{b} = \frac{1}{\det A} {}^t \mathcal{C}_A \mathbf{b}.$$

L' i -simo elemento di tale colonna è

$$\frac{1}{\det A} \left({}^t \mathcal{C}_A \right)^{(i)} \mathbf{b} = \frac{1}{\det A} {}^t [(\mathcal{C}_A)_{(i)}] \mathbf{b} = \frac{1}{\det A} (\alpha_{1i} \dots \alpha_{ni}) \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = \frac{1}{\det A} \det B_i.$$

Questo prova la formula.

Osservazione 1. È opportuno osservare che la nozione di determinante proviene storicamente dalla formula di Cramer [anche se noi abbiamo presentato questi due argomenti nell'ordine inverso]. Ad esempio, considerato un $SL(2, 2, K)$ $AX = \mathbf{b}$, con $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \in \mathbf{GL}_2(K)$ e $\mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$, la soluzione di tale SL è data dalla coppia (x_1, x_2) , con

$$x_1 = \frac{b_1 a_{22} - b_2 a_{12}}{a_{11} a_{22} - a_{12} a_{21}}, \quad x_2 = \frac{b_2 a_{11} - b_1 a_{21}}{a_{11} a_{22} - a_{12} a_{21}}.$$

Formule più complicate (ma analoghe) per SL quadrati di ordine 3, 4 ecc. hanno portato, per interpretare i numeratori e i denominatori delle precedenti frazioni, a definire la nozione di determinante di matrici quadrate.

Veniamo ora al teorema di Rouché-Capelli. Tale teorema fornisce un criterio per riconoscere se un SL è compatibile. Inoltre, con l'aiuto dell'algoritmo di Gauss, individua "l'infinità delle soluzioni" di un SL compatibile, cioè ci dice "quante" soluzioni ha il sistema.

Teorema 2. (*Teorema di Rouché-Capelli*) Sia $AX = \mathbf{b}$ un $SL(m, n, K)$ e sia $(A \ \mathbf{b})$ la sua matrice completa. Risulta:

$$\text{il } SL \ AX = \mathbf{b} \text{ è compatibile} \iff \text{rg}(A) = \text{rg}((A \ \mathbf{b})).$$

Se il SL è compatibile, ammette $\infty^{n-\text{rg}(A)}$ soluzioni.

Dim. Nel primo paragrafo di questo capitolo abbiamo osservato che una soluzione di un $SL \ AX = \mathbf{b}$ esprime la colonna \mathbf{b} come combinazione lineare delle colonne della matrice A . Pertanto:

$$\begin{aligned} \text{il } SL \ \text{è compatibile} &\iff \text{la colonna } \mathbf{b} \text{ è combinazione lineare delle colonne di } A \iff \\ &\iff \mathbf{b} \in \langle A_{(1)}, \dots, A_{(n)} \rangle \iff \langle A_{(1)}, \dots, A_{(n)}, \mathbf{b} \rangle = \langle A_{(1)}, \dots, A_{(n)} \rangle \iff \text{rg}((A \ \mathbf{b})) = \text{rg}(A). \end{aligned}$$

[Si noti che l'ultima implicazione \iff proviene dall'**Osserv. 3** di **Cap. 2.5**: due sottospazi vettoriali contenuti uno nell'altro ed aventi la stessa dimensione coincidono].

Supponiamo ora che il $SL \ AX = \mathbf{b}$ sia compatibile e poniamo $r := \text{rg}(A) = [\text{rg}((A \ \mathbf{b}))]$. Per semplificare le notazioni, possiamo senz'altro assumere che le prime r righe di A siano linearmente indipendenti [ciò può essere ottenuto con opportuni scambi di righe della matrice $(A \ \mathbf{b})$]. È evidente che anche le prime r righe di $(A \ \mathbf{b})$ sono linearmente indipendenti. Le rimanenti $m - r$ righe sono quindi loro combinazioni lineari. Dunque, per $s = r + 1, \dots, m$:

$$(A^{(s)} \ b_s) = \sum_{i=1}^r c_{si} (A^{(i)} \ b_i) = \left(\sum_{i=1}^r c_{si} A^{(i)} \ \sum_{i=1}^r c_{si} b_i \right).$$

Denotiamo con A^* la matrice formata dalle prime r righe di A e con \mathbf{b}^* i primi r elementi di \mathbf{b} . Otteniamo il $SL(r, n, K)$

$$A^* X = \mathbf{b}^*$$

e vogliamo verificare che tale SL è equivalente al $SL(m, n, K)$ assegnato. Denotati con Σ e Σ^* rispettivamente gli insiemi delle soluzioni di $AX = \mathbf{b}$ e di $A^* X = \mathbf{b}^*$, dobbiamo verificare che $\Sigma = \Sigma^*$. Ovviamente $\Sigma \subseteq \Sigma^*$. Viceversa, sia $\underline{z} \in \Sigma^*$ [e quindi $A^{(i)} \underline{z} = b_i$, per $i = 1, \dots, r$]. Allora, per $s = r + 1, \dots, m$:

$$A^{(s)} \underline{z} = \sum_{i=1}^r c_{si} A^{(i)} \underline{z} = \sum_{i=1}^r c_{si} b_i = b_s.$$

Dunque $\underline{z} \in \Sigma$ e pertanto $\Sigma = \Sigma^*$.

Possiamo quindi risolvere il $SL \ A^* X = \mathbf{b}^*$ in luogo del SL assegnato $AX = \mathbf{b}$. Applichiamo ad $A^* X = \mathbf{b}^*$ l'algoritmo di Gauss. Perverremo ad un SL a scala. Inoltre, poiché il rango si conserva con operazioni elementari riga e poiché il $SL \ A^* X = \mathbf{b}^*$ ha un numero di equazioni uguale al rango

della matrice, nel corso dell'algoritmo nessuna equazione può annullarsi identicamente. Si perviene perciò ad un SL a scala avente ancora r equazioni [ed n incognite]. Tale SL ammette, come noto, ∞^{n-r} soluzioni e dunque il SL $AX = \mathbf{b}$ ammette ∞^{n-r} soluzioni.

Come si risolve quindi un $SL(m, n, K)$ $AX = \mathbf{b}$? Dobbiamo innanzitutto verificare che risulta:

$$rg((A \ \mathbf{b})) = rg(A).$$

Verificato questo, se $rg(A) = r$, scegliamo in A r righe linearmente indipendenti e consideriamo le corrispondenti r righe di $(A \ \mathbf{b})$, eliminando invece le residue $m - r$ righe di $(A \ \mathbf{b})$.

Otteniamo così un $SL(r, n, K)$ $A^*X = \mathbf{b}^*$ equivalente al SL assegnato. La matrice A^* , avendo rango r , possiede r colonne linearmente indipendenti e sia B la matrice (invertibile) formata da tali colonne. In A^* ci sono $n - r$ colonne "esterne" a B . Attribuiamo valori parametrici indipendenti alle corrispondenti incognite e portiamo tali espressioni a secondo membro. Quel che si ottiene è un $SL(r, r, K)$ avente matrice dei coefficienti B (invertibile) e termini noti dipendenti da $n - r$ parametri. Con la formula di Cramer otterremo per tale sistema un'unica soluzione. Unica sì, ma dipendente da $n - r$ parametri. Si tratta quindi delle ∞^{n-r} soluzioni del sistema cercate.

La matrice B sopra considerata è detta *sottomatrice fondamentale* del SL . Essa non è in generale unica, ma dipende da varie scelte fatte [la scelta delle righe di A linearmente indipendenti e la scelta delle r colonne linearmente indipendenti di A^*]. È evidente che la scelta di B influenza il modo con cui vengono presentate le soluzioni del sistema, ma ovviamente non il complesso delle soluzioni stesse.

Un esempio. Vogliamo risolvere il $SL(3, 5, \mathbf{R})$:

$$\begin{cases} x_1 - x_2 - x_4 + 2x_5 = 0 \\ -x_1 + x_2 + x_3 + x_5 = -1 \\ 3x_1 - 3x_2 - x_3 - 2x_4 + 3x_5 = 1. \end{cases}$$

La matrice A dei coefficienti e la matrice completa M del SL sono rispettivamente

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

Fissiamo in A la sottomatrice $B = A(1, 2 | 3, 4) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ (avente determinante non nullo). I suoi tre orlati sono tutti nulli, come si può verificare. Dunque $rg(A) = 2$.

Anche la matrice M ha rango 2. Infatti, considerati i quattro orlati di $M(1, 2 | 3, 4)$, tre di essi coincidono con quelli di B in A , [e sono quindi già nulli], ed il quarto è $|M(1, 2, 3 | 3, 4, 6)|$ ed è anch'esso nullo, come si può verificare.

Dal teorema di Rouché-Capelli, il SL assegnato è compatibile ed ha $\infty^{5-2} = \infty^3$ soluzioni. Per ottenere tali soluzioni eliminiamo la terza equazione del SL ed assegniamo valori parametrici alle incognite x_1, x_2, x_5 [estranee alle colonne della sottomatrice fondamentale $B = A(1, 2 | 3, 4)$].

Posto $x_1 = t_1, x_2 = t_2, x_5 = t_3$, si ottiene il $SL(2, 2, \mathbf{R})$

$$\begin{cases} t_1 - t_2 - x_4 + 2t_3 = 0 \\ -t_1 + t_2 + x_3 + t_3 = -1, \end{cases} \quad \text{cioè} \quad \begin{cases} x_3 = t_1 - t_2 - t_3 - 1 \\ x_4 = t_1 - t_2 + 2t_3. \end{cases}$$

Le soluzioni del SL sono date quindi dall'insieme

$$\Sigma = \{(t_1, t_2, t_1 - t_2 - t_3 - 1, t_1 - t_2 + 2t_3, t_3), \quad \forall t_1, t_2, t_3 \in \mathbf{R}\}.$$

Ogni $SLO(m, n, K)$ $AX = \mathbf{0}$ è ovviamente sempre compatibile [in quanto ammette almeno la soluzione banale $\underline{0}$]. Il teorema di Rouché-Capelli conferma tale fatto [infatti $rg((A \ \mathbf{0})) = rg(A)$, $\forall A \in \mathfrak{M}_{m,n}(K)$]. Inoltre afferma che tale SLO ammette $\infty^{n-rg(A)}$ soluzioni. In particolare, il SLO è privo di autosoluzioni $\iff n = rg(A)$.

Un caso importante da esaminare è quello dei $SLO(n-1, n, K)$ $AX = \mathbf{0}$, con $rg(A) = n-1$, cioè con rango massimo possibile. Ci servono alcune notazioni.

La matrice A possiede esattamente n minori di ordine $n-1$, non tutti nulli. Ciascuno di essi è ottenuto eliminando da A una colonna. Denoteremo con \mathcal{A}_i il minore ottenuto da A eliminando

la i -sima colonna di A . Attribuiamo ai minori $\mathcal{A}_1, \dots, \mathcal{A}_n$ alternativamente segno + e segno - ed otteniamo così la n -pla (non nulla):

$$\underline{z} := (\mathcal{A}_1, -\mathcal{A}_2, \mathcal{A}_3, \dots, (-1)^{n+1}\mathcal{A}_n).$$

Proposizione 1. Sia $AX = \mathbf{0}$ un $SLO(n-1, n, K)$ con $rg(A) = n-1$. Tale SLO ammette ∞^1 soluzioni, proporzionali all'autosoluzione \underline{z} sopra definita.

Dim. In base al teorema di Rouché-Capelli, il SLO assegnato ha $\infty^{n-(n-1)} = \infty^1$ soluzioni. È noto inoltre che le soluzioni di un SLO formano uno spazio vettoriale. In questo caso si tratta di un sottospazio vettoriale 1-dimensionale di K^n , che denoteremo, al solito, Σ_0 .

Basterà quindi verificare che la n -pla \underline{z} sopra definita è una soluzione di $AX = \mathbf{0}$, cioè che risulta:

$$A^{(i)} \mathbf{z} = 0, \quad \forall i = 1, \dots, n-1,$$

[dove \mathbf{z} è la colonna corrispondente alla n -pla \underline{z}], per concludere che $\Sigma_0 = \langle \underline{z} \rangle$.

Per $i = 1, \dots, n-1$, consideriamo la matrice

$$C_i := \begin{pmatrix} A^{(i)} \\ A \end{pmatrix},$$

ottenuta premettendo la riga $A^{(i)}$ alla matrice A . Si tratta di una matrice quadrata di ordine n con due righe uguali e quindi con determinante nullo. Sviluppiamo $\det(C_i)$ rispetto alla prima riga. Osserviamo che

$$\alpha_{11} = (-1)^{1+1}\mathcal{A}_1 = \mathcal{A}_1, \quad \alpha_{12} = (-1)^{1+2}\mathcal{A}_2 = -\mathcal{A}_2, \dots, \alpha_{1n} = (-1)^{1+n}\mathcal{A}_n.$$

Ne segue che:

$$0 = \det(C_i) = a_{i1}\mathcal{A}_1 + a_{i2}(-\mathcal{A}_2) + a_{i3}\mathcal{A}_3 + \dots + a_{in}((-1)^{n+1}\mathcal{A}_n) = A^{(i)} \mathbf{z},$$

come richiesto.

Due esempi. (1) Risolviamo, al variare di $a \in \mathbf{R}$, il $SLO(3, 4, \mathbf{R})$:

$$\begin{cases} ax_2 + 2x_3 - x_4 = 0 \\ x_1 + ax_3 = 0 \\ ax_2 + ax_3 + x_4 = 0. \end{cases}$$

Tale SLO ha matrice dei coefficienti

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

Senza prima discutere il rango di A , calcoliamo la 4-pla $\underline{z} = (\mathcal{A}_1, -\mathcal{A}_2, \mathcal{A}_3, -\mathcal{A}_4)$. Si ha:

$$\begin{aligned} \mathcal{A}_1 &= \begin{vmatrix} a & 2 & -1 \\ 0 & a & 0 \\ a & a & 1 \end{vmatrix} = 2a^2, \quad \mathcal{A}_2 = \begin{vmatrix} 0 & 2 & -1 \\ 1 & a & 0 \\ 0 & a & 1 \end{vmatrix} = -a - 2, \\ \mathcal{A}_3 &= \begin{vmatrix} 0 & a & -1 \\ 1 & 0 & 0 \\ 0 & a & 1 \end{vmatrix} = -2a, \quad \mathcal{A}_4 = \begin{vmatrix} 0 & a & 2 \\ 1 & 0 & a \\ 0 & a & a \end{vmatrix} = 2a - a^2. \end{aligned}$$

Ne segue:

$$\underline{z} = (2a^2, a + 2, -2a, a^2 - 2a).$$

Si osserva subito che, per ogni $a \in \mathbf{R}$, risulta: $\underline{z} \neq \underline{0}$ [infatti ad esempio la seconda e terza componente di \underline{z} non si annullano simultaneamente]. Pertanto $rg(A) = 3, \forall a \in \mathbf{R}$.

Si conclude che il SLO assegnato ha soluzioni descritte dal sottospazio vettoriale di \mathbf{R}^4 :

$$\Sigma_0 = \langle (2a^2, a + 2, -2a, a^2 - 2a) \rangle.$$

(2) Risolviamo il $SLO(3, 4, \mathbf{Z}_2)$:

$$\begin{cases} x_1 + x_2 + x_3 = 0 \\ x_2 + x_3 + x_4 = 0 \\ x_1 + x_2 + x_4 = 0. \end{cases}$$

Tale SLO ha matrice

$$A = \begin{pmatrix} \bar{1} & \bar{1} & \bar{1} & \bar{0} \\ \bar{0} & \bar{1} & \bar{1} & \bar{1} \\ \bar{1} & \bar{1} & \bar{0} & \bar{1} \end{pmatrix} \in \mathfrak{M}_{3,4}(\mathbf{Z}_2).$$

Tale matrice ha rango 3. Infatti la sottomatrice formata dalle prime tre colonne di A ha determinante

$$\begin{vmatrix} \bar{1} & \bar{1} & \bar{1} \\ \bar{0} & \bar{1} & \bar{1} \\ \bar{1} & \bar{1} & \bar{0} \end{vmatrix} = \bar{1} + \bar{1} + \bar{1} = \bar{1}.$$

In base alla **Prop. 1**, le soluzioni del SLO sono proporzionali all'autosoluzione

$$\left(\begin{vmatrix} \bar{1} & \bar{1} & \bar{1} \\ \bar{0} & \bar{1} & \bar{1} \\ \bar{1} & \bar{1} & \bar{0} \end{vmatrix}, - \begin{vmatrix} \bar{1} & \bar{1} & \bar{0} \\ \bar{0} & \bar{1} & \bar{1} \\ \bar{1} & \bar{0} & \bar{1} \end{vmatrix}, \begin{vmatrix} \bar{1} & \bar{1} & \bar{0} \\ \bar{0} & \bar{1} & \bar{1} \\ \bar{1} & \bar{1} & \bar{1} \end{vmatrix}, - \begin{vmatrix} \bar{1} & \bar{1} & \bar{0} \\ \bar{1} & \bar{1} & \bar{1} \\ \bar{1} & \bar{0} & \bar{1} \end{vmatrix} \right) = (\bar{1}, \bar{0}, \bar{1}, \bar{1}).$$

Pertanto lo spazio vettoriale delle soluzioni del SLO è

$$\Sigma_0 = \langle (\bar{1}, \bar{0}, \bar{1}, \bar{1}) \rangle = \{(\bar{0}, \bar{0}, \bar{0}, \bar{0}), (\bar{1}, \bar{0}, \bar{1}, \bar{1})\}.$$

[Si noti che le soluzioni sono 2. Infatti $|\Sigma_0| = |\mathbf{Z}_2|^1 = 2$].

ESERCIZI PROPOSTI

3.4.1 Risolvere il seguente $SLO(4, 3, \mathbf{R})$:

$$\begin{cases} x + y + z = 0 \\ -y + z = 0 \\ -x - 2y = 0 \\ 2x + 3y + z = 0. \end{cases}$$

3.4.2 Risolvere, al variare di $a, b \in \mathbf{R}$, il seguente $SL(3, 3, \mathbf{R})$:

$$\begin{cases} ax + by = 0 \\ y + bz = a \\ x - az = b. \end{cases}$$

3.4.3. Utilizzando il teorema di Rouché-Capelli discutere la risoluzione, al variare di tre parametri $a, b, c \in \mathbf{R}$, del $SL(1, 2, \mathbf{R})$: $\{ax + by = c\}$.

3.4.4. Rifare l'**Esercizio 3.1.3** (senza usare l'algoritmo di Gauss).

3.4.5. Rifare l'**Esercizio 3.1.4** (senza usare l'algoritmo di Gauss).

3.4.6 È assegnato il $SL(3, 3, \mathbf{R})$:

$$\begin{cases} ay + bz = c \\ -ax + az = -c \\ -bx - ay = 0, \end{cases}$$

dipendente da tre parametri $a, b, c \in \mathbf{R}$, con $a, b \neq 0$. Determinare per quali valori dei parametri il SL è compatibile e scriverne l'insieme Σ delle soluzioni.

3.4.7 È assegnato il $SLO(2, 3, \mathbf{R})$:

$$\begin{cases} bx + ay = 0 \\ y + az = 0, \end{cases}$$

dipendente da due parametri $a, b \in \mathbf{R}$. Risolvere tale SLO , al variare dei parametri.

3.4.8. Risolvere il seguente $SL(3, 5, \mathbf{Z}_3)$:

$$\begin{cases} x_1 + x_3 + x_5 = \bar{2} \\ x_2 + x_3 = \bar{1} \\ x_1 + x_4 + x_5 = \bar{0}. \end{cases}$$

3.4.9. È assegnato il seguente $SL(3, 3, \mathbf{Z}_3)$, dipendente da un parametro $a \in \mathbf{Z}_3$:

$$\begin{cases} x + \bar{2}z = a \\ \bar{2}x + \bar{2}y + z = a \\ \bar{2}x + z = \bar{2} + a. \end{cases}$$

Determinare per quali eventuali $a \in \mathbf{Z}_3$ il SL è compatibile e calcolarne le soluzioni.

3.4.10 È assegnato il $SL(3, 3, \mathbf{R})$:

$$\begin{cases} ax - y + bz = -a \\ x - bz = 0 \\ ax + 2y = b, \end{cases}$$

dipendente da due parametri $a, b \in \mathbf{R}$. Risolvere tale SL , al variare dei parametri.

3.4.11. Al variare dei parametri $a, b \in \mathbf{R}$, risolvere il seguente $SL(2, 2, \mathbf{R})$

$$\begin{cases} ax + by = a - b \\ bx + ay = b - a. \end{cases}$$

Capitolo 4

APPLICAZIONI LINEARI

1. Notazioni "matriciali" negli spazi vettoriali

Introdurremo in questo paragrafo delle notazioni che ci verranno utili nel seguito del capitolo. Con queste notazioni potremo anche interpretare le nozioni di indipendenza lineare e di sistema di generatori di uno spazio vettoriale (già studiate nei capitoli precedenti) in termini di semplici condizioni relative al rango di matrici.

Il prodotto righe per colonne tra matrici si presta in modo naturale a scrivere le combinazioni lineari di vettori in un K -spazio vettoriale V . Siano infatti $\underline{u}_1, \dots, \underline{u}_t \in V$ e $c_1, \dots, c_t \in K$. Poniamo

$$\mathbf{U} := (\underline{u}_1 \ \dots \ \underline{u}_t) \in \mathfrak{M}_{1,t}(V), \quad \mathbf{c} := \begin{pmatrix} c_1 \\ \vdots \\ c_t \end{pmatrix} \in \mathfrak{M}_{t,1}(K).$$

Vogliamo definire il prodotto righe per colonne tra la matrice-riga \mathbf{U} di vettori e la matrice-colonna \mathbf{c} di scalari. A tale scopo interpreteremo la moltiplicazione tra il vettore \underline{u}_i ed il corrispondente scalare c_i come *moltiplicazione per uno scalare*, cioè $\underline{u}_i c_i$. Ponremo quindi:

$$\mathbf{U}\mathbf{c} := \underline{u}_1 c_1 + \dots + \underline{u}_t c_t \in V.$$

Dunque $\mathbf{U}\mathbf{c}$ è la "notazione matriciale" che ci permette di scrivere compattamente il vettore combinazione lineare dei vettori $\underline{u}_1, \dots, \underline{u}_t$, con coefficienti c_1, \dots, c_t .

Estendiamo tale notazione al caso del prodotto di una riga di vettori per una matrice (a più colonne) di scalari. Considerata la matrice $C \in \mathfrak{M}_{t,s}(K)$, poniamo:

$$\mathbf{U}C := (\mathbf{U}C_{(1)} \ \dots \ \mathbf{U}C_{(s)}) \in \mathfrak{M}_{1,s}(V).$$

Ad esempio, se

$$\mathbf{U} = (\underline{u}_1 \ \underline{u}_2 \ \underline{u}_3) \in \mathfrak{M}_{1,3}(V) \quad \text{e} \quad C = \begin{pmatrix} 1 & 0 & -1 & 1 \\ -1 & 1 & 0 & -2 \\ 0 & 1 & 2 & 0 \end{pmatrix} \in \mathfrak{M}_{3,4}(K),$$

allora

$$\mathbf{U}C = (\underline{u}_1 - \underline{u}_2 \ \underline{u}_2 + \underline{u}_3 \ -\underline{u}_1 + 2\underline{u}_3 \ \underline{u}_1 - 2\underline{u}_2) \in \mathfrak{M}_{1,4}(V).$$

Oltre alle matrici $\mathbf{U} \in \mathfrak{M}_{1,t}(V)$, $C \in \mathfrak{M}_{t,s}(K)$, sia assegnata una matrice $D \in \mathfrak{M}_{s,r}(K)$. Poiché le "misure" delle matrici lo consentono, sono definiti i prodotti righe per colonne $(\mathbf{U}C)D$ e $\mathbf{U}(CD)$. Si verifica subito che risulta:

$$(\mathbf{U}C)D = \mathbf{U}(CD).$$

Ciò segue facilmente dall'assioma [di spazio vettoriale]: $(\underline{u}c)d = \underline{u}(cd)$, $\forall \underline{u} \in V$, $\forall c, d \in K$. Utilizzeremo tale proprietà, indicandola come *proprietà associativa* del prodotto righe per colonne.

Se $V = V_K^n$ e se $\{\underline{e}_1, \dots, \underline{e}_n\}$ è una base di V , è noto che il generico vettore $\underline{v} \in V$ si scrive in modo unico come combinazione lineare di $\underline{e}_1, \dots, \underline{e}_n$, con coefficienti $x_1, \dots, x_n \in K$. Come già sappiamo, la n -pla (x_1, \dots, x_n) è detta n -pla delle coordinate di \underline{v} rispetto a tale base. Se poniamo

$$\mathbf{E} := (\underline{e}_1 \ \dots \ \underline{e}_n) \in \mathfrak{M}_{1,n}(V), \quad \mathbf{x} := \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathfrak{M}_{n,1}(K),$$

risulta subito:

$$\underline{v} = \mathbf{E} \mathbf{x}$$

e la matrice-colonna \mathbf{x} è detta *colonna delle coordinate di \underline{v} in base \mathbf{E}* . In particolare, l'unicità di scrittura di un vettore in base \mathbf{E} si traduce nell'implicazione:

$$(*) \quad \mathbf{E} \mathbf{x} = \mathbf{E} \mathbf{y} \implies \mathbf{x} = \mathbf{y}, \quad \forall \mathbf{x}, \mathbf{y} \in \mathfrak{M}_{n,1}(K).$$

Assegnati t vettori $\underline{u}_1, \underline{u}_2, \dots, \underline{u}_t \in V$, sia $\mathbf{U} := (\underline{u}_1 \ \underline{u}_2 \ \dots \ \underline{u}_t) \in \mathfrak{M}_{1,t}(V)$ la matrice riga di tali vettori. Supponiamo note le coordinate di tali vettori in base \mathbf{E} . Dunque, per $i = 1, \dots, t$, sia $\underline{u}_i = \mathbf{E} \mathbf{b}_i$, con $\mathbf{b}_i \in \mathfrak{M}_{n,1}(K)$, e si ponga:

$$B := (\mathbf{b}_1 \ \mathbf{b}_2 \ \dots \ \mathbf{b}_t) \in \mathfrak{M}_{n,t}(K).$$

La matrice B è detta *matrice delle coordinate dei vettori $\underline{u}_1, \underline{u}_2, \dots, \underline{u}_t$ in base \mathbf{E}* . Risulta:

$$\mathbf{U} = \mathbf{E} B$$

$$[\text{infatti } \mathbf{E} B = \mathbf{E} (\mathbf{b}_1 \ \mathbf{b}_2 \ \dots \ \mathbf{b}_t) = (\mathbf{E} \mathbf{b}_1 \ \mathbf{E} \mathbf{b}_2 \ \dots \ \mathbf{E} \mathbf{b}_t) = (\underline{u}_1 \ \underline{u}_2 \ \dots \ \underline{u}_t) = \mathbf{U}].$$

Ad esempio, fissata in $V = V_{\mathbf{R}}^4$ una base $\mathbf{E} = (\underline{e}_1 \ \underline{e}_2 \ \underline{e}_3 \ \underline{e}_4)$ ed assegnati i tre vettori

$$\underline{u}_1 = \underline{e}_1 - \underline{e}_3 + \underline{e}_4, \quad \underline{u}_2 = 2\underline{e}_1 + \underline{e}_3, \quad \underline{u}_3 = \underline{e}_1 + \underline{e}_2,$$

la matrice delle coordinate dei tre vettori in base \mathbf{E} è

$$B = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 0 & 1 \\ -1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \in \mathfrak{M}_{4,3}(\mathbf{R}) \quad \text{e si scriverà} \quad (\underline{u}_1 \ \underline{u}_2 \ \underline{u}_3) = \mathbf{E} B.$$

Proveremo ora che, assegnati t vettori in uno spazio vettoriale n -dimensionale, le proprietà di essere linearmente indipendenti o di essere un sistema di generatori, si interpretano come condizioni sul rango della matrice delle loro coordinate (rispetto ad una base assegnata).

Proposizione 1. Sia $\dim(V) = n$ e sia $\mathbf{E} = (\underline{e}_1 \ \dots \ \underline{e}_n)$ una base di V . Siano $\underline{u}_1, \dots, \underline{u}_t \in V$ e si ponga $\mathbf{U} := (\underline{u}_1 \ \dots \ \underline{u}_t)$. Sia poi $\mathbf{U} = \mathbf{E} B$, con $B \in \mathfrak{M}_{n,t}(K)$. Risulta:

- (i) $\underline{u}_1, \dots, \underline{u}_t$ sono linearmente indipendenti $\iff \text{rg}(B) = t$.
- (ii) $\{\underline{u}_1, \dots, \underline{u}_t\}$ è un sistema di generatori di V $\iff \text{rg}(B) = n$.
- (iii) $\{\underline{u}_1, \dots, \underline{u}_t\}$ è una base di V $\iff t = n$ e $B \in \mathbf{GL}_n(K)$.

Dim. (i) Come noto:

$$\underline{u}_1, \dots, \underline{u}_t \text{ sono linearmente indipendenti} \iff \left[\sum_{i=1}^t c_i \underline{u}_i = \underline{0} \implies c_1 = \dots = c_t = 0 \right].$$

Siano $\mathbf{c} = \begin{pmatrix} c_1 \\ \vdots \\ c_t \end{pmatrix} \in \mathfrak{M}_{t,1}(K)$, $\mathbf{0}_t = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \in \mathfrak{M}_{t,1}(K)$ e $\mathbf{0} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \in \mathfrak{M}_{n,1}(K)$ [matrici colonne nulle]. Ovviamente $\underline{0} = \mathbf{E} \mathbf{0}$. La precedente equivalenza diventa [tenuto conto dell'associatività del prodotto righe per colonne e della regola (*) relativa all'unicità di scrittura in base \mathbf{E}]:

$$\begin{aligned} \underline{u}_1, \dots, \underline{u}_t \text{ sono linearmente indipendenti} &\iff [\mathbf{U} \mathbf{c} = \underline{0} \implies \mathbf{c} = \mathbf{0}_t] \\ &\iff [(\mathbf{E} B) \mathbf{c} = \mathbf{E} \mathbf{0} \implies \mathbf{c} = \mathbf{0}_t] \\ &\iff [\mathbf{E} (B \mathbf{c}) = \mathbf{E} \mathbf{0} \implies \mathbf{c} = \mathbf{0}_t] \\ &\iff [B \mathbf{c} = \mathbf{0} \implies \mathbf{c} = \mathbf{0}_t]. \end{aligned}$$

L'implicazione $B \mathbf{c} = \mathbf{0} \implies \mathbf{c} = \mathbf{0}_t$ equivale a dire che il $SLO(n, t, K)$ $BX = \mathbf{0}$ non ha autosoluzioni. In base al teorema di Rouché-Capelli, ciò avviene $\iff t - \text{rg}(B) = 0$. Pertanto concludiamo che

$$\underline{u}_1, \dots, \underline{u}_t \text{ sono linearmente indipendenti} \iff \text{rg}(B) = t.$$

(ii) Come noto:

$$\langle \underline{u}_1, \dots, \underline{u}_t \rangle = V \iff \forall \underline{v} \in V, \underline{v} = c_1 \underline{u}_1 + \dots + c_t \underline{u}_t, \exists c_1, \dots, c_t \in K.$$

Assumiamo che il generico vettore $\underline{v} \in V$ abbia colonna delle coordinate \mathbf{x} , cioè che sia $\underline{v} = \mathbf{E}\mathbf{x}$, con $\mathbf{x} \in \mathfrak{M}_{n,1}(K)$. Con le notazioni già introdotte in (i), si ha:

$$\begin{aligned} \langle \underline{u}_1, \dots, \underline{u}_t \rangle = V &\iff \forall \underline{v} \in V, \underline{v} = \mathbf{U}\mathbf{c}, \exists \mathbf{c} \in \mathfrak{M}_{t,1}(K) \\ &\iff \forall \mathbf{E}\mathbf{x} \in V, \mathbf{E}\mathbf{x} = (\mathbf{E}\mathbf{B})\mathbf{c}, \exists \mathbf{c} \in \mathfrak{M}_{t,1}(K) \\ &\iff \forall \mathbf{E}\mathbf{x} \in V, \mathbf{E}\mathbf{x} = \mathbf{E}(\mathbf{B}\mathbf{c}), \exists \mathbf{c} \in \mathfrak{M}_{t,1}(K) \\ &\iff \forall \mathbf{x} \in \mathfrak{M}_{n,1}(K), \mathbf{x} = \mathbf{B}\mathbf{c}, \exists \mathbf{c} \in \mathfrak{M}_{t,1}(K) \\ &\iff \forall \mathbf{x} \in \mathfrak{M}_{n,1}(K), \text{ il } SL(n, t, K) \text{ } BX = \mathbf{x} \text{ è compatibile.} \end{aligned}$$

Ricordato che un SL è compatibile \iff la colonna dei termini noti è combinazione lineare delle colonne della matrice dei coefficienti, si ha:

$$\begin{aligned} \langle \underline{u}_1, \dots, \underline{u}_t \rangle = V &\iff \forall \mathbf{x} \in \mathfrak{M}_{n,1}(K), \mathbf{x} \in \langle B_{(1)}, \dots, B_{(t)} \rangle \\ &\iff \langle B_{(1)}, \dots, B_{(t)} \rangle = \mathfrak{M}_{n,1}(K) \\ &\iff \dim(\langle B_{(1)}, \dots, B_{(t)} \rangle) = \dim(\mathfrak{M}_{n,1}(K)) \\ &\iff rg(B) = n. \end{aligned}$$

(iii) Si tratta di un'immmediata conseguenza delle due proposizioni precedenti, tenendo conto anche del fatto che una matrice quadrata di ordine n è invertibile se e solo se ha rango n .

Ad esempio i tre vettori considerati nel precedente esempio sono linearmente indipendenti, in quanto, come subito si osserva, $rg(B) = 3$. Non sono invece un sistema di generatori di V (perché sono solo tre e quindi $rg(B) < 4$).

Si noti che, se $\mathbf{U} = \mathbf{E}\mathbf{B}$ e $rg(B) = r$, una base di $\langle \underline{u}_1, \dots, \underline{u}_t \rangle$ è formata da r vettori del sistema di generatori $\{\underline{u}_1, \dots, \underline{u}_t\}$, corrispondenti a r colonne linearmente indipendenti di B .

È noto dalla **Prop. 1** di **Cap. 3.1** che le soluzioni di un SLO in n incognite formano un sottospazio vettoriale di K^n . Proveremo ora che, viceversa, assegnata una base \mathbf{E} di un K -spazio vettoriale $V = V_K^n$, ogni sottospazio vettoriale di V può essere espresso come insieme delle soluzioni di un opportuno SLO in n incognite, che fornisce le cosiddette *equazioni cartesiane di U (in base \mathbf{E})*.

Proposizione 2. Sia $U = \langle \underline{u}_1, \dots, \underline{u}_t \rangle$ un sottospazio vettoriale di $V = V_K^n$ e sia $\mathbf{E} = (\underline{e}_1 \dots \underline{e}_n)$ una base di V . Sia, al solito, $\mathbf{U} := (\underline{u}_1 \dots \underline{u}_t) = \mathbf{E}\mathbf{B}$, con $B \in \mathfrak{M}_{n,t}(K)$. Indicato con $\underline{v} = \mathbf{E}\mathbf{x}$ un generico vettore di V , risulta:

$$\underline{v} \in U \iff rg((B \mathbf{x})) = rg(B).$$

Posto $r = rg(B)$, l'uguaglianza $rg((B \mathbf{x})) = r$ si traduce in un $SLO(n - r, n, K)$, avente matrice di rango $n - r$, le cui soluzioni sono le coordinate (in base \mathbf{E}) di tutti e soli i vettori di U . Tale SLO è detto *sistema di equazioni cartesiane di U in base \mathbf{E}* .

Dim. Si ha infatti, in base al teorema di Rouché-Capelli:

$$\begin{aligned} \underline{v} \in U &\iff \underline{v} = \mathbf{U}\mathbf{c}, \exists \mathbf{c} \in \mathfrak{M}_{t,1}(K) \\ &\iff \mathbf{E}\mathbf{x} = \mathbf{E}\mathbf{B}\mathbf{c}, \exists \mathbf{c} \in \mathfrak{M}_{t,1}(K) \\ &\iff B\mathbf{c} = \mathbf{x}, \exists \mathbf{c} \in \mathfrak{M}_{t,1}(K) \\ &\iff rg(B) = rg((B \mathbf{x})). \end{aligned}$$

Fissiamo ora in B una sottomatrice quadrata invertibile B_1 di ordine $r (= rg(B))$. In base al principio degli orlati, la condizione $rg((B \mathbf{x})) = r$ equivale all'annullarsi degli orlati di B_1 . Ma possiamo limitarci a considerare i soli orlati di B_1 che coinvolgono la colonna \mathbf{x} [in quanto gli altri (se ce ne sono) sono nulli]. Questi orlati sono $n - r$ e possono essere interpretati come polinomi di primo grado nelle incognite x_1, \dots, x_n di \mathbf{x} . Dunque formano un $SLO(n - r, n, K)$, che possiamo ad esempio denotare $CX = \mathbf{0}$. Tale SLO ammette come soluzioni tutte e sole le n -ple che corrispondono (in base \mathbf{E}) ai vettori di U . Poiché (come sopra osservato) $\dim(U) = r$, il SLO ha ∞^r soluzioni. D'altra parte ha $\infty^{n-rg(C)}$ soluzioni. Quindi $n - rg(C) = r$, cioè $rg(C) = n - r$.

N.B. Ogni $SLO(m, n, K)$ $AX = \mathbf{0}$ equivalente a tale $SLO(n - r, n, K)$ fornisce ancora un sistema di equazioni cartesiane di U . Dunque le equazioni cartesiane di un sottospazio vettoriale non sono

uniche. Deve comunque risultare $m \geq n - r$ [in quanto $n - rg(A) = r$ e $rg(A) \leq m$] e pertanto il $SLO(n - r, n, K)$ ottenuto ha il numero minimo di equazioni necessarie per rappresentare U .

Due esempi. (1) Torniamo ancora all'esempio precedente. Il sottospazio $U = \langle \underline{u}_1, \underline{u}_2, \underline{u}_3 \rangle$ di $V = V_{\mathbf{R}}^4$ ha equazioni ottenute imponendo la condizione $rg((B \ x)) = 3$, ovvero $\det((B \ x)) = 0$ [infatti la matrice $(B \ x) \in \mathfrak{M}_4(\mathbf{R})$ ha un solo minore di ordine 4: $\det((B \ x))$]. Si ha:

$$\begin{vmatrix} 1 & 2 & 1 & x_1 \\ 0 & 0 & 1 & x_2 \\ -1 & 1 & 0 & x_3 \\ 1 & 0 & 0 & x_4 \end{vmatrix} = - \begin{vmatrix} 2 & 1 & x_1 \\ 0 & 1 & x_2 \\ 1 & 0 & x_3 \end{vmatrix} + x_4 \begin{vmatrix} 1 & 2 & 1 \\ 0 & 0 & 1 \\ -1 & 1 & 0 \end{vmatrix} = x_1 - x_2 - 2x_3 - 3x_4 = 0.$$

Dunque U ha equazione cartesiana $x_1 - x_2 - 2x_3 - 3x_4 = 0$.

(2) In $V = V_{\mathbf{R}}^5$ è assegnato il sottospazio vettoriale $U = \langle \underline{u}_1, \underline{u}_2, \underline{u}_3 \rangle$, i cui tre generatori hanno, rispetto ad una base $\mathbf{E} = (\underline{e}_1 \ \underline{e}_2 \ \underline{e}_3 \ \underline{e}_4 \ \underline{e}_5)$ di V , matrice delle coordinate

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

Vogliamo determinare un sistema di equazioni cartesiane di U rispetto ad \mathbf{E} .

Calcoliamo $rg(B)$. Scelta in B la sottomatrice invertibile $B_1 = B(1, 2|1, 2)$, si verifica subito che i suoi tre orlati sono nulli. Dunque $rg(B) = 2$. Ne segue che $\dim(U) = 2$ e $U = \langle \underline{u}_1, \underline{u}_2 \rangle$.

Imponiamo ora la condizione $rg((B \ x)) = 2$. Tale condizione equivale all'annullamento degli orlati di B_1 che coinvolgono la colonna \mathbf{x} , cioè

$$\begin{vmatrix} 1 & 2 & x_1 \\ 0 & -1 & x_2 \\ -1 & 1 & x_3 \end{vmatrix} = 0, \quad \begin{vmatrix} 1 & 2 & x_1 \\ 0 & -1 & x_2 \\ 2 & 1 & x_4 \end{vmatrix} = 0, \quad \begin{vmatrix} 1 & 2 & x_1 \\ 0 & -1 & x_2 \\ 0 & 1 & x_5 \end{vmatrix} = 0.$$

Si ottiene quindi il $SLO(3, 5, \mathbf{R})$:

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

che fornisce equazioni cartesiane di U .

Vogliamo ora scrivere le *formule di cambiamento di base* e *di coordinate di vettore* in uno spazio vettoriale $V = V_K^n$. Siano \mathbf{E}, \mathbf{F} due basi di V . In base a Prop. 1 (iii) risulta:

$$\mathbf{F} = \mathbf{E} B, \text{ con } B \in \mathbf{GL}_n(K).$$

Tale formula è detta *formula del cambiamento di base da \mathbf{E} a \mathbf{F}* e la matrice B è detta *matrice del cambiamento di base da \mathbf{E} a \mathbf{F}* . Risulta:

$$\mathbf{F} B^{-1} = (\mathbf{E} B) B^{-1} = \mathbf{E} (B B^{-1}) = \mathbf{E} I_n = \mathbf{E}.$$

Abbiamo così ottenuto

$$\mathbf{E} = \mathbf{F} B^{-1},$$

detta *formula del cambiamento di base da \mathbf{F} a \mathbf{E}* [ed è l'inversa della formula precedente].

Sia ora \underline{v} un generico vettore di V . Esprimiamolo nelle due basi \mathbf{E}, \mathbf{F} :

$$\underline{v} = \mathbf{E} \mathbf{x} = \mathbf{F} \mathbf{y}, \text{ con } \mathbf{x}, \mathbf{y} \in \mathfrak{M}_{n,1}(K).$$

Usando la prima delle due formule del cambiamento di base, si ottiene $\mathbf{E} \mathbf{x} = \mathbf{E} B \mathbf{y}$ e quindi, per unicità di scrittura,

$$\mathbf{x} = B \mathbf{y},$$

detta *formula del cambiamento di coordinate di vettore da \mathbf{F} a \mathbf{E}* . Moltiplicando questa uguaglianza a sinistra per B^{-1} si ottiene

$$\mathbf{y} = B^{-1} \mathbf{x},$$

detta *formula del cambiamento di coordinate di vettore da \mathbf{E} a \mathbf{F}* .

N.B. Si osservi che nelle due formule di cambiamento di base, la matrice invertibile B (o la sua inversa B^{-1}) è un fattore a destra, mentre nelle due formule di cambiamento di coordinate è un fattore a sinistra. Ciò non è privo di conseguenze: nelle formule di cambiamento di coordinate di vettore intervengono le righe di B (o B^{-1}) [come coefficienti di combinazioni lineari], mentre nelle formule di cambiamento di base intervengono le colonne di B (o B^{-1}).

Ciò può essere osservato confrontando ad esempio le due formule $\mathbf{F} = \mathbf{E}B$ e $\mathbf{x} = B\mathbf{y}$, scritte in forma non compatta:

$$\begin{cases} \underline{f}_1 = \underline{e}_1 b_{11} + \underline{e}_2 b_{21} + \dots + \underline{e}_n b_{n1} \\ \dots \\ \underline{f}_n = \underline{e}_1 b_{1n} + \underline{e}_2 b_{2n} + \dots + \underline{e}_n b_{nn}, \end{cases} \quad \begin{cases} x_1 = b_{11} y_1 + b_{12} y_2 + \dots + b_{1n} y_n \\ \dots \\ x_n = b_{n1} y_1 + b_{n2} y_2 + \dots + b_{nn} y_n. \end{cases}$$

Un esempio. Consideriamo in $V = V_{\mathbf{R}}^2$ due basi $\mathbf{E} = (\underline{e}_1 \ \underline{e}_2)$, $\mathbf{F} = (\underline{f}_1 \ \underline{f}_2)$, tali che

$$\underline{f}_1 = \underline{e}_1 + \underline{e}_2, \quad \underline{f}_2 = -\underline{e}_1 - 2\underline{e}_2.$$

Vogliamo ottenere la formula del cambiamento di base da \mathbf{F} a \mathbf{E} e le due formule del cambiamento di coordinate.

I dati sopra forniti ci danno il cambiamento di base da \mathbf{E} a \mathbf{F} e dunque corrispondono alla formula $\mathbf{F} = \mathbf{E}B$. Risulta:

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

Il cambiamento di base da \mathbf{F} a \mathbf{E} è dato da $\mathbf{E} = \mathbf{F}B^{-1}$. Poiché $B^{-1} = \begin{pmatrix} 2 & -1 \\ 1 & -1 \end{pmatrix}$, si ha:

$$\begin{cases} \underline{e}_1 = 2\underline{f}_1 + \underline{f}_2 \\ \underline{e}_2 = -\underline{f}_1 - \underline{f}_2. \end{cases}$$

N.B. Tale formula poteva naturalmente essere ottenuta *direttamente* dai dati sopra assegnati. Infatti: $\underline{f}_1 + \underline{f}_2 = -\underline{e}_2$, quindi $\underline{e}_2 = -\underline{f}_1 - \underline{f}_2$; allora $\underline{e}_1 = \underline{f}_1 - \underline{e}_2 = 2\underline{f}_1 + \underline{f}_2$.

È evidente però che se si opera su uno spazio vettoriale di dimensione elevata, queste trasformazioni sono più complicate e quindi è preferibile svolgere i calcoli utilizzando le matrici.

Il cambiamento di coordinate da \mathbf{F} a \mathbf{E} è dato da $\mathbf{x} = B\mathbf{y}$, cioè

$$\begin{cases} x_1 = y_1 - y_2 \\ x_2 = y_1 - 2y_2, \end{cases}$$

mentre il cambiamento di coordinate da \mathbf{E} a \mathbf{F} è dato da $\mathbf{y} = B^{-1}\mathbf{x}$, cioè

$$\begin{cases} y_1 = 2x_1 - x_2 \\ y_2 = x_1 - x_2. \end{cases}$$

ESERCIZI PROPOSTI

4.1.1. Sia $V = V_{\mathbf{R}}^3$, con base $\mathbf{E} = (\underline{e}_1 \ \underline{e}_2 \ \underline{e}_3)$. Sono assegnati in V i tre vettori

$$\underline{u}_1 = \underline{e}_1 - \underline{e}_2, \quad \underline{u}_2 = \underline{e}_2 - \underline{e}_3, \quad \underline{u}_3 = \underline{e}_3 + \underline{e}_1.$$

(i) Verificare che $\mathbf{F} = (\underline{u}_1 \ \underline{u}_2 \ \underline{u}_3)$ è una base di V .

(ii) Esprimere in base \mathbf{F} il vettore $\underline{v} = \underline{e}_1 + 2\underline{e}_2 + 3\underline{e}_3$.

(iii) Scrivere la formula di cambiamento di coordinate di vettore dalla base \mathbf{E} alla base \mathbf{F} .

(iv) Esiste in V un vettore non nullo avente le stesse coordinate nelle due basi?

4.1.2. Sia $V = V_{\mathbf{R}}^3$, con base \mathbf{E} . Sia $\mathbf{F} = (\underline{f}_1 \ \underline{f}_2 \ \underline{f}_3)$ un'altra base. Indichiamo con \mathbf{x} la colonna delle coordinate di un generico vettore $\underline{v} \in V$ rispetto alla base \mathbf{E} e con \mathbf{y} la colonna delle coordinate dello stesso vettore rispetto alla base \mathbf{F} . Determinare la base \mathbf{F} in funzione di \mathbf{E} , sapendo che

$$\begin{cases} y_1 = x_1 - x_2 \\ y_2 = x_2 - x_3 \\ y_3 = x_1 + x_3 \end{cases}$$

4.1.3. Siano $\{\underline{e}_1, \underline{e}_2, \underline{e}_3\}$ e $\{\underline{f}_1, \underline{f}_2, \underline{f}_3\}$ due basi di uno spazio vettoriale $V = V_{\mathbf{R}}^3$. Sia:

$$\underline{f}_1 = \underline{e}_1 - \underline{e}_2, \quad \underline{f}_2 = -\underline{e}_1 + \underline{e}_3, \quad \underline{f}_3 = 2\underline{e}_1 + \underline{e}_2.$$

Assegnato il sottospazio vettoriale $W = \langle \underline{e}_1 - \underline{e}_3, \underline{e}_2 + \underline{e}_3 \rangle$, determinarne un sistema di generatori espresso rispetto alla base $\{\underline{f}_1, \underline{f}_2, \underline{f}_3\}$.

4.1.4. Sono assegnate in $V = \mathfrak{M}_2(\mathbf{R})$ le quattro matrici

$$J_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad J_2 = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \quad J_3 = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}, \quad J_4 = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

(i) Verificare che $\{J_1, J_2, J_3, J_4\}$ è una base di V .

(ii) Esprimere in tale base la matrice $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$.

4.1.5. Sia $V = V_{\mathbf{R}}^3$, con base \mathbf{E} . Siano $\underline{f}_1, \underline{f}_2 \in V$ tali che

$$\underline{f}_1 = \underline{e}_1 - \underline{e}_2 + \underline{e}_3, \quad \underline{f}_2 = \underline{e}_2 + \underline{e}_3.$$

Verificare se esiste una base $\mathbf{F} = (\underline{f}_1, \underline{f}_2, \underline{f}_3)$ tale che un vettore $\underline{v} \in V$ abbia in base \mathbf{E} coordinate $(1, 1, 1)$ ed in base \mathbf{F} coordinate $(0, 1, 1)$.

4.1.6. Rifare l'**Esercizio 2.5.5**.

4.1.7. In $V = V_{\mathbf{R}}^5$, con base \mathbf{E} , è assegnato il sottospazio vettoriale $U = \langle \underline{u}_1, \underline{u}_2, \underline{u}_3 \rangle$, con

$$\underline{u}_1 = \underline{e}_1 - \underline{e}_2 + \underline{e}_5, \quad \underline{u}_2 = 2\underline{e}_1 - \underline{e}_3 + \underline{e}_4, \quad \underline{u}_3 = -2\underline{e}_2 + \underline{e}_3 - \underline{e}_4 + 2\underline{e}_5.$$

Determinare equazioni cartesiane di U , cioè un *SLO* in cinque incognite ed a valori in \mathbf{R} , le cui soluzioni siano tutte e sole le coordinate dei vettori di U , in base \mathbf{E} .

2. Applicazioni lineari e relativa rappresentazione

Abbiamo già introdotto, nel **Cap. 2.4**, la definizione di applicazione lineare (o omomorfismo) tra due spazi vettoriali. Richiamiamola: *se V, W sono due K -spazi vettoriali, un'applicazione $T : V \rightarrow W$ è detta applicazione lineare se verifica le due condizioni:*

$$T(\underline{v}_1 + \underline{v}_2) = T(\underline{v}_1) + T(\underline{v}_2), \quad T(c\underline{v}) = cT(\underline{v}), \quad \forall \underline{v}, \underline{v}_1, \underline{v}_2 \in V, \quad \forall c \in K.$$

Le due condizioni precedenti possono essere riunificate nella sola condizione:

$$T(c_1\underline{v}_1 + c_2\underline{v}_2) = c_1T(\underline{v}_1) + c_2T(\underline{v}_2), \quad \forall \underline{v}_1, \underline{v}_2 \in V, \quad \forall c_1, c_2 \in K,$$

ovvero nella condizione

$$T(c_1\underline{v}_1 + \dots + c_n\underline{v}_n) = c_1T(\underline{v}_1) + \dots + c_nT(\underline{v}_n), \quad \forall \underline{v}_1, \dots, \underline{v}_n \in V, \quad \forall c_1, \dots, c_n \in K.$$

Come si vede, un'applicazione T è lineare se e solo se "conserva le combinazioni lineari", nel senso che l'immagine di una combinazione lineare di vettori è la combinazione lineare dei vettori immagine, con gli stessi coefficienti.

Un'applicazione lineare di V in V viene chiamata, come già detto nel **Cap. 2.4**, *operatore lineare* (o anche *endomorfismo*) di V . Un operatore lineare biettivo è detto *automorfismo* di V .

Ricordiamo alcuni fatti già noti (cfr. **Prop. 1 e 6 di Cap. 2.4**) relativi alle le applicazioni lineari.

1. *Se $T : V \rightarrow W$ e $S : W \rightarrow U$ sono applicazioni lineari, anche $S \circ T : V \rightarrow U$ è un'applicazione lineare.*

2. *Se $T : V \rightarrow W$ è un isomorfismo di K -spazi vettoriali, anche l'applicazione $T^{-1} : W \rightarrow V$ è un isomorfismo di K -spazi vettoriali.*

3. *Ogni applicazione lineare $T : V \rightarrow W$ definisce i due sottospazi vettoriali*

$$\text{Ker}(T) = T^{-1}(\underline{0}_W) = \{\underline{v} \in V \mid T(\underline{v}) = \underline{0}_W\}, \quad \text{detto nucleo di } T;$$

$$\text{Im}(T) = T(V) = \{T(\underline{v}), \quad \forall \underline{v} \in V\}, \quad \text{detto immagine di } T.$$

La precedente proprietà **3** può facilmente generalizzata, verificando che, assegnata un'applicazione lineare, l'immagine e la controimmagine di sottospazi vettoriali sono sottospazi vettoriali. Vale infatti il seguente risultato.

Proposizione 1. *Sia $T : V \rightarrow W$ un'applicazione lineare tra K -spazi vettoriali. Per ogni sottospazio vettoriale V' di V ed ogni sottospazio vettoriale W' di W , risulta:*

$T(V')$ è un sottospazio vettoriale di W ; $T^{-1}(W')$ è un sottospazio vettoriale di V .

Dim. Presi comunque $\underline{w}_1 = T(\underline{v}_1), \underline{w}_2 = T(\underline{v}_2) \in T(V')$ [con $\underline{v}_1, \underline{v}_2 \in V'$] e presi comunque $c, d \in K$, si ha (tenuto conto che $c\underline{v}_1 + d\underline{v}_2 \in V'$):

$$c\underline{w}_1 + d\underline{w}_2 = cT(\underline{v}_1) + dT(\underline{v}_2) = T(c\underline{v}_1 + d\underline{v}_2) \in T(V').$$

Analogamente, presi comunque $\underline{v}_1, \underline{v}_2 \in T^{-1}(W')$ [con $T(\underline{v}_1) = \underline{w}_1, T(\underline{v}_2) = \underline{w}_2$] e presi comunque $c, d \in K$, si ha:

$$T(c\underline{v}_1 + d\underline{v}_2) = cT(\underline{v}_1) + dT(\underline{v}_2) = c\underline{w}_1 + d\underline{w}_2 \in W'$$

e quindi $c\underline{v}_1 + d\underline{v}_2 \in T^{-1}(W')$.

Vediamo qualche esempio di applicazione lineare e di operatore lineare.

(i) Siano V, W due K -spazi vettoriali. L'applicazione

$$\mathbf{0} : V \rightarrow W \quad \text{tale che } \mathbf{0}(\underline{v}) = \underline{0}_W, \quad \forall \underline{v} \in V,$$

[dove $\underline{0}_W$ ovviamente è il vettore nullo di W] è, come subito si verifica, un'applicazione lineare da V a W , detta *applicazione lineare nulla*.

(ii) Sia V un K -spazio vettoriale e si fissi in K uno scalare c . L'applicazione

$$c\mathbf{1}_V : V \rightarrow V \text{ tale che } c\mathbf{1}_V(\underline{v}) = c\underline{v}, \forall \underline{v} \in V$$

è, come subito si verifica, un operatore lineare di V , detto *operatore di moltiplicazione per lo scalare* c . Si osserva poi subito che, se $c \neq 0$, $c\mathbf{1}_V$ è un automorfismo di V [con inverso $\frac{1}{c}\mathbf{1}_V$]. Per $c = 1$, abbiamo l'*operatore identità* $\mathbf{1}_V$ di V . Per $c = 0$, abbiamo l'*operatore nullo* $\mathbf{0}_V$ di V .

(iii) Siano $T : V \rightarrow W$, $S : V \rightarrow W$ due applicazioni lineari. È definita l'applicazione

$$T + S : V \rightarrow W \text{ tale che } (T + S)(\underline{v}) = T(\underline{v}) + S(\underline{v}), \forall \underline{v} \in V.$$

Si verifica che $T + S$ è un'applicazione lineare da V a W , detta *applicazione lineare somma di* T , S . Analogamente, se $c \in K$ e $T : V \rightarrow W$ è un'applicazione lineare, è definita l'applicazione

$$cT : V \rightarrow W \text{ tale che } (cT)(\underline{v}) = cT(\underline{v}), \forall \underline{v} \in V.$$

Anch'essa è un'applicazione lineare [infatti coincide con $T \circ c\mathbf{1}_V$], ed è detta *moltiplicazione di* T per lo scalare c . Valgono le seguenti proprietà che sono di immediata verifica: $\forall c, d \in K, \forall S, T, R$ applicazioni lineari da V a W ,

$$\begin{aligned} T + \mathbf{0} &= \mathbf{0} + T = T, \quad T + (-T) = (-T) + T = \mathbf{0}, \quad (T + S) + R = T + (S + R), \quad T + S = S + T, \\ c(T + S) &= cT + cS, \quad (c + d)T = cT + dT, \quad c(dT) = (cd)T, \quad 1T = T. \end{aligned}$$

Si conclude che l'insieme delle applicazioni lineari da V a W è esso stesso un K -spazio vettoriale, rispetto all'operazione di somma [rispetto a cui è un gruppo commutativo] ed alla moltiplicazione per uno scalare. Tale spazio vettoriale viene denotato $\mathcal{H}om_K(V, W)$.

Osserviamo poi che, se $W = V$, lo spazio vettoriale $\mathcal{H}om_K(V, V)$ è dotato anche dell'operazione di composizione di applicazione lineari \circ . Poiché, $\forall T, R, S \in \mathcal{H}om_K(V, V)$, si ha:

$\mathbf{1}_V \circ T = T = T \circ \mathbf{1}_V$, $(T \circ R) \circ S = T \circ (R \circ S)$, $(T + S) \circ R = T \circ R + S \circ R$ e $T \circ (R + S) = T \circ R + T \circ S$, allora $\mathcal{H}om_K(V, V)$ è un anello unitario, di solito denotato $\mathcal{E}nd_K(V)$ e detto *anello degli operatori lineari* (o degli endomorfismi) di V . Infine, il gruppo degli elementi invertibili di $\mathcal{E}nd_K(V)$ (cioè degli automorfismi di V) viene usualmente denotato $\mathcal{A}ut_K(V)$.

Ci chiediamo quante siano le applicazioni lineari tra due spazi vettoriali o, più precisamente, quale sia la dimensione dello spazio vettoriale $\mathcal{H}om_K(V, W)$. Abbiamo un significativo risultato, nell'ipotesi che almeno il primo dei due spazi vettoriali abbia dimensione finita.

Sia $V = V_K^n$ e sia $\{\underline{e}_1, \underline{e}_2, \dots, \underline{e}_n\}$ una sua base. Ogni applicazione lineare $T : V \rightarrow W$ è completamente individuata conoscendo gli n vettori

$$T(\underline{e}_1), T(\underline{e}_2), \dots, T(\underline{e}_n) \in W.$$

Infatti, $\forall \underline{v} \in V$, se $\underline{v} = \sum_{i=1}^n c_i \underline{e}_i$, si ha: $T(\underline{v}) = \sum_{i=1}^n c_i T(\underline{e}_i)$. Dunque $T(\underline{v})$ è calcolabile conoscendo i vettori $T(\underline{e}_i)$ (e ovviamente le coordinate di \underline{v}).

Viceversa, scelti arbitrariamente in W n vettori $\underline{w}_1, \underline{w}_2, \dots, \underline{w}_n$, possiamo considerare l'applicazione $T : V \rightarrow W$ così definita:

$$T\left(\sum_{i=1}^n c_i \underline{e}_i\right) = \sum_{i=1}^n c_i \underline{w}_i, \quad \forall \sum_{i=1}^n c_i \underline{e}_i \in V.$$

Tale applicazione è "tautologicamente" lineare in quanto è - come si dice - "costruita per linearità" a partire dalle n assegnazioni $T(\underline{e}_i) = \underline{w}_i, \forall i = 1, \dots, n$.

Abbiamo così costruito due applicazioni (dipendenti dalla scelta di una base $\{\underline{e}_1, \underline{e}_2, \dots, \underline{e}_n\}$ di V). La prima:

$$\Phi : \mathcal{H}om_K(V, W) \rightarrow W^n$$

associa a $T \in \mathcal{H}om_K(V, W)$ la n -pla $\Phi(T) = (T(\underline{e}_1), T(\underline{e}_2), \dots, T(\underline{e}_n)) \in W^n$. La seconda:

$$\Psi : W^n \rightarrow \mathcal{H}om_K(V, W)$$

associa a $(\underline{w}_1, \underline{w}_2, \dots, \underline{w}_n) \in W^n$ l'applicazione (costruita per linearità) $\Psi((\underline{w}_1, \underline{w}_2, \dots, \underline{w}_n))$ [che denotiamo T] tale che $T(\underline{e}_i) = \underline{w}_i, \forall i = 1, \dots, n$.

Queste due applicazioni sono manifestamente inverse l'una dell'altra e quindi creano una biiezione tra $\mathcal{H}om_K(V, W)$ e W^n .

Ci chiediamo allora se tali applicazioni siano isomorfismi di spazi vettoriali. Intanto osserviamo subito che W^n è dotato in modo ovvio di struttura di spazio vettoriale [rispetto alle due operazioni

$$(\underline{w}_1, \dots, \underline{w}_n) + (\underline{w}'_1, \dots, \underline{w}'_n) = (\underline{w}_1 + \underline{w}'_1, \dots, \underline{w}_n + \underline{w}'_n), \quad c(\underline{w}_1, \dots, \underline{w}_n) = (c\underline{w}_1, \dots, c\underline{w}_n).$$

Si ha poi, $\forall c, d \in K, \forall T, S \in \mathcal{H}om_K(V, W)$:

$$\begin{aligned} \Phi(cT + dS) &= \\ &= ((cT + dS)(\underline{e}_1), \dots, (cT + dS)(\underline{e}_n)) = \\ &= c(T(\underline{e}_1), \dots, T(\underline{e}_n)) + d(S(\underline{e}_1), \dots, S(\underline{e}_n)) = c\Phi(T) + d\Phi(S). \end{aligned}$$

Dunque Φ è un isomorfismo ed abbiamo così provato il seguente risultato.

Proposizione 2. *Sia $V = V_K^n$ e sia $\{\underline{e}_1, \underline{e}_2, \dots, \underline{e}_n\}$ una sua base. Per ogni K -spazio vettoriale W esiste un isomorfismo (dipendente dalla base scelta) tra $\mathcal{H}om_K(V_K^n, W)$ e W^n .*

Osservazione 1. Si noti che, se $\dim(W) = m$, allora $\dim(W^n) = nm$.

Se infatti $\{\underline{f}_1, \dots, \underline{f}_m\}$ è una base di W , si può verificare che le nm n -ple:

$$(\underline{f}_{i_1}, \underline{0}, \dots, \underline{0}), (\underline{0}, \underline{f}_{i_2}, \underline{0}, \dots, \underline{0}), \dots, (\underline{0}, \underline{0}, \dots, \underline{0}, \underline{f}_{i_n}), \text{ con } i_1, i_2, \dots, i_n \in \{1, 2, \dots, m\},$$

formano una base di W^n . Segue dalla **Prop. 2** che

$$\dim(\mathcal{H}om_K(V_K^n, W_K^m)) = nm.$$

Ora ci poniamo il problema di rappresentare le applicazioni lineari tra spazi vettoriali di dimensione finita.

Sia $T : V \rightarrow W$ un'applicazione线are, con $V = V_K^n$ e $W = W_K^m$. Supponiamo che siano assegnate una base $\{\underline{e}_1, \dots, \underline{e}_n\}$ di V ed una base $\{\underline{f}_1, \dots, \underline{f}_m\}$ di W . Con le notazioni introdotte nel paragrafo precedente, poniamo:

$$\mathbf{E} := (\underline{e}_1 \ \dots \ \underline{e}_n) \in \mathfrak{M}_{1,n}(V), \quad \mathbf{F} := (\underline{f}_1 \ \dots \ \underline{f}_m) \in \mathfrak{M}_{1,m}(W).$$

Al variare di $\underline{v} \in V$, sia

$$\underline{w} = T(\underline{v})$$

la sua immagine. Vogliamo tradurre tale uguaglianza tra vettori nella corrispondente uguaglianza tra le loro coordinate (nelle basi \mathbf{E}, \mathbf{F}).

Siano $\underline{v} = \mathbf{E}\mathbf{x}$ e $\underline{w} = \mathbf{F}\mathbf{y}$, con $\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathfrak{M}_{n,1}(K)$ e $\mathbf{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} \in \mathfrak{M}_{m,1}(K)$. Si ha:

$$(\bullet) \quad T(\underline{v}) = T\left(\sum_{i=1}^n \underline{e}_i x_i\right) = \sum_{i=1}^n T(\underline{e}_i) x_i = (T(\underline{e}_1) \ \dots \ T(\underline{e}_n)) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = (T(\underline{e}_1) \ \dots \ T(\underline{e}_n)) \mathbf{x}.$$

Come già osservato, gli n vettori $T(\underline{e}_1), T(\underline{e}_2), \dots, T(\underline{e}_n) \in W$ individuano completamente T . Esprimiamo tali vettori nella base \mathbf{F} . Sia

$$T(\underline{e}_1) = \mathbf{F}\mathbf{a}_1, \quad T(\underline{e}_2) = \mathbf{F}\mathbf{a}_2, \quad \dots, \quad T(\underline{e}_n) = \mathbf{F}\mathbf{a}_n,$$

con $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n \in \mathfrak{M}_{m,1}(K)$. Riunendo tali colonne si ottiene la matrice

$$A = (\mathbf{a}_1 \ \mathbf{a}_2 \ \dots \ \mathbf{a}_n) \in \mathfrak{M}_{m,n}(K),$$

che viene detta *matrice dell'applicazione lineare T rispetto alle basi \mathbf{E} ed \mathbf{F}* .

Riunendo i vettori $T(\underline{e}_1), T(\underline{e}_2), \dots, T(\underline{e}_n)$ in un vettore riga, abbiamo

$$(T(\underline{e}_1) \ T(\underline{e}_2) \ \dots \ T(\underline{e}_n)) = \mathbf{F} A.$$

Se poi, per semplificare ulteriormente le notazioni, poniamo

$$T(\mathbf{E}) := (T(\underline{e}_1) \ T(\underline{e}_2) \ \dots \ T(\underline{e}_n)),$$

otteniamo la formula

$$T(\mathbf{E}) = \mathbf{F} A,$$

che fornisce l'azione di T sui vettori della base \mathbf{E} e che chiameremo *formula di definizione di T rispetto alle basi \mathbf{E} ed \mathbf{F}* .

N.B. La notazione $T(\mathbf{E})$ va utilizzata con molta precisione ed accuratezza: si faccia attenzione a non scriverla (o interpretarla) nella forma $T((\underline{e}_1 \ \underline{e}_2 \ \dots \ \underline{e}_n))$ o nella forma $T(\underline{e}_1 \ \underline{e}_2 \ \dots \ \underline{e}_n)$, che non hanno senso [in quanto T agisce su singoli vettori (e non su righe di vettori)]. Con tale notazione, la precedente formula (•) si scrive nella forma:

$$T(\mathbf{E} \mathbf{x}) = T(\mathbf{E}) \mathbf{x}.$$

Torniamo all'uguaglianza $\underline{w} = T(\underline{v})$. Si ha:

$$\mathbf{F} \mathbf{y} = \underline{w} = T(\underline{v}) = T(\mathbf{E} \mathbf{x}) = T(\mathbf{E}) \mathbf{x} = \mathbf{F} A \mathbf{x}$$

e quindi $\mathbf{F} \mathbf{y} = \mathbf{F} A \mathbf{x}$. Per l'unicità di scrittura in base \mathbf{F} si conclude che

$$\mathbf{y} = A \mathbf{x}.$$

Interpretando \mathbf{x} ed \mathbf{y} come colonne di incognite, diremo che tale uguaglianza fornisce le *equazioni di T , nelle basi \mathbf{E}, \mathbf{F}* .

Il significato di tali equazioni è chiaro: sostituendo le coordinate di un vettore $\underline{v} \in V$ al posto della colonna \mathbf{x} , la corrispondente colonna \mathbf{y} fornisce le coordinate di $T(\underline{v})$.

In forma non compatta le precedenti equazioni si scrivono nella forma:

$$\begin{cases} y_1 = a_{11} x_1 + a_{12} x_2 + \dots + a_{1n} x_n \\ y_2 = a_{21} x_1 + a_{22} x_2 + \dots + a_{2n} x_n \\ \vdots \\ y_m = a_{m1} x_1 + a_{m2} x_2 + \dots + a_{mn} x_n. \end{cases}$$

Un esempio. Sia $T : \mathbf{R}^3 \rightarrow \mathbf{R}^4$ l'applicazione lineare definita, rispetto alle basi canoniche \mathbf{E} di \mathbf{R}^3 ed \mathbf{F} di \mathbf{R}^4 , dalle seguenti condizioni:

$$T(\underline{e}_1) = (1, 0, 0, 1), \quad T(\underline{e}_2) = (0, -1, 1, 0), \quad T(\underline{e}_3) = (1, 2, 0, 1).$$

La matrice di T rispetto alle due basi fissate è

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

e le equazioni di T (sempre rispetto a tali basi) sono

$$\begin{cases} y_1 = x_1 + x_3 \\ y_2 = -x_2 + 2x_3 \\ y_3 = x_2 \\ y_4 = x_1 + x_3. \end{cases}$$

Assegnato ad esempio il vettore $\underline{v} = (1, 2, 3) \in \mathbf{R}^3$, il vettore $\underline{w} = T(\underline{v})$ è ottenuto calcolando $A \mathbf{v}$ [dove \mathbf{v} è la colonna corrispondente al vettore \underline{v}]. Si ha:

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

e dunque $\underline{w} = (4, 4, 2, 4)$.

Potremmo anche chiederci quali altri eventuali vettori di \mathbf{R}^3 abbiano per immagine \underline{w} . Ciò significa ovviamente determinare $T^{-1}(\underline{w})$.

Per ottenere tali vettori basta risolvere il $SL(4, 3, \mathbf{R})$ $AX = \mathbf{w}$, cioè

$$\begin{cases} x_1 + x_3 = 4 \\ -x_2 + 2x_3 = 4 \\ x_2 = 2 \\ x_1 + x_3 = 4. \end{cases}$$

Tale SL è compatibile [in quanto ha soluzione $(1, 2, 3)$]. Poiché $rg(A) = 3$, il SL ha ∞^{3-3} soluzioni, cioè un'unica soluzione [che è quindi $(1, 2, 3)$]. Dunque $T^{-1}(\underline{w}) = \{\underline{v}\}$.

La matrice A di un'applicazione lineare $T : V_K^n \rightarrow W_K^m$ [sempre relativamente a due basi assegnate \mathbf{E}, \mathbf{F}] consente di determinare facilmente il nucleo $Ker(T)$ e l'immagine $Im(T)$ di T .

Si ha infatti (con le notazioni usate in precedenza):

$$\underline{v} \in Ker(T) \iff T(\underline{v}) = \underline{0}_W \iff \mathbf{F} A \mathbf{x} = \mathbf{F} \mathbf{0} \iff A \mathbf{x} = \mathbf{0}$$

[dove ovviamente $\mathbf{0} \in \mathfrak{M}_{m,1}(K)$]. Dunque

$Ker(T)$ ha (in base \mathbf{E}) equazioni cartesiane date dal $SLO(m, n, K)$ $AX = \mathbf{0}$.

Segue, dal teorema di Rouché-Capelli, che

$$\dim(Ker(T)) = n - rg(A).$$

Relativamente ad $Im(T)$, osserviamo che ogni vettore $T(\underline{v}) \in Im(T)$ è combinazione lineare di $T(\underline{e}_1), \dots, T(\underline{e}_n)$. Pertanto:

$$Im(T) = \langle T(\underline{e}_1), \dots, T(\underline{e}_n) \rangle.$$

Quindi una base di $Im(T)$ è ottenuta scegliendo il massimo numero di vettori linearmente indipendenti tra $T(\underline{e}_1), \dots, T(\underline{e}_n)$. Poiché

$$(T(\underline{e}_1) \ \dots \ T(\underline{e}_n)) = \mathbf{F} A,$$

basta scegliere il massimo numero possibile di colonne di A linearmente indipendenti ed i corrispondenti vettori $T(\underline{e}_i)$ formano una base di $Im(T)$. In particolare abbiamo ottenuto:

$$\dim(Im(T)) = rg(A).$$

Un esempio. Sia $V = V_{\mathbf{R}}^4$ e sia $T : V \rightarrow V$ l'operatore lineare definito, rispetto ad una stessa base \mathbf{E} di V [sia nello spazio di partenza che in quello di arrivo] dalla matrice:

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

Vogliamo determinare $Ker(T)$ e $Im(T)$.

$Ker(T)$ è ottenuto risolvendo il $SLO(4, 4, \mathbf{R})$ $AX = \mathbf{0}$.

Poiché la matrice A ha rango 3 [infatti si osserva che le prime tre righe di A sono linearmente indipendenti, mentre la quarta è la somma delle prime tre], il SLO ha $\infty^{4-3} = \infty^1$ soluzioni, ovvero lo spazio vettoriale Σ_0 delle sue soluzioni è un sottospazio 1-dimensionale di \mathbf{R}^4 .

Per ottenere una base di Σ_0 , cioè un'autosoluzione, conviene utilizzare la **Prop. 1** di **Cap. 3.4**. Si elimina dal SLO l'ultima equazione (che è superflua) e si determina il vettore $\underline{z} \in \mathbf{R}^4$ formato dai minori di ordine 3 (delle prime tre righe di A), presi a segni alterni. Si ottiene:

$$\left(\begin{vmatrix} 0 & -1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{vmatrix}, - \begin{vmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{vmatrix}, \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 1 & 1 \end{vmatrix}, - \begin{vmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \\ -2 & 1 & 0 \end{vmatrix} \right) = (1, -1, 1, 3).$$

Dunque $\Sigma_0 = \langle (1, -1, 1, 3) \rangle$ e pertanto

$$Ker(T) = \langle \underline{e}_1 - \underline{e}_2 + \underline{e}_3 + 3\underline{e}_4 \rangle.$$

Per ottenere una base di $Im(T)$ basta scegliere tre colonne di A linearmente indipendenti. Abbiamo varie alternative. Ad esempio possiamo scegliere le ultime tre colonne. Otteniamo quindi

$$Im(T) = \langle \underline{e}_2 + \underline{e}_3 + 2\underline{e}_4, -\underline{e}_1 + \underline{e}_2, \underline{e}_3 + \underline{e}_4 \rangle.$$

Lo studente avrà forse notato (sia nell'ultimo esempio che nelle considerazioni precedenti) che risulta

$$\dim(Ker(T)) + \dim(Im(T)) = \dim(V)$$

[infatti $(n - rg(A)) + rg(A) = n$]. Questa formula può essere dimostrata anche in forma "intrinseca", cioè senza utilizzare la matrice A [che non è unica, in quanto dipende dalla scelta delle basi \mathbf{E} di V ed \mathbf{F} di W], e vale, più generalmente, anche se W ha dimensione non finita. Ridimostriamo dunque in forma intrinseca tale risultato, che è noto con il nome di *teorema della nullità più rango* [dove "nullità" sta per "dimensione del nucleo"].

Teorema 1. Sia $T : V \rightarrow W$ un'applicazione lineare, con $V = V_K^n$. Risulta:

$$\dim(Ker(T)) + \dim(Im(T)) = \dim(V) [= n].$$

Dim. Assumiamo che $Ker(T)$ abbia dimensione s e scegliamone una base $\{\underline{u}_1, \dots, \underline{u}_s\}$. In base al teorema del completamento, possiamo aggiungere a tali vettori altri $n - s$ vettori $\underline{v}_1, \dots, \underline{v}_{n-s}$ di V , sino ad ottenere la base $\{\underline{u}_1, \dots, \underline{u}_s, \underline{v}_1, \dots, \underline{v}_{n-s}\}$ di V . Basterà ora verificare che i vettori

$$T(\underline{v}_1), \dots, T(\underline{v}_{n-s})$$

formano una base di $Im(T)$, e la formula sarà dimostrata.

Per ogni $\underline{w} \in Im(T)$, si ha:

$$\underline{w} = T(\underline{v}) = T\left(\sum_{i=1}^s a_i \underline{u}_i + \sum_{j=1}^{n-s} b_j \underline{v}_j\right) = \sum_{i=1}^s a_i T(\underline{u}_i) + \sum_{j=1}^{n-s} b_j T(\underline{v}_j) = \sum_{j=1}^{n-s} b_j T(\underline{v}_j)$$

[in quanto ogni $T(\underline{u}_i) = \underline{0}_W$]. Abbiamo quindi verificato che $\underline{w} \in \langle T(\underline{v}_1), \dots, T(\underline{v}_{n-s}) \rangle$, cioè che $\{T(\underline{v}_1), \dots, T(\underline{v}_{n-s})\}$ è un sistema di generatori di W .

Verifichiamo ora che i vettori $T(\underline{v}_1), \dots, T(\underline{v}_{n-s})$ sono linearmente indipendenti. Sia

$$\sum_{j=1}^{n-s} b_j T(\underline{v}_j) = \underline{0}_W.$$

Ne segue che $T\left(\sum_{j=1}^{n-s} b_j \underline{v}_j\right) = \underline{0}_W$, cioè che $\sum_{j=1}^{n-s} b_j \underline{v}_j \in Ker(T)$. Allora,

$$\sum_{j=1}^{n-s} b_j \underline{v}_j = \sum_{i=1}^s a_i \underline{u}_i, \text{ cioè } \sum_{j=1}^{n-s} b_j \underline{v}_j - \sum_{i=1}^s a_i \underline{u}_i = \underline{0},$$

per opportuni coefficienti $a_i \in K$. Poiché $\underline{u}_1, \dots, \underline{u}_s, \underline{v}_1, \dots, \underline{v}_{n-s}$ sono linearmente indipendenti, tutti i coefficienti b_j [e a_i] sono nulli. Pertanto $T(\underline{v}_1), \dots, T(\underline{v}_{n-s})$ sono linearmente indipendenti.

Osservazione 2. Ritorniamo ad un'applicazione lineare $T : V_K^n \rightarrow W_K^m$ che, rispetto alle basi \mathbf{E} di V ed \mathbf{F} di W , abbia matrice $A \in \mathfrak{M}_{m,n}(K)$.

Analogamente a ciò che è stato fatto per $Ker(T)$, vogliamo ottenere equazioni cartesiane di $Im(T)$, cioè un *SLO* ad m incognite, le cui soluzioni siano le coordinate, in base \mathbf{F} , di tutti e soli i vettori di $Im(T)$.

$$\begin{aligned} \text{Risulta: } \underline{w} \in Im(T) &\iff \underline{w} = T(\underline{v}), \exists \underline{v} \in V \\ &\iff \mathbf{F}\mathbf{y} = \mathbf{F}\mathbf{A}\mathbf{x}, \exists \mathbf{E}\mathbf{x} \in V \\ &\iff \mathbf{A}\mathbf{x} = \mathbf{y}, \exists \mathbf{x} \in \mathfrak{M}_{n,1}(K) \\ &\iff \text{il } SL(m, n, K) \text{ } AX = \mathbf{y} \text{ è compatibile} \\ &\iff rg((A \ \mathbf{y})) = rg(A). \end{aligned}$$

Sia $r = rg(A)$ e sia B una sottomatrice quadrata invertibile di ordine r in A . La condizione $rg((A \ \mathbf{y})) = r$ equivale a che siano tutti nulli gli orlati di B . Tra questi, quelli (se ce ne sono) che non coinvolgono la colonna \mathbf{y} sono già nulli [in quanto $rg(A) = r$]. Quelli che coinvolgono la colonna \mathbf{y} sono esattamente $m - r$. Interpretando gli elementi di \mathbf{y} come incognite, l'annullamento di questi $m - r$ minori crea un *SLO* ($m - r, m, K$), le cui soluzioni sono le coordinate, in base \mathbf{F} , di tutti e soli i vettori di $Im(T)$.

Facciamo un esempio. Sia $T : \mathbf{R}^4 \rightarrow \mathbf{R}^5$ l'applicazione lineare avente, rispetto alle basi canoniche \mathbf{E} di \mathbf{R}^4 ed \mathbf{F} di \mathbf{R}^5 , matrice

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

Si osserva che $rg(A) = 2$ [infatti ad esempio le colonne $A_{(1)}, A_{(3)}$ sono linearmente indipendenti, mentre $A_{(2)} = A_{(3)} - A_{(1)}$, $A_{(4)} = A_{(1)} + A_{(3)}$].

Scegliamo come sottomatrice B di A la matrice $B = A(1, 2 | 1, 3) = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$. La condizione $rg((A \ y)) = 2$ ci dà l'annullamento dei tre orlati [coinvolgenti la colonna di incognite $\mathbf{y} \in \mathfrak{M}_{5,1}(\mathbf{R})$]:

$$\begin{vmatrix} 1 & 2 & y_1 \\ 0 & 1 & y_2 \\ -1 & 0 & y_3 \end{vmatrix} = 0, \quad \begin{vmatrix} 1 & 2 & y_1 \\ 0 & 1 & y_2 \\ 2 & 0 & y_4 \end{vmatrix} = 0, \quad \begin{vmatrix} 1 & 2 & y_1 \\ 0 & 1 & y_2 \\ 0 & 1 & y_5 \end{vmatrix} = 0,$$

cioè il $SLO(3, 5, \mathbf{R})$

$$\begin{cases} y_1 - 2y_2 + y_3 = 0 \\ -2y_1 + 4y_2 + y_4 = 0 \\ -y_2 = y_5 = 0. \end{cases}$$

Si noti che una base di $Im(T)$ è data da $\{(1, 0, -1, 2, 0), (2, 1, 0, 0, 1)\}$ [prima e terza colonna di A]. Ovviamente queste due 5-pie sono una base di autosoluzioni del SLO .

Lasciamo al lettore la verifica che una base di $Ker(T)$ è ad esempio $\{(1, 1, -1, 0), (2, 1, 0, -1)\}$.

Il **Teor. 1** ha un importante corollario nel caso in cui V, W hanno la stessa dimensione.

Corollario 1. Sia $dim(V) = dim(W) = n$ e sia $T : V \rightarrow W$ un'applicazione lineare. Risulta:

$$T \text{ è iniettiva} \iff T \text{ è suriettiva.}$$

[Dunque T è un isomorfismo \iff è un monomorfismo \iff è un epimorfismo].

Dim. Dal **Teor. 1**, $dim(Ker(T)) = 0 \iff dim(Im(T)) = n$. Quindi

$$Ker(T) = \{\underline{0}\} \iff Im(T) = W.$$

Per concludere basta ricordare (cfr. **Cap. 2.4**) che T è iniettiva $\iff Ker(T) = \{\underline{0}\}$.

Nei due risultati che seguono otteniamo la rappresentazione della combinazione lineare di due applicazioni lineari e della composizione di due applicazioni lineari.

Proposizione 3. Siano $V = V_K^n$ e $W = W_K^m$, con basi rispettivamente \mathbf{E}, \mathbf{F} . Siano $c, d \in K$ e siano $S, T : V \rightarrow W$ due applicazioni lineari aventi rispettivamente matrici A, B (rispetto ad \mathbf{E}, \mathbf{F}). Allora:

$$cS + dT : V \rightarrow W \text{ ha matrice } cA + dB \text{ (rispetto ad } \mathbf{E}, \mathbf{F}).$$

Dim. Per $i = 1, \dots, n$ risulta:

$$(cS + dT)(\underline{e}_i) = cS(\underline{e}_i) + dT(\underline{e}_i) = c\mathbf{F}A_{(i)} + d\mathbf{F}B_{(i)} = \mathbf{F}(cA_{(i)} + dB_{(i)}) = \mathbf{F}(cA + dB)_{(i)}.$$

Dunque

$$\left((cS + dT)(\underline{e}_1) \ \dots \ (cS + dT)(\underline{e}_n) \right) = \mathbf{F}(cA + dB).$$

N.B. Utilizzando le formule $T(\mathbf{E}) = \mathbf{F}A$ e $S(\mathbf{E}) = \mathbf{F}B$, si ha subito:

$$(cS + dT)(\mathbf{E}) = cS(\mathbf{E}) + dT(\mathbf{E}) = c\mathbf{F}A + d\mathbf{F}B = \mathbf{F}(cA + dB).$$

Proposizione 4. Siano $V = V_K^n$, $W = W_K^m$ e $U = U_K^p$, con basi rispettivamente $\mathbf{E}, \mathbf{F}, \mathbf{G}$. Siano $T : V \rightarrow W$ e $S : W \rightarrow U$ due applicazioni lineari, aventi rispettivamente matrici A, B (rispetto alle basi considerate). Allora:

$S \circ T : V \rightarrow U$ ha matrice BA (rispetto ad \mathbf{E}, \mathbf{G}).

Dim. Per $i = 1, \dots, n$ risulta:

$$(S \circ T)(\underline{e}_i) = S(T(\underline{e}_i)) = S(\mathbf{F} A_{(i)}) = (S(\underline{f}_1) \ \dots \ S(\underline{f}_m)) A_{(i)} = (\mathbf{G} B) A_{(i)} = \mathbf{G}(B A_{(i)}).$$

Dunque

$$\left((S \circ T)(\underline{e}_1) \ \dots \ (S \circ T)(\underline{e}_n) \right) = \left(\mathbf{G}(B A_{(1)}) \ \dots \ \mathbf{G}(B A_{(n)}) \right) = \mathbf{G}(B A_{(1)} \ \dots \ B A_{(n)}) = \mathbf{G} B A.$$

N.B. Utilizzando le formule $T(\mathbf{E}) = \mathbf{F} A$ e $S(\mathbf{F}) = \mathbf{G} B$, si ha subito:

$$(S \circ T)(\mathbf{E}) = S(T(\mathbf{E})) = S(\mathbf{F} A) = S(\mathbf{F})A = (\mathbf{G} B)A = \mathbf{G}(B A).$$

Un esempio. Siano $T : \mathbf{R}^2 \rightarrow \mathbf{R}^3$ e $S : \mathbf{R}^3 \rightarrow \mathbf{R}^2$ applicazioni lineari definite, rispetto alle basi canoniche \mathbf{E} di \mathbf{R}^2 ed \mathbf{F} di \mathbf{R}^3 , rispettivamente dalle matrici

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

Vogliamo calcolare la matrice di $S \circ T : \mathbf{R}^2 \rightarrow \mathbf{R}^2$, rispetto alla base \mathbf{E} .

Procediamo utilizzando la proposizione precedente. $S \circ T$ ha matrice BA e si ha:

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

Possiamo ottenere la matrice di $S \circ T$ anche con un calcolo diretto. Risulta, dai dati contenuti nelle due matrici:

$$\begin{aligned} T(\underline{e}_1) &= \underline{f}_1 - \underline{f}_3, & T(\underline{e}_2) &= 2\underline{f}_2 + \underline{f}_3, \\ S(\underline{f}_1) &= \underline{e}_2, & S(\underline{f}_2) &= 2\underline{e}_1, & S(\underline{f}_3) &= \underline{e}_1 - \underline{e}_2. \end{aligned}$$

Si ha quindi:

$$\begin{aligned} (S \circ T)(\underline{e}_1) &= S(T(\underline{e}_1)) = S(\underline{f}_1 - \underline{f}_3) = S(\underline{f}_1) - S(\underline{f}_3) = \underline{e}_2 - (\underline{e}_1 - \underline{e}_2) = -\underline{e}_1 + 2\underline{e}_2; \\ (S \circ T)(\underline{e}_2) &= S(T(\underline{e}_2)) = S(2\underline{f}_2 + \underline{f}_3) = 2S(\underline{f}_2) + S(\underline{f}_3) = 4\underline{e}_1 + (\underline{e}_1 - \underline{e}_2) = 5\underline{e}_1 - \underline{e}_2. \end{aligned}$$

La matrice di $S \circ T$ è quindi la matrice BA già ottenuta. Si osservi infine che $BA \in \mathbf{GL}_2(\mathbf{R})$. Ne segue che $\text{Ker}(S \circ T) = \{\underline{0}\}$ e $\text{Im}(S \circ T) = \mathbf{R}^2$.

Lasciamo al lettore il compito di verificare che invece $T \circ S : \mathbf{R}^3 \rightarrow \mathbf{R}^3$ ha matrice (rispetto ad \mathbf{F})

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

Poiché $\text{rg}(AB) < 3$, $\det(AB) = 0$ e quindi $T \circ S$ non è un automorfismo. Si verifichi allora che:

- $\text{Ker}(T \circ S)$ ha equazioni cartesiane $\begin{cases} 2y_2 + y_3 = 0 \\ y_1 - y_3 = 0 \end{cases}$ e che $\text{Ker}(T \circ S) = \langle (2, -1, 2) \rangle$;
- $\text{Im}(T \circ S)$ ha equazione cartesiana $2y_1 - y_2 + 2y_3 = 0$ e che $\text{Im}(T \circ S) = \langle (0, 2, 1), (1, 0, -1) \rangle$.

Concludiamo il paragrafo affrontando il seguente problema: se cambiamo le basi degli spazi vettoriali, come cambia la matrice dell'applicazione lineare?

Studieremo tale problema in un caso particolare, che è l'unico che ci interesserà nel seguito.

Assumiamo che T sia un operatore lineare di $V = V_K^n$ e che $\mathbf{E} = (e_1 \ \dots \ e_n)$ sia la base assegnata in V [sia come spazio di partenza che come spazio d'arrivo dell'operatore]. Assumiamo poi che T abbia matrice $A \in \mathfrak{M}_n(K)$ rispetto a tale base, cioè

$$(T(e_1) \ \dots \ T(e_n)) = \mathbf{E} A.$$

Sia ora $\mathbf{F} = (f_1 \ \dots \ f_n)$ un'altra base di V e sia

$$(T(f_1) \ \dots \ T(f_n)) = \mathbf{F} B, \text{ con } B \in \mathfrak{M}_n(K).$$

Se \mathbf{E} ed \mathbf{F} sono legate da una matrice di cambiamento di base C , cioè $\mathbf{F} = \mathbf{E}C$, con $C \in \mathbf{GL}_n(K)$, proveremo che risulta:

$$B = C^{-1}AC.$$

Poiché $\underline{f}_i = \mathbf{E}C_{(i)}$, allora

$$T(\underline{f}_i) = T(\mathbf{E}C_{(i)}) = (T(e_1) \dots T(e_n)) C_{(i)} = (\mathbf{E}A)C_{(i)} = \mathbf{E}(AC_{(i)}).$$

Pertanto

$$\mathbf{F}B = (T(\underline{f}_1) \dots T(\underline{f}_n)) = (\mathbf{E}(AC_{(1)}) \dots \mathbf{E}(AC_{(n)})) = \mathbf{E}AC.$$

Da $\mathbf{F} = \mathbf{E}C$ segue che $\mathbf{E}CB = \mathbf{E}AC$ e quindi $CB = AC$. Moltiplicando a sinistra per C^{-1} concludiamo che $B = C^{-1}AC$, come preannunciato.

N.B. Utilizzando le formule $T(\mathbf{E}) = \mathbf{E}A$ e $T(\mathbf{F}) = \mathbf{F}B$, si ha subito:

$$\mathbf{F}B = T(\mathbf{F}) = T(\mathbf{E}C) = T(\mathbf{E})C = (\mathbf{E}A)C = \mathbf{E}(AC) = (\mathbf{F}C^{-1})AC = \mathbf{F}(C^{-1}AC)$$

e quindi $B = C^{-1}AC$.

Definizione 1. Due matrici $A, A' \in \mathfrak{M}_n(K)$ sono dette *simili* se esiste $C \in \mathbf{GL}_n(K)$ tale che $A' = C^{-1}AC$. Si può facilmente verificare che l’”essere simili” è una relazione di equivalenza in $\mathfrak{M}_n(K)$.

Abbiamo provato che le matrici che rappresentano un operatore lineare di $V = V_K^n$ sono simili. Ma si può dimostrare anche il viceversa: se due matrici $A, B \in \mathfrak{M}_n(K)$ sono simili e $B = C^{-1}AC$, allora esse rappresentano lo stesso operatore lineare T , avente matrice A in base \mathbf{E} e matrice B in base $\mathbf{E}C$ [infatti risulta: $T(\mathbf{E}C) = T(\mathbf{E})C = (\mathbf{E}A)C = \mathbf{E}(CC^{-1}AC) = (\mathbf{E}C)C^{-1}AC = (\mathbf{E}C)B$].

Osservazione 3. Si verifica subito che una matrice $A \in \mathfrak{M}_n(K)$ è simile solo a se stessa \iff commuta con ogni matrice di $\mathbf{GL}_n(K)$ [infatti $C^{-1}AC = A \iff AC = CA, \forall C \in \mathbf{GL}_n(K)$].

Si verifica poi subito che ogni matrice scalare aI_n commuta con ogni matrice di $M \in \mathfrak{M}_n(K)$ [infatti risulta: $(aI_n)M = aM = M(aI_n)$]. Viceversa si può dimostrare (ma non lo facciamo) che se A commuta con ogni matrice $C \in \mathbf{GL}_n(K)$, A è una matrice scalare.

ESERCIZI PROPOSTI

4.2.1. Siano V, W due K -spazi vettoriali di dimensione n , con basi rispettivamente \mathbf{E}, \mathbf{F} . Sia $T : V \rightarrow W$ un’applicazione lineare avente matrice A , rispetto alle basi \mathbf{E}, \mathbf{F} .

Verificare che T è un isomorfismo $\iff A \in \mathbf{GL}_n(K)$.

4.2.2. Sia \mathbf{E} la base canonica di \mathbf{R}^3 e sia $\mathbf{F} = (\underline{f}_1 \ \underline{f}_2 \ \underline{f}_3) \in \mathfrak{M}_{1,3}(\mathbf{R}^3)$, con

$$\underline{f}_1 = (1, 0, 1), \quad \underline{f}_2 = (0, 1, -2), \quad \underline{f}_3 = (1, 1, 0).$$

Sia $T : \mathbf{R}^3 \rightarrow \mathbf{R}^3$ un operatore lineare tale che

$$T(\underline{f}_1) = (1, 0, 0), \quad T(\underline{f}_2) = (1, 1, 0), \quad T(\underline{f}_3) = (1, 1, 1).$$

Determinare la matrice A di T rispetto alla base canonica \mathbf{E} .

4.2.3 Siano $S : \mathbf{R}^3 \rightarrow \mathbf{R}^5$ e $T : \mathbf{R}^5 \rightarrow \mathbf{R}^3$ applicazioni lineari così definite:

$$S((a_1, a_2, a_3)) = (0, a_1, a_2, a_3, 0), \quad T((b_1, b_2, b_3, b_4, b_5)) = (b_1, b_3, b_5),$$

$$\forall (a_1, a_2, a_3) \in \mathbf{R}^3, \quad \forall (b_1, b_2, b_3, b_4, b_5) \in \mathbf{R}^5.$$

(i) Scrivere le matrici di $T \circ S$ e di $S \circ T$ rispetto alle basi canoniche \mathbf{E} di \mathbf{R}^3 ed \mathbf{F} di \mathbf{R}^5 .

(ii) Verificare se $\text{Ker}(T \circ S)$ e $\text{Im}(T \circ S)$ sono supplementari in \mathbf{R}^3 e se $\text{Ker}(S \circ T)$ e $\text{Im}(S \circ T)$ lo sono in \mathbf{R}^5 .

4.2.4 È assegnata la matrice $A = \begin{pmatrix} 0 & -1 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 2 \end{pmatrix} \in \mathfrak{M}_{3,4}(\mathbf{R})$. Sia $T : \mathbf{R}^4 \rightarrow \mathbf{R}^3$ l’applicazione lineare definita dalla matrice A , rispetto alle basi canoniche \mathbf{E} di \mathbf{R}^4 ed \mathbf{F} di \mathbf{R}^3 .

lineare definita dalla matrice A , rispetto alle basi canoniche \mathbf{E} di \mathbf{R}^4 ed \mathbf{F} di \mathbf{R}^3 .

Sia poi $S : \mathbf{R}^3 \rightarrow \mathbf{R}^4$ l'applicazione lineare definita (sempre rispetto ad \mathbf{F} ed \mathbf{E}) dalla matrice ${}^t A$.

Determinare la matrice dell'operatore lineare $S \circ T$ di \mathbf{R}^4 (rispetto ad \mathbf{E}) e calcolare dimensioni e basi di $\text{Ker}(S \circ T)$ e $\text{Im}(S \circ T)$.

4.2.5 Siano $V = V_{\mathbf{R}}^3$ e $W = W_{\mathbf{R}}^4$, con basi rispettivamente \mathbf{E} ed \mathbf{F} . Sia $T : V \rightarrow W$ definito rispetto a tali basi dalla matrice

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \in \mathfrak{M}_{4,3}(\mathbf{R}).$$

Sia W_1 il sottospazio vettoriale di W generato dai vettori $\underline{w}_1, \underline{w}_2$, aventi in base \mathbf{F} rispettivamente coordinate $(-1, 0, 1, 2), (1, 1, 0, 1)$.

Determinare una base e la dimensione di $T^{-1}(W_1)$.

4.2.6 Sia $V = V_{\mathbf{R}}^4$ con base \mathbf{E} . Sia $T : V \rightarrow V$ un operatore lineare avente in base \mathbf{E} matrice

$$A = \begin{pmatrix} 1 & 0 & 1 & -1 \\ 1 & 2 & 0 & -1 \\ 0 & -2 & 1 & 0 \\ 2 & 0 & 2 & -2 \end{pmatrix} \in \mathfrak{M}_4(\mathbf{R}).$$

Sia U il sottospazio vettoriale di V generato dai tre vettori $\underline{e}_1 + \underline{e}_4, \underline{e}_2 - \underline{e}_3, \underline{e}_2 - \underline{e}_4$.

(i) Determinare la dimensione ed una base di $T(U)$.

(ii) Determinare equazioni cartesiane di $T(U)$, cioè un SLO in quattro incognite le cui soluzioni sono le coordinate dei vettori di $T(U)$.

4.2.7 Sia T un'applicazione lineare da $V = V_{\mathbf{R}}^3$ a $W = W_{\mathbf{R}}^4$ che, espressa rispetto alle basi $\mathbf{E} = (\underline{e}_1 \ \underline{e}_2 \ \underline{e}_3)$ di V ed $\mathbf{F} = (\underline{f}_1 \ \underline{f}_2 \ \underline{f}_3 \ \underline{f}_4)$ di W , è definita da

$$T(\underline{e}_1) = \underline{f}_1 - \underline{f}_2, \quad T(\underline{e}_2) = \underline{f}_2 - \underline{f}_3, \quad T(\underline{e}_3) = \underline{f}_3 - \underline{f}_4.$$

(i) Scrivere la matrice A e le equazioni di T , rispetto alle basi \mathbf{E}, \mathbf{F} . Calcolare poi equazioni cartesiane e basi di $\text{Ker}(T)$ e di $\text{Im}(T)$.

(ii) Sia $S : W \rightarrow V$ l'applicazione lineare definita dalla matrice $B = {}^t A$, rispetto alle basi \mathbf{F}, \mathbf{E} . Determinare la formula di definizione e le equazioni di S , rispetto alle basi \mathbf{F}, \mathbf{E} . Calcolare poi equazioni cartesiane e basi di $\text{Ker}(S)$ e di $\text{Im}(S)$.

(iii) Determinare matrice ed equazioni di $T \circ S$. Calcolare poi equazioni cartesiane e basi di $\text{Ker}(T \circ S)$ e di $\text{Im}(T \circ S)$.

4.2.8 In \mathbf{R}^3 , con base canonica \mathbf{E} , sono assegnati i due sottospazi vettoriali:

U , rappresentato dal $SLO(1, 3, \mathbf{R})$ $\{x_1 - x_2 = 0\}$;

W , rappresentato dal $SLO(1, 3, \mathbf{R})$ $\{x_2 - x_3 = 0\}$.

Determinare un isomorfismo $T : \mathbf{R}^3 \rightarrow \mathbf{R}^3$ tale che $T(U) = W$. Esprimere T in base \mathbf{E} .

4.2.9. Sia $V = (\mathbf{Z}_2)^2 [= \mathbf{Z}_2 \times \mathbf{Z}_2]$.

(i) Determinare la cardinalità dell'anello $\mathcal{E}nd(V) [= \mathcal{E}nd_{\mathbf{Z}_2}(V)]$ degli operatori lineari di V .

(ii) Determinare il gruppo $\mathcal{U}(\mathcal{E}nd(V))$ degli elementi invertibili di $\mathcal{E}nd(V)$.

4.2.10. Siano \mathbf{E} ed \mathbf{F} due basi di $V = V_K^n$, tali che $\mathbf{F} = \mathbf{E}C$, con $C \in \mathbf{GL}_n(K)$. Sia T un operatore lineare di V definito da $T(\mathbf{E}) = \mathbf{F}A$. Determinare la matrice di T rispetto:

- alla base \mathbf{E} sia nello spazio vettoriale di partenza che in quello di arrivo,
- alla base \mathbf{F} sia nello spazio vettoriale di partenza che in quello di arrivo,
- alla base \mathbf{F} nello spazio vettoriale di partenza ed alla base \mathbf{E} in quello di arrivo.

4.2.11. Sono assegnati i due spazi vettoriali

$V = V_K^n$, con basi \mathbf{E} ed \mathbf{E}' tali che $\mathbf{E}' = \mathbf{E}C$;

$W = W_K^m$, con basi \mathbf{F} ed \mathbf{F}' tali che $\mathbf{F}' = \mathbf{F}D$.

Sia $T : V \rightarrow W$ un'applicazione lineare avente matrice $A \in \mathfrak{M}_{m,n}(K)$, rispetto alle basi \mathbf{E} ed \mathbf{F} . Qual'è la matrice di T rispetto alle basi \mathbf{E}' ed \mathbf{F}' ?

3. Diagonalizzazione di operatori lineari

Nel precedente paragrafo abbiamo visto come sia possibile rappresentare un operatore lineare T di $V = V_K^n$ tramite una matrice. Tale matrice non è unica, ma varia con la scelta della base di V .

Qual'è la base migliore rispetto a cui rappresentare T ? Intuitivamente si tratta di scegliere una base rispetto a cui la matrice di T sia la più semplice possibile. Un obiettivo ambizioso è sperare che la matrice possa essere diagonale.

Se, rispetto ad una base $\mathbf{F} = (\underline{f}_1 \ \dots \ \underline{f}_n)$ di V , T è rappresentato da una matrice diagonale

$$D = \begin{pmatrix} c_1 & 0 & 0 & \dots & 0 \\ 0 & c_2 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & c_n \end{pmatrix},$$

risulta:

$$T(\underline{f}_1) = c_1 \underline{f}_1, \quad T(\underline{f}_2) = c_2 \underline{f}_2, \quad \dots, \quad T(\underline{f}_n) = c_n \underline{f}_n.$$

Diciamo subito che una tale base (e quindi una rappresentazione di T con matrice diagonale) in generale non esiste. Ci occuperemo allora del problema di stabilire in quali occasioni una tale base esista e di come fare per ottenerla.

Partiamo quindi dallo studio di vettori $\underline{v} \in V$ che si comportano come i vettori \underline{f}_i , cioè che sono proporzionali alla propria immagine. Sono detti *autovettori* di T . Ecco la definizione.

Definizione 1. Sia $T : V \rightarrow V$ un operatore lineare e sia $\underline{v} \in V$ un vettore non nullo. Diciamo che \underline{v} è un *autovettore* di T se esiste uno scalare $\lambda \in K$ tale che

$$T(\underline{v}) = \lambda \underline{v}.$$

Lo scalare λ è detto *autovalore* di T associato a \underline{v} .

Analogamente, un elemento $\lambda \in K$ è detto *autovalore* di T se esiste un vettore non nullo $\underline{v} \in V$ tale che $T(\underline{v}) = \lambda \underline{v}$. In tal caso \underline{v} è detto *autovettore* di T associato a λ .

L'insieme degli autovalori di T [che potrebbe anche essere vuoto] è chiamato *spettro* di T ed è denotato $\Lambda(T)$.

[L'uso della lettera λ per indicare un autovalore proviene dalla Fisica ed è universalmente accettato].

Osservazione 1. (i) La richiesta che un autovettore sia un vettore non nullo è fatta per non banalizzare il concetto di autovalore. Se infatti accettassimo come autovettore anche $\underline{0}$, ogni scalare $\lambda \in K$ sarebbe un autovalore di T [in quanto $T(\underline{0}) = \underline{0} = \lambda \underline{0}$, $\forall \lambda \in K$].

(ii) Le definizioni di autovettore e di autovalore sono intrinseche, cioè non legate alla scelta di una base. In effetti queste definizioni valgono per ogni spazio vettoriale, anche di dimensione non finita.

(iii) Si verifica subito che l'autovalore associato ad un autovettore \underline{v} è unico. Se infatti $T(\underline{v}) = \lambda \underline{v} = \mu \underline{v}$, allora $(\lambda - \mu) \underline{v} = \underline{0}$. Poiché $\underline{v} \neq \underline{0}$, allora $\lambda - \mu = 0$, cioè $\lambda = \mu$.

Invece di autovettori associati ad un autovalore λ ce ne sono tanti. Formano in effetti [se aggiungiamo il vettore nullo] un sottospazio vettoriale di V . Proviamo questo fatto nella proposizione che segue.

Proposizione 1. Sia $T : V \rightarrow V$ un operatore lineare e sia $\lambda \in K$ un suo autovalore. L'insieme

$$\{\underline{v} \in V \mid T(\underline{v}) = \lambda \underline{v}\}$$

è un sottospazio vettoriale di V . Tale sottospazio coincide con il nucleo dell'operatore lineare $T - \lambda \mathbf{1}_V$.

Dim. È sufficiente verificare soltanto l'ultima affermazione. L'operatore lineare $T - \lambda \mathbf{1}_V : V \rightarrow V$ è ovviamente così definito:

$$(T - \lambda \mathbf{1}_V)(\underline{v}) = T(\underline{v}) - \lambda \underline{v}, \quad \forall \underline{v} \in V.$$

Pertanto

$$Ker(T - \lambda \mathbf{1}_V) = \{\underline{v} \mid T(\underline{v}) - \lambda \underline{v} = \underline{0}\} = \{\underline{v} \in V \mid T(\underline{v}) = \lambda \underline{v}\}.$$

Definizione 2. Sia $T : V \rightarrow V$ un operatore lineare e sia $\lambda \in K$ un suo autovalore. Il sottospazio vettoriale $Ker(T - \lambda \mathbf{1}_V)$ è detto *autospazio di T associato all'autovalore λ* e viene denotato $\mathbf{E}_\lambda(T)$ (ovvero, più brevemente, \mathbf{E}_λ). [L'iniziale \mathbf{E} di \mathbf{E}_λ segue dal fatto che gli autovettori in inglese sono chiamati "eigenvectors" e gli autovalori "eigenvalues"].

La dimensione di $\mathbf{E}_\lambda(T)$ viene chiamata *multiplicità geometrica di λ* (come autovalore di T) e denotata d_λ .

Osservazione 2. Sia T un operatore lineare di V . Come troviamo i suoi autovalori? Per ogni $c \in K$ possiamo considerare l'operatore lineare $T - c\mathbf{1}_V$. Risulta:

$$c \text{ è un autovalore di } T \iff Ker(T - c\mathbf{1}_V) \neq \{\underline{0}\}.$$

$$\text{Infatti } c \text{ è un autovalore di } T \iff \exists \underline{v} \neq \underline{0} \mid T(\underline{v}) = c\underline{v} \iff \exists \underline{v} \neq \underline{0} \mid \underline{v} \in Ker(T - c\mathbf{1}_V).$$

Esempi. (i) Per ogni $c \in K$, l'operatore $c\mathbf{1}_V$ ammette soltanto l'autovalore c , con autospazio $\mathbf{E}_c = V$. Infatti, $\forall \underline{v} \in V$, $(c\mathbf{1}_V)(\underline{v}) = c\underline{v}$. Invece, per ogni $d \neq c$, $c\mathbf{1}_V - d\mathbf{1}_V = (c-d)\mathbf{1}_V$ è un automorfismo e quindi $Ker(c\mathbf{1}_V - d\mathbf{1}_V) = \{\underline{0}\}$. Dunque d non è autovalore di $c\mathbf{1}_V$.

(ii) Esistono operatori lineari privi di autovalori. Consideriamo infatti l'operatore lineare T di \mathbf{R}^2 definito, rispetto alla base canonica di \mathbf{R}^2 , dalla matrice

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

Assumiamo che, posto $\underline{v} = (x, y)$, risulti, per qualche $\lambda \in \mathbf{R}$, $T(\underline{v}) = \lambda \underline{v}$. Questa uguaglianza vettoriale si traduce nell'uguaglianza matriciale

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \lambda \begin{pmatrix} x \\ y \end{pmatrix},$$

cioè nel $SLO(2, 2, \mathbf{R})$

$$\begin{cases} y = \lambda x \\ -x = \lambda y, \end{cases} \text{ ovvero } \begin{cases} -\lambda x + y = 0 \\ -x - \lambda y = 0. \end{cases}$$

La matrice di tale SLO ha determinante $\lambda^2 + 1$. Tale numero reale non può mai essere nullo (essendo somma di quadrati e dunque positivo). Ne segue che il SLO è privo di autosoluzioni e dunque che nessun $\lambda \in \mathbf{R}$ ammette autovettori associati. Pertanto T è privo di autovalori.

(iii) Ora saliamo un po' rispetto al nostro standard, per dare un'idea dell'importanza del concetto di autovettore.

Denotiamo con \mathfrak{C} (o \mathfrak{C}_I) l'insieme delle funzioni reali definite su un intervallo aperto $I \subseteq \mathbf{R}$ e dotate di derivata di ogni ordine. Si può verificare con facilità che \mathfrak{C} è un sottospazio vettoriale di \mathfrak{F}_I [\mathbf{R} -spazio vettoriale di tutte le funzioni reali definite su I (cfr. **Osserv. 2(ii)** di **Cap. 2.5**)].

Sia ora $D : \mathfrak{C} \rightarrow \mathfrak{C}$ l'operatore di derivazione, che associa ad ogni funzione $f = f(x) \in \mathfrak{C}$, la sua derivata prima $D(f)$ [che si tratti di un operatore lineare è ben noto al lettore; infatti - come si impara nei corsi di Calcolo - vale la proprietà di "linearità" della derivata]:

$$D(a f + b g) = a D(f) + b D(g), \quad \forall a, b \in \mathbf{R}, \quad \forall f, g \in \mathfrak{C}.$$

Cosa significa che f è un autovettore associato ad un autovalore λ ? Ovviamente che risulta $D(f) = \lambda f$, ovvero che f è una soluzione dell'equazione differenziale del primo ordine

$$y' = \lambda y.$$

Dunque risolvere un'equazione differenziale (del primo ordine) significa determinare un'autospazio di D . Si noti poi che ogni $\lambda \in \mathbf{R}$ è un autovalore di D . Infatti $f(x) = e^{\lambda x} \in \mathbf{E}_\lambda(D)$ [in quanto, come noto, $D(e^{\lambda x}) = \lambda e^{\lambda x}$].

Ora ci muoviamo alla ricerca di quanti più autovettori di T linearmente indipendenti sia possibile

determinare in V . La speranza sarebbe quella di costruire una base di autovettori di T .

Proposizione 2. Sia $T : V \rightarrow V$ un operatore lineare e siano $\lambda_1, \dots, \lambda_s$ suoi autovalori a due a due distinti. Siano $\underline{v}_1, \dots, \underline{v}_s$ rispettivi autovettori associati. Tali autovettori sono linearmente indipendenti.

Dim. Si procede per induzione sul numero s degli autovalori.

Se $s = 1$, abbiamo un unico autovettore. Essendo un autovettore, è un vettore non nullo e quindi è un vettore linearmente indipendente (cfr. **Osserv. 1(i)** di **Cap. 2.5**).

Sia $s > 1$ e si supponga il risultato vero per $s-1$ autovalori: dunque $\underline{v}_1, \dots, \underline{v}_{s-1}$ sono linearmente indipendenti. Proviamo che anche $\underline{v}_1, \dots, \underline{v}_s$ sono linearmente indipendenti. Sia

$$\sum_{i=1}^s c_i \underline{v}_i = \underline{0}.$$

Allora

$$T\left(\sum_{i=1}^s c_i \underline{v}_i\right) = \begin{cases} = \sum_{i=1}^s c_i T(\underline{v}_i) = \sum_{i=1}^s c_i \lambda_i \underline{v}_i \\ = T(\underline{0}) = \lambda_s \underline{0} = \lambda_s \left(\sum_{i=1}^s c_i \underline{v}_i\right) = \sum_{i=1}^s c_i \lambda_s \underline{v}_i \end{cases}$$

e quindi

$$\sum_{i=1}^s c_i \lambda_i \underline{v}_i = \sum_{i=1}^s c_i \lambda_s \underline{v}_i.$$

Cancellando l'ultimo addendo di queste due sommatorie [che è lo stesso vettore $c_s \lambda_s \underline{v}_s$] si ottiene

$$\sum_{i=1}^{s-1} c_i \lambda_i \underline{v}_i = \sum_{i=1}^{s-1} c_i \lambda_s \underline{v}_i, \text{ cioè } \sum_{i=1}^{s-1} c_i (\lambda_i - \lambda_s) \underline{v}_i = \underline{0}.$$

Poiché $\underline{v}_1, \dots, \underline{v}_{s-1}$ sono linearmente indipendenti,

$$c_1 (\lambda_1 - \lambda_s) = 0, \dots, c_{s-1} (\lambda_{s-1} - \lambda_s) = 0$$

e, poiché $\lambda_1 - \lambda_s, \dots, \lambda_{s-1} - \lambda_s$ sono non nulli, allora $c_1 = \dots = c_{s-1} = 0$. Dalla relazione $\sum_{i=1}^s c_i \underline{v}_i = \underline{0}$ segue allora che anche $c_s = 0$. Si conclude che $\underline{v}_1, \dots, \underline{v}_s$ sono linearmente indipendenti.

La proposizione precedente può essere migliorata nella seguente direzione: invece di considerare un solo autovettore in ogni autospazio \mathbf{E}_{λ_i} , consideriamo più autovettori linearmente indipendenti in ciascun \mathbf{E}_{λ_i} . Risulterà che questi autovettori (tutti insieme) sono linearmente indipendenti. Quanti autovettori linearmente indipendenti potremo poi scegliere in \mathbf{E}_{λ_i} ? Ovviamente al più d_{λ_i} (dimensione dell'autospazio). Ecco il risultato in questione.

Proposizione 3. Sia $T : V \rightarrow V$ un operatore lineare e siano $\lambda_1, \dots, \lambda_s$ suoi autovalori a due a due distinti. Per $i = 1, \dots, s$, si scelgano nell'autospazio $\mathbf{E}_{\lambda_i}(T)$ n_i autovettori linearmente indipendenti $\underline{v}_{i,1}, \dots, \underline{v}_{i,n_i}$ [con $1 \leq n_i \leq d_{\lambda_i}$]. Risulta: i vettori

$$\underline{v}_{11}, \dots, \underline{v}_{1n_1}, \underline{v}_{21}, \dots, \underline{v}_{2n_2}, \dots, \underline{v}_{s1}, \dots, \underline{v}_{sn_s}$$

sono linearmente indipendenti.

Dim. Bisogna verificare che

$$\sum_{i=1}^s \sum_{j=1}^{n_i} c_{ij} \underline{v}_{ij} = \underline{0} \implies c_{ij} = 0, \forall i = 1, \dots, s, \forall j = 1, \dots, n_i.$$

Poniamo $\underline{u}_i := \sum_{j=1}^{n_i} c_{ij} \underline{v}_{ij}$. Allora $\underline{u}_1 + \dots + \underline{u}_s = \underline{0}$ e quindi $\underline{u}_1, \dots, \underline{u}_s$ sono linearmente dipendenti.

Se $\underline{u}_i = \underline{0}$, allora (essendo $\underline{v}_{i1}, \dots, \underline{v}_{in_i}$ linearmente indipendenti) $c_{i1} = \dots = c_{in_i} = 0$. Pertanto, se ogni $\underline{u}_i = \underline{0}$, la tesi è dimostrata.

Se invece, per assurdo, esistesse qualche $\underline{u}_i \neq \underline{0}$, si avrebbe:

$$T(\underline{u}_i) = T\left(\sum_{j=1}^{n_i} c_{ij} \underline{v}_{ij}\right) = \sum_{j=1}^{n_i} c_{ij} T(\underline{v}_{ij}) = \sum_{j=1}^{n_i} c_{ij} \lambda_i \underline{v}_{ij} = \lambda_i \underline{u}_i$$

e quindi \underline{u}_i sarebbe un autovettore associato a λ_i . Tali \underline{u}_i non nulli, in base alla proposizione precedente sarebbero linearmente indipendenti. Ma la loro somma è $\underline{0}$ e dunque sono linearmente dipendenti: assurdo.

Sia ora $V = V_K^n$ uno spazio vettoriale n -dimensionale e sia T un operatore lineare di V .

La **Prop. 2** ci garantisce che T possiede al più n autovalori distinti, cioè che $|\Lambda(T)| \leq n$. Se infatti per assurdo T avesse $n+1$ autovalori distinti, avrebbe in corrispondenza $n+1$ vettori linearmente indipendenti [i rispettivi autovettori].

La **Prop. 3** ci dice qualcosa di più:

$$\sum_{\lambda \in \Lambda(T)} d_\lambda \leq n.$$

Infatti, se per ogni $\lambda \in \Lambda(T)$ scegliamo d_λ autovettori linearmente indipendenti, questi vettori (tutti insieme) sono linearmente indipendenti e dunque il loro numero complessivo è $\leq n$.

Se in particolare $\sum_{\lambda \in \Lambda(T)} d_\lambda = n$, tali vettori formano una base di V . Dunque V ammette una base di autovettori di T .

Vale anche il viceversa: assumiamo che V ammetta una base di autovettori di T . Tale base deve contenere autovettori di ogni autospazio di T [in base alla **Prop. 3**]. Se $\Lambda(T) = \{\lambda_1, \dots, \lambda_s\}$, i vettori di tale base si ripartiscono quindi in questo modo:

$$\underline{v}_{1,1}, \dots, \underline{v}_{1,n_1} \in \mathbf{E}_{\lambda_1}(T), \quad \underline{v}_{2,1}, \dots, \underline{v}_{2,n_2} \in \mathbf{E}_{\lambda_2}(T), \quad \dots, \quad \underline{v}_{s,1}, \dots, \underline{v}_{s,n_s} \in \mathbf{E}_{\lambda_s}(T),$$

con $n_1 + n_2 + \dots + n_s = n$. Poiché ogni $n_i \leq d_{\lambda_i}$, si ha:

$$n = \sum_{i=1}^s n_i \leq \sum_{i=1}^s d_{\lambda_i} = \sum_{\lambda \in \Lambda(T)} d_\lambda \leq n \quad \text{e dunque} \quad \sum_{\lambda \in \Lambda(T)} d_\lambda = n.$$

Abbiamo così dimostrato il seguente risultato.

Proposizione 4. Sia $V = V_K^n$ e T un operatore lineare di V . Risulta:

$$\text{esiste in } V \text{ una base di autovettori di } T \iff \sum_{\lambda \in \Lambda(T)} d_\lambda = n.$$

Definizione 3. Un operatore lineare T di $V = V_K^n$ è detto *diagonalizzabile* se V ammette una base di autovettori di T . Una tale base di autovettori è detta *base diagonalizzante* (per T).

Se quindi T è diagonalizzabile ed ammette una base di autovettori \mathbf{F} , la sua matrice in base \mathbf{F} è una matrice diagonale e sulla diagonale compaiono tutti gli autovalori di T , ciascuno un numero di volte pari alla rispettiva molteplicità geometrica.

Osservazione 3. (i) Un operatore lineare di $V = V_K^n$ che ammetta n autovalori distinti è certamente diagonalizzabile [infatti $\sum_{i=1}^n d_{\lambda_i} = n$ (e ogni $d_{\lambda_i} = 1$)]. Ovviamente esistono operatori lineari diagonalizzabili che posseggono meno di n autovalori distinti: ad esempio sia $T : \mathbf{R}^3 \rightarrow \mathbf{R}^3$ avente (rispetto alla base canonica \mathbf{E} di \mathbf{R}^3) matrice

$$B = \begin{pmatrix} a & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & b \end{pmatrix}, \quad \text{con } a \neq b.$$

Tale operatore ha due soli autovalori: a (con $d_a = 2$) e b (con $d_b = 1$). Si noti che, essendo B diagonale, una base di autovettori di T è proprio \mathbf{E} . In particolare $\mathbf{E}_a(T) = \langle \underline{e}_1, \underline{e}_2 \rangle$ e $\mathbf{E}_b(T) = \langle \underline{e}_3 \rangle$.

(ii) Ovviamente un operatore lineare privo di autovalori [ad esempio quello di **Esempi(ii)**] non è diagonalizzabile. Esistono comunque operatori lineari non diagonalizzabili che posseggono qualche

autovalore. Eccone un esempio: Si consideri l'operatore lineare $T : \mathbf{R}^2 \rightarrow \mathbf{R}^2$ avente (rispetto alla base canonica \mathbf{E} di \mathbf{R}^2) matrice

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

Sia $T(\underline{v}) = \lambda \underline{v}$, con $\lambda \in \mathbf{R}$. Posto $\underline{v} = (x, y)$, tale uguaglianza si scrive nella forma

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \lambda \begin{pmatrix} x \\ y \end{pmatrix},$$

che si traduce nel $SLO(2, 2, \mathbf{R})$

$$\begin{cases} (1 - \lambda)x + y = 0 \\ (1 - \lambda)y = 0. \end{cases}$$

La matrice di tale SLO è

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

Il rango di tale matrice è 2 se $\lambda \neq 1$, ed è 1 se $\lambda = 1$. Ne segue che il SLO è privo di autosoluzioni se $\lambda \neq 1$, ed ha ∞^1 soluzioni se $\lambda = 1$. Tali soluzioni sono proporzionali ad \underline{e}_1 .

Pertanto T ammette il solo autovalore 1 con molteplicità geometrica $d_1 = 1$ [e con $\mathbf{E}_1(T) = \langle \underline{e}_1 \rangle$]. Dunque T è non è diagonalizzabile.

Ci poniamo ora il problema di determinare tutti gli autovalori di un operatore lineare T , a partire da una sua matrice.

Sia $V = V_K^n$ e sia T un operatore lineare di V . Sia \mathbf{E} una base di V e sia $A \in \mathfrak{M}_n(K)$ la matrice di T rispetto a tale base. Sia $\lambda \in K$. Dall'**Osserv. 2**,

$$\lambda \in \Lambda(T) \iff \text{Ker}(T - \lambda \mathbf{1}_V) \neq \{\underline{0}\}.$$

Osserviamo subito che $\lambda \mathbf{1}_V$ ha matrice scalare λI_n in base \mathbf{E} [in effetti $\lambda \mathbf{1}_V$ ha sempre matrice λI_n , rispetto ad ogni base di V , in quanto $\lambda \mathbf{1}_V(\underline{v}) = \lambda \underline{v}, \forall \underline{v} \in V$].

Segue dalla **Prop. 3** del precedente paragrafo che $T - \lambda \mathbf{1}_V$ ha matrice (in base \mathbf{E}) $A - \lambda I_n$ e quindi che $\text{Ker}(T - \lambda \mathbf{1}_V)$ ha equazioni cartesiane (in base \mathbf{E}) date dal $SLO(n, n, K)$

$$(A - \lambda I_n) X = \mathbf{0}.$$

Tale SLO ammette autosoluzioni $\iff rg(A - \lambda I_n) < n \iff \det(A - \lambda I_n) = 0$.

Interpretiamo ora λ come un'incognita. Allora

$$\det(A - \lambda I_n) = \begin{vmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} - \lambda \end{vmatrix}$$

[opportunamente sviluppato secondo la definizione di determinante] è un polinomio nell'incognita λ ed a coefficienti in K . Tale polinomio viene chiamato *polinomio caratteristico di A* e viene denotato con P o P_A . Abbiamo provato il seguente fatto.

Proposizione 5. Sia T un operatore lineare di $V = V_K^n$, avente matrice A (rispetto ad una base \mathbf{E} di V) e polinomio caratteristico P . Sia $\lambda \in K$. Risulta:

$$\lambda \in \Lambda(T) \iff P(\lambda) = 0.$$

In altri termini, gli autovalori di T sono tutti e soli gli zeri del polinomio caratteristico di T .

Se lo studente ritorna agli esempi numerici di questo paragrafo, verificherà che:

$T : \mathbf{R}^2 \rightarrow \mathbf{R}^2$, di matrice A , ha polinomio caratteristico $P_A = \lambda^2 + 1$ (privi di zeri in \mathbf{R});

$T : \mathbf{R}^3 \rightarrow \mathbf{R}^3$, di matrice B , ha polinomio caratteristico $P_B = (a - \lambda)^2(b - \lambda)$ (con zeri $a, b \in \mathbf{R}$);

$T : \mathbf{R}^2 \rightarrow \mathbf{R}^2$, di matrice C , ha polinomio caratteristico $P_C = (1 - \lambda)^2$ (con unico zero $1 \in \mathbf{R}$).

Proveremo ora che il polinomio caratteristico è un *invariante di T*, cioè non dipende dalla base

scelta per definire la matrice di T . Ciò in qualche misura è un fatto inaspettato, visto che abbiamo usato A per definirlo. Alla luce di questo risultato, parleremo di polinomio caratteristico di T (e non più di A) e lo denoteremo con P_T (piuttosto che con P_A).

Proposizione 6. *Sia T un operatore lineare di $V = V_K^n$. Il suo polinomio caratteristico è un invariante di T , cioè è indipendente dalla base \mathbf{E} di V rispetto a cui viene rappresentato T .*

Dim. Assumiamo che \mathbf{E}, \mathbf{F} siano due basi di V , collegate dalla formula di cambiamento di base $\mathbf{F} = \mathbf{E}C$, con $C \in \mathbf{GL}_n(K)$. Supponiamo che, rispetto alla base \mathbf{E} , T abbia matrice A e che, rispetto alla base \mathbf{F} , T abbia matrice B . Abbiamo provato, alla fine del precedente paragrafo, che risulta $B = C^{-1}AC$. Dobbiamo ora provare che

$$\det(A - \lambda I_n) = \det(B - \lambda I_n).$$

Ricordiamo (cfr. l'**Osserv. 3** del paragrafo precedente) che ogni matrice scalare λI_n commuta con ogni matrice di $\mathfrak{M}_n(K)$. Si ha quindi, usando le proprietà distributive a destra e sinistra del prodotto righe per colonne:

$$\begin{aligned} B - \lambda I_n &= (C^{-1}AC) - \lambda I_n = (C^{-1}AC) - (C^{-1}C)\lambda I_n = (C^{-1}AC) - C^{-1}(C\lambda I_n) = \\ &= (C^{-1}AC) - C^{-1}(\lambda I_n C) = (C^{-1}AC) - (C^{-1}\lambda I_n C) = C^{-1}(A - \lambda I_n)C. \end{aligned}$$

In base al teorema di Binet ed alle proprietà del determinante,

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

Si conclude che $\det(B - \lambda I_n) = \det(A - \lambda I_n)$.

Cosa possiamo dire a proposito del polinomio caratteristico? Innanzitutto il suo grado è n . Infatti, considerato lo sviluppo del determinante $|A - \lambda I_n|$, si osserva che uno degli addendi è dato da

$$\prod_{i=1}^n (a_{ii} - \lambda) = (-1)^n \lambda^n + \dots.$$

Tutti gli altri addendi del determinante sono formati da fattori di cui almeno due si trovano al di fuori della diagonale. Per questo motivo il loro grado in λ è $\leq n-2$.

Il coefficiente direttore del polinomio caratteristico è quindi $(-1)^n$, mentre il coefficiente del termine di grado $n-1$ [essendo ottenuto dal solo primo addendo] è

$$(-1)^{n-1}(a_{11} + a_{22} + \dots + a_{nn}).$$

La somma degli elementi della diagonale di A è chiamata *traccia di A* e denotata $Tr(A)$.

Infine il termine noto è $\det(A)$. Infatti il termine noto di un qualsiasi polinomio coincide con il valore che il polinomio assume in 0 ed ovviamente $P_T(0) = |A - 0I_n| = \det(A)$. Abbiamo pertanto

$$P_T = (-1)^n \lambda^n + (-1)^{n-1} Tr(A) \lambda^{n-1} + \dots + \det(A).$$

[Non abbiamo dato informazioni sugli altri coefficienti di P_T , anche se non sarebbe difficile trovarne].

Si noti che, essendo il polinomio caratteristico un invariante di T , anche tutti i suoi coefficienti sono invarianti di T . In particolare quindi lo sono la traccia ed il determinante.

Un esempio. Vogliamo verificare che l'operatore lineare $T : \mathbf{R}^3 \rightarrow \mathbf{R}^3$ definito rispetto alla base canonica \mathbf{E} di \mathbf{R}^3 dalla matrice

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

è diagonalizzabile, calcolandone prima il polinomio caratteristico, quindi gli autovalori ed infine una base di ciascun autospazio.

Risulta:

$$\begin{aligned} P_T &= \begin{vmatrix} \frac{5}{3} - \lambda & -\frac{2}{3} & \frac{2}{3} \\ -\frac{1}{3} & \frac{4}{3} - \lambda & -\frac{1}{3} \\ 0 & 0 & 1 - \lambda \end{vmatrix} = (1 - \lambda) \left(\left(\frac{5}{3} - \lambda \right) \left(\frac{4}{3} - \lambda \right) - \frac{2}{9} \right) = (1 - \lambda) (\lambda^2 - 3\lambda + 2) = \\ &= (1 - \lambda)(\lambda - 1)(\lambda - 2) = -(\lambda - 1)^2(\lambda - 2) = -\lambda^3 + 4\lambda^2 - 5\lambda + 2. \end{aligned}$$

[Si noti che $Tr(A) = 4$, $\det(A) = 2$]. Si verifica subito che T ammette gli autovalori 1, 2.

Determiniamo una base dell'autospazio $\mathbf{E}_1(T)$. Si ha

$$A - I_3 = \begin{pmatrix} \frac{2}{3} & -\frac{2}{3} & \frac{2}{3} \\ -\frac{1}{3} & \frac{1}{3} & -\frac{1}{3} \\ 0 & 0 & 0 \end{pmatrix}.$$

Tale matrice ha rango 1 ed il corrispondente SLO si riduce al $SLO(1, 3, \mathbf{R})$

$$\{x - y + z = 0\}.$$

Tale SLO ammette ∞^2 soluzioni, generate ad esempio dai due vettori $(1, 1, 0), (1, 0, -1)$. Pertanto

$$\mathbf{E}_1(T) = \langle (1, 1, 0), (1, 0, -1) \rangle.$$

Determiniamo ora una base dell'autospazio $\mathbf{E}_2(T)$. Si ha

$$A - 2I_3 = \begin{pmatrix} -\frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \\ -\frac{1}{3} & -\frac{2}{3} & -\frac{1}{3} \\ 0 & 0 & -1 \end{pmatrix}.$$

Tale matrice ha rango 2 e ad esempio le ultime due righe sono linearmente indipendenti. Il corrispondente SLO si riduce al $SLO(2, 3, \mathbf{R})$

$$\begin{cases} -\frac{1}{3}x - \frac{2}{3}y - z = 0 \\ -z = 0, \end{cases} \quad \text{che è equivalente a} \quad \begin{cases} x + 2y = 0 \\ z = 0. \end{cases}$$

Tale SLO ha ∞^1 soluzioni, proporzionali a $(2, -1, 0)$. Pertanto

$$\mathbf{E}_2(T) = \langle (2, -1, 0) \rangle.$$

Una base di autovettori di T è quindi

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

e rispetto a tale base T ha matrice diagonale

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

[Verifica: posto $C = \begin{pmatrix} 1 & 1 & 2 \\ 1 & 0 & -1 \\ 0 & -1 & 0 \end{pmatrix}$, deve risultare $C^{-1}AC = D$].

Riprendiamo le nostre considerazioni sul polinomio caratteristico P_T . Se $\lambda_0 \in \Lambda(T)$, il polinomio $\lambda - \lambda_0$ è un fattore di P_T . Può avvenire che anche $(\lambda - \lambda_0)^2, (\lambda - \lambda_0)^3, \dots, (\lambda - \lambda_0)^h$ siano fattori di P_T , mentre $(\lambda - \lambda_0)^{h+1}$ non è fattore di P_T . Diremo in tal caso che λ_0 ha *moltiplicità algebrica* h , (*come zero di* P_T).

Nell'ultimo esempio sopra considerato, l'autovalore 1 ha molteplicità algebrica 2, mentre l'autovalore 2 ha molteplicità algebrica 1.

Formalizziamo il concetto di molteplicità algebrica.

Definizione 4. Sia P_T il polinomio caratteristico di un operatore lineare T di $V = V_K^n$. Sia $\lambda_0 \in \Lambda(T)$. Diciamo che λ_0 ha molteplicità algebrica $h = h_{\lambda_0}$ se $(\lambda - \lambda_0)^h$ è un fattore di P_T , mentre $(\lambda - \lambda_0)^{h+1}$ non è un fattore di P_T . [In altri termini, h è il massimo esponente i tale che $(\lambda - \lambda_0)^i$ è un fattore di P_T].

Si noti che $n \geq \sum h_\lambda$ e che tale diseguaglianza è stretta $\iff P_T$ possiede almeno un fattore irriducibile di grado ≥ 2 .

Confrontiamo ora la molteplicità algebrica con quella geometrica.

Proposizione 7. Sia T un operatore lineare di $V = V_K^n$. Sia $\lambda_0 \in \Lambda(T)$. La molteplicità geometrica d_{λ_0} è minore o uguale alla molteplicità algebrica h_{λ_0} .

Dim. Per comodità di scrittura scriviamo h_0 in luogo di h_{λ_0} e d_0 in luogo di d_{λ_0} . Scegliamo una base $\{\underline{v}_1, \dots, \underline{v}_{d_0}\}$ di $\mathbf{E}_{\lambda_0}(T)$ e completiamola ad una base di V . Sia \mathbf{F} la base ottenuta. In base \mathbf{F} la matrice di T è la seguente:

$$\begin{pmatrix} \lambda_0 & 0 & \dots & 0 & \dots & \dots \\ 0 & \lambda_0 & \dots & 0 & \dots & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \lambda_0 & \dots & \dots \\ 0 & 0 & \vdots & 0 & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \vdots & 0 & \vdots & \vdots \end{pmatrix}$$

[si noti che relativamente alle ultime $n - d_0$ colonne di tale matrice non abbiamo alcuna informazione]. Il polinomio caratteristico di T è quindi

$$P_T = (\lambda_0 - \lambda)^{d_0} Q$$

dove Q è un polinomio in λ di grado $n - d_0$ [ricavato dal calcolo del determinante della precedente matrice]. Poiché $(\lambda_0 - \lambda)^{d_0}$ è un fattore di P_T , allora $h_0 \geq d_0$.

Osservazione 4. Se $h_\lambda > d_\lambda$ (anche per un solo autovalore λ) l'operatore T non può essere diagonalizzabile [in quanto $\sum_{\Lambda \in \lambda(T)} d_\lambda < \sum_{\Lambda \in \lambda(T)} h_\lambda \leq n$].

Analogamente, T non può essere diagonalizzabile se il polinomio caratteristico P_T ha un fattore irriducibile di grado ≥ 2 [infatti in questo caso $\sum_{\Lambda \in \lambda(T)} d_\lambda \leq \sum_{\Lambda \in \lambda(T)} h_\lambda < n$].

Concludiamo con l'esempio di un operatore lineare avente una molteplicità geometrica strettamente inferiore alla corrispondente molteplicità algebrica ed anche un fattore irriducibile di grado ≥ 2 nel polinomio caratteristico.

Sia $T : \mathbf{R}^5 \rightarrow \mathbf{R}^5$ definito rispetto alla base canonica \mathbf{E} di \mathbf{R}^5 dalla matrice

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

Il polinomio caratteristico è

$$P_T = (1 - \lambda)^3 (\lambda^2 + 1).$$

Il polinomio $\lambda^2 + 1$ è irriducibile su \mathbf{R} e dunque non produce autovalori [e già per questo motivo T non può essere diagonalizzabile]. L'unico autovalore di T è $\lambda = 1$, con $h_1 = 3$. Ne segue subito che $1 \leq d_1 \leq 3$. Per determinare d_1 calcoliamo una base di $\mathbf{E}_1(T)$.

Tale autospazio è descritto dal $SLO(5, 5, \mathbf{R})$ avente matrice

$$A - I_3 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & -1 & -1 \end{pmatrix}.$$

Tale matrice ha rango 3. Eliminando le due righe nulle, si ottiene il $SLO(3, 5, \mathbf{R})$

$$\begin{cases} x_3 = 0 \\ -x_4 + x_5 = 0 \\ -x_4 - x_5 = 0, \end{cases}$$

il quale ammette ∞^2 soluzioni, generate ad esempio da $\{\underline{e}_1, \underline{e}_2\}$. Si conclude che $\mathbf{E}_1(T) = \langle \underline{e}_1, \underline{e}_2 \rangle$ e $2 = d_1 < h_1 = 3$.

ESERCIZI PROPOSTI

4.3.1. Sia $T : V_K^n \rightarrow V_K^n$ un operatore lineare.

(i) Verificare che, se λ è un autovalore di T , λ^2 è un autovalore di T^2 .

(ii) È vero che, se T è diagonalizzabile, anche T^2 lo è ?

(iii) È vero che, se T^2 è diagonalizzabile, anche T lo è ?

4.3.2. Sia $V = V_K^n$ e T un operatore lineare di V , tale che

$$T \neq \mathbf{1}_V, T \neq \mathbf{0}_V \text{ e } T^2 = T.$$

(i) Verificare che 0, 1 sono autovalori di T .

(ii) Verificare che gli autospazi $\mathbf{E}_0(T)$, $\mathbf{E}_1(T)$ sono sottospazi vettoriali supplementari di V .

4.3.3. Al variare di $a, b \in \mathbf{R}$, sia $T : \mathbf{R}^3 \rightarrow \mathbf{R}^3$ l'operatore lineare avente matrice

$$A = \begin{pmatrix} 0 & a & b \\ -a & 0 & -b \\ -b & b & 0 \end{pmatrix}$$

[rispetto alla base canonica \mathbf{E} di \mathbf{R}^3]. Per quali $a, b \in \mathbf{R}$ l'operatore T è diagonalizzabile ?

4.3.4. Sia $T : \mathbf{R}^3 \rightarrow \mathbf{R}^3$ l'operatore lineare definito, rispetto alla base canonica \mathbf{E} , dalla matrice

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

dipendente da un parametro reale a . Determinare per quali valori di a T è diagonalizzabile.

4.3.5. Sia $T : \mathbf{R}^4 \rightarrow \mathbf{R}^4$ l'operatore lineare definito, rispetto alla base canonica \mathbf{E} , dalla matrice

$$A = \begin{pmatrix} -5 & 0 & 4 & 4 \\ -10 & 5 & 6 & -4 \\ 0 & 0 & 3 & 8 \\ 0 & 0 & 2 & -3 \end{pmatrix}.$$

Determinare gli autovalori di T e, se esiste, una base di autovettori di T .

4.3.6. È assegnata l'applicazione lineare $T : \mathbf{R}^3 \rightarrow \mathbf{R}^2$ definita, rispetto alle basi canoniche dei due spazi vettoriali, dalla matrice

$$A = \begin{pmatrix} 0 & a & 1 \\ 0 & 1 & a \end{pmatrix}, \text{ dipendente da un parametro } a \in \mathbf{R}.$$

Sia $S : \mathbf{R}^2 \rightarrow \mathbf{R}^3$ l'applicazione lineare definita dalla matrice ${}^t A$ (sempre rispetto alle basi canoniche). Determinare per quali eventuali $a \in \mathbf{R}$ l'operatore lineare $S \circ T$ non è diagonalizzabile.

4.3.7. Sia $T : \mathbf{R}^2 \rightarrow \mathbf{R}^2$ l'operatore lineare definito dai seguenti dati:

$$\underline{v} = (1, 2), \quad T(\underline{v}) = (2, 1), \quad T^2(\underline{v}) = (-2, -1).$$

(i) Determinare la matrice A di T rispetto alla base canonica \mathbf{E} di \mathbf{R}^2 .

(ii) Determinare gli autovalori di T e, se esiste, una base di autovettori di T .

4.3.8. Sia $T : \mathbf{R}^3 \rightarrow \mathbf{R}^3$ l'operatore lineare individuato dai seguenti dati:

$$T \text{ è diagonalizzabile, } \Lambda(T) = \{0, 1\}, \quad \mathbf{E}_1(T) = \langle (2, 0, 1) \rangle,$$

$$T(W) \subseteq W, \text{ con } W = \langle (1, 1, 1), (0, 1, -1) \rangle.$$

Determinare la matrice di T rispetto alla base canonica \mathbf{E} di \mathbf{R}^3 .

4.3.9. In $V = V_{\mathbf{R}}^3$ sono assegnate due basi $\mathbf{E} = (\underline{e}_1 \ \underline{e}_2 \ \underline{e}_3)$, $\mathbf{F} = (\underline{f}_1 \ \underline{f}_2 \ \underline{f}_3)$ tali che:

$$\underline{e}_1 = \underline{f}_2, \quad \underline{e}_2 = -\underline{f}_1 + \underline{f}_3, \quad \underline{e}_3 = \underline{f}_1 + \underline{f}_2.$$

Sia $T : V \rightarrow V$ l'operatore lineare tale che

$$T(\underline{f}_1) = \underline{e}_1 - \underline{e}_2, \quad T(\underline{f}_2) = \underline{e}_2 - \underline{e}_3, \quad T(\underline{f}_3) = \underline{e}_3 - \underline{e}_1.$$

(i) Esprimere T in base \mathbf{E} ed in base \mathbf{F} .

(ii) Verificare se T è diagonalizzabile e calcolare una base di ogni autospazio.

(iii) Sia U il sottospazio vettoriale di V definito in base \mathbf{F} dal $SLO(1, 3, \mathbf{R})$ $\{y_1 - y_2 = 0\}$. Determinare una base di $T(U)$ (in base \mathbf{F}).

4.3.10. Sia $T : \mathbf{R}^3 \rightarrow \mathbf{R}^3$ l'operatore lineare definito dai seguenti dati: posto $\underline{x} = (1, 0, -1) \in \mathbf{R}^3$,

$$T(\underline{x}) = (0, 1, 2), \quad T^2(\underline{x}) = (1, 1, 0), \quad T^3(\underline{x}) = (0, 1, 2).$$

- (i) Determinare la matrice A di T rispetto alla base canonica \mathbf{E} di \mathbf{R}^3 .
- (ii) Verificare che T è diagonalizzabile e indicarne gli autovalori.
- (iii) Assegnato il sottospazio $W = \langle T(\underline{x}), T^2(\underline{x}) \rangle$, determinare una base di $T^{-1}(W)$.

4.3.11. Sia $V = V_K^n$ e sia $T : V \rightarrow V$ l'operatore lineare definito, rispetto ad una base \mathbf{E} di V , dalla matrice $A \in \mathfrak{M}_n(K)$. Sia $T' : V \rightarrow V$ l'operatore lineare definito, sempre rispetto ad \mathbf{E} , dalla matrice ${}^t A$ (trasposta di A).

- (i) Verificare che $P_T = P_{T'}$.
- (ii) Verificare che T è diagonalizzabile $\iff T'$ è diagonalizzabile.

4.3.12. Determinare, in una base opportuna, gli operatori lineari $T : \mathbf{R}^3 \rightarrow \mathbf{R}^3$ definiti dalle seguenti tre condizioni:

$$\mathbf{E}_1(T) = \langle (1, 0, -2) \rangle, \quad \mathbf{E}_0(T) = \langle (1, 1, 1) \rangle, \quad h_0 = 2$$

[dove h_0 denota la molteplicità algebrica di 0 come autovalore di T].

4.3.13. Sia T un operatore lineare di $V = V_K$, tale che $T^3 \neq \mathbf{0}$ e $T^4 = \mathbf{0}$ (operatore nullo).

- (i) Verificare che $\Lambda(T) = \{0\}$.
- (ii) Verificare che T non è diagonalizzabile.
- (iii) Determinare un esempio di un siffatto operatore T di \mathbf{R}^4 .

4.3.14. È assegnata la matrice

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

Verificare che esiste un polinomio non nullo $M = M(x) = a_0 + a_1x + a_2x^2 + a_3x^3$, a coefficienti in \mathbf{R} , tale che

$$\sum_{i=0}^3 a_i A^i = \mathbf{0}$$

[dove $\mathbf{0}$ denota la matrice nulla di ordine 3]. Confrontare il polinomio M con il polinomio caratteristico P_A .

Capitolo 5

ELEMENTI DI TEORIA DEI GRUPPI

1. Gruppi ciclici

Sia (G, \cdot) un gruppo [con elemento neutro $1 = 1_G$]. Preso un elemento $x \in G$, consideriamone le *potenze ad esponente intero* x^t ($\forall t \in \mathbf{Z}$), che sono così definite:

$$\begin{aligned} x^0 &= 1, \quad x^1 = x, \quad x^{-1} = (x)^{-1} \quad [\text{cioè } x^{-1} \text{ è l'inverso di } x \text{ in } G], \\ x^t &= \underbrace{x \cdot x \cdot \dots \cdot x}_{t \text{ fattori}}, \quad \text{se } t \geq 2, \quad x^t = \underbrace{x^{-1} \cdot x^{-1} \cdot \dots \cdot x^{-1}}_{-t \text{ fattori}}, \quad \text{se } t \leq -2. \end{aligned}$$

Da tali definizioni segue, $\forall t, s \in \mathbf{Z}$:

$$x^t \cdot x^s = x^{t+s} = x^s \cdot x^t, \quad (x^t)^s = (x^s)^t = x^{ts}$$

[e quindi, in particolare, $(x^{-1})^t = (x^t)^{-1} = x^{-t}$].

Si noti che $1^t = 1$, $\forall t \in \mathbf{Z}$. Per ogni $x \neq 1$, si hanno due eventualità:

$$\exists t > 0 \mid x^t = 1; \quad \forall t > 0, x^t \neq 1.$$

Nel primo caso saremo interessati a considerare il minimo intero positivo t tale che $x^t = 1$. Tale esponente t sarà detto *periodo di* x , come precisato nella seguente definizione.

Definizione 1. Sia (G, \cdot) un gruppo e sia $x \in G$. Diremo che x ha periodo t se t è il minimo intero positivo tale che $x^t = 1$ [cioè $x, x^2, \dots, x^{t-1} \neq 1$, mentre $x^t = 1$]. Scrivremo $\circ(x) = t$.

Diremo invece che x ha periodo infinito se $x^t \neq 1$, $\forall t > 0$, e scrivremo in tal caso $\circ(x) = \infty$.

Ovviamente $\circ(1) = 1$ [infatti $1^1 = 1$]. Viceversa, se $\circ(x) = 1$, allora $x^1 = 1$ e quindi $x = 1$. Dunque 1 è caratterizzato come l'unico elemento di G di periodo 1. Si verifica poi subito che:

$$\circ(x) = 2 \iff x \neq 1 \text{ e } x^{-1} = x.$$

Valgono i seguenti semplici risultati relativi al periodo ed alle potenze di un elemento.

Proposizione 1. Sia (G, \cdot) un gruppo e sia $x \in G$ un elemento di periodo finito t . Se $x^h = 1$, allora $t \mid h$ [cioè h è un multiplo di t]. Si ha poi, $\forall h, k \in \mathbf{Z}$:

$$x^h = x^k \iff h \equiv k \pmod{t}.$$

Dim. Dividiamo h per t :

$$h = tq + r \quad \text{con } q, r \in \mathbf{Z} \quad \text{e } 0 \leq r < t.$$

Va provato che $r = 0$ (da cui $t \mid h$). Si ha:

$$1 = x^h = x^{tq+r} = x^{tq} x^r = (x^t)^q x^r = 1^q x^r = x^r.$$

Se fosse $r > 0$, da $x^r = 1$ seguirebbe che $\circ(x) \leq r$ e quindi $\circ(x) < t$: assurdo. Dunque $r = 0$.

Se $x^h = x^k$, allora $x^{h-k} = 1$ e quindi $t \mid h-k$, cioè $h \equiv k \pmod{t}$. Viceversa, sia $h \equiv k \pmod{t}$. Allora $h - k = tq$ e quindi $x^{h-k} = x^{tq} = 1$, cioè $1 = x^{h-k}$. Moltiplicando tale uguaglianza per x^k si ottiene: $x^k = x^k \cdot x^{h-k} = x^{k+h-k} = x^h$.

Proposizione 2. Sia (G, \cdot) un gruppo e sia $x \in G$. L'insieme $\{x^h, \forall h \in \mathbf{Z}\}$ è un sottogruppo di (G, \cdot) , detto sottogruppo ciclico generato da x e denotato $\langle x \rangle$.

Se $\circ(x) = t$, tale sottogruppo ha ordine t ed è formato dalle t potenze $1, x, x^2, \dots, x^{t-1}$.

Se $\circ(x) = \infty$, le potenze x^t sono a due a due distinte e quindi $|\langle x \rangle| = \infty$.

Dim. Per dimostrare che $\langle x \rangle$ è un sottogruppo di G basta verificare che, $\forall x^h, x^k \in \langle x \rangle$, anche $x^h(x^k)^{-1} \in \langle x \rangle$. Infatti $x^h(x^k)^{-1} = x^h x^{-k} = x^{h-k} \in \langle x \rangle$.

Sia ora $\circ(x) = t$. Consideriamo un'arbitraria potenza $x^h \in \langle x \rangle$ e dividiamo h per t . Sia $h = tq + r$, con $0 \leq r \leq t - 1$. Procedendo come nella proposizione precedente, si ottiene $x^h = x^r$. Dunque $\langle x \rangle \subseteq \{1, x, x^2, \dots, x^{t-1}\}$ e quindi (essendo l'inclusione opposta ovvia)

$$\langle x \rangle = \{1, x, x^2, \dots, x^{t-1}\}.$$

Resta da verificare che queste t potenze di x sono a due a due distinte. Infatti, per assurdo, sia $x^h = x^k$, con $0 \leq h < k \leq t - 1$. Dalla **Prop. 1**, $k - h = tq$, con $q \in \mathbf{Z}$. Poiché $0 < k - h < t$, allora $0 < tq < t$: assurdo.

Sia infine $\circ(x) = \infty$ e, per assurdo, $x^h = x^k$, con $h < k$. Procedendo come sopra si ottiene che $x^{k-h} = 1$ e dunque $\circ(x) \leq k - h$: assurdo.

Corollario 1. Sia (G, \cdot) un gruppo finito di ordine n . Ogni elemento di G ha periodo $\leq n$.

Dim. Poiché $\circ(x) = |\langle x \rangle|$ e $|\langle x \rangle| \leq |G| = n$, ogni elemento di G ha periodo finito $\leq n$.

N.B. Dimostreremo nel prossimo paragrafo un risultato più forte: $\circ(x)$ è un divisore di n .

Ora traduciamo i risultati precedenti nella notazione additiva: poiché i prodotti diventano somme, le potenze ad esponenti interi di x diventeranno i *multipli interi* di x .

Sia dunque $(G, +)$ un gruppo [con elemento neutro $0 = 0_G$]. Per ogni $x \in G$, se ne considerino i *multipli interi* tx ($\forall t \in \mathbf{Z}$), che sono così definiti:

$$\begin{aligned} 0x &= 0, \quad 1x = x, \quad (-1)x = -x \quad [\text{cioè } (-1)x \text{ è l'opposto di } x \text{ in } G], \\ tx &= \underbrace{x + x + \dots + x}_{t \text{ addendi}}, \quad \text{se } t \geq 2, \quad tx = \underbrace{(-x) + (-x) + \dots + (-x)}_{-t \text{ addendi}}, \quad \text{se } t \leq -2. \end{aligned}$$

Segue che, $\forall t, s \in \mathbf{Z}$:

$$tx + sx = (t+s)x = sx + tx, \quad t(sx) = s(tx) = (ts)x$$

[e quindi, in particolare, $t(-x) = -(tx) = (-t)x$].

Ovviamente $t0 = 0$, $\forall t \in \mathbf{Z}$, e, per ogni $x \neq 0$, si hanno due eventualità:

$$\exists t > 0 \mid tx = 0; \quad \forall t > 0, tx \neq 0.$$

Sussistono la seguente definizione ed i seguenti risultati, perfettamente analoghi ai precedenti. Invitiamo lo studente a rifarne comunque le dimostrazioni.

Definizione 1'. Sia $(G, +)$ un gruppo e sia $x \in G$. Diciamo che x ha periodo t se t è il minimo intero positivo tale che $tx = 0$; scriveremo in tal caso $\circ(x) = t$. Diremo invece che x ha periodo infinito se $tx \neq 0$, $\forall t > 0$, e scriveremo in tal caso $\circ(x) = \infty$.

Proposizione 1'. Sia $(G, +)$ un gruppo e sia $x \in G$. Se x ha periodo finito t e $hx = 0$, allora $t \mid h$ [cioè h è un multiplo di t]. Si ha poi, $\forall h, k \in \mathbf{Z}$:

$$hx = kx \iff h \equiv k \pmod{t}.$$

Proposizione 2'. Sia $(G, +)$ un gruppo e sia $x \in G$. L'insieme $\{hx, \forall h \in \mathbf{Z}\}$ è un sottogruppo di $(G, +)$, detto *sottogruppo ciclico generato da x* e denotato $\langle x \rangle$.

Se $\circ(x) = t$, tale sottogruppo ha ordine t ed è formato dai t multipli $0, x, 2x, \dots, (t-1)x$.

Se $\circ(x) = \infty$, i multipli tx sono a due a due distinti e quindi $|\langle x \rangle| = \infty$.

Corollario 1'. Sia $(G, +)$ un gruppo finito di ordine n . Ogni elemento di G ha periodo $\leq n$.

Veniamo ora ad alcuni esempi.

(1) In $(\mathbf{Z}, +)$ ogni intero non nullo ha periodo infinito. Se infatti $n \in \mathbf{Z}^*$, si ha: $hn \neq 0$, $\forall h > 0$.

(2) In S_n ogni k -ciclo ha periodo k . Sia infatti $\gamma = (c_1 c_2 \dots c_k)$ un k -ciclo di S_n . Risulta:

$$\begin{aligned}\gamma(c_1) &= c_2, \\ \gamma^2(c_1) &= \gamma(c_2) = c_3, \\ \gamma^3(c_1) &= \gamma(\gamma^2(c_1)) = \gamma(c_3) = c_4,\end{aligned}$$

...

$$\begin{aligned}\gamma^{k-1}(c_1) &= \gamma(\gamma^{k-2}(c_1)) = \gamma(c_{k-1}) = c_k, \\ \gamma^k(c_1) &= \gamma(\gamma^{k-1}(c_1)) = \gamma(c_1) = c_1.\end{aligned}$$

Dunque $\gamma^k(c_1) = c_1$ e, analogamente, si verifica che $\gamma^k(c_i) = c_i$. Pertanto

$$\gamma^k = \mathbf{1}_{S_n}, \text{ mentre } \gamma^i \neq \mathbf{1}_{S_n}, \quad \forall i = 1, \dots, k-1.$$

Quindi $\circ(\gamma) = k$. Si può inoltre verificare che, se γ_1, γ_2 sono rispettivamente un k_1 -ciclo ed un k_2 -ciclo disgiunti, allora

$$\circ(\gamma_1 \gamma_2) = mcm(k_1, k_2)$$

[dove $mcm(k_1, k_2)$ denota il *minimo comune multiplo* di k_1, k_2]. Più generalmente, se una permutazione $\sigma \in S_n$ è prodotto dei cicli disgiunti $\gamma_1, \dots, \gamma_i$ di lunghezze risp. k_1, \dots, k_i , allora

$$\circ(\sigma) = mcm(k_1, \dots, k_i).$$

(3) Calcoliamo i periodi degli elementi di $(\mathbf{Z}_{12}, +)$. Risulta:

$$\begin{aligned}\circ(\bar{1}) &= 12 \quad [\text{infatti } 12\bar{1} = \bar{12} = \bar{0}, \text{ mentre } h\bar{1} = \bar{h} \neq \bar{0}, \quad \forall h = 1, \dots, 11]; \\ \circ(\bar{2}) &= 6 \quad [\text{infatti } 6\bar{2} = \bar{12} = \bar{0}, \text{ mentre } h\bar{2} = \bar{2h} \neq \bar{0}, \quad \forall h = 1, \dots, 5].\end{aligned}$$

Analogamente,

$$\begin{aligned}\circ(\bar{3}) &= 4, \quad \circ(\bar{4}) = 3, \quad \circ(\bar{5}) = 12, \quad \circ(\bar{6}) = 2, \quad \circ(\bar{7}) = 12, \\ \circ(\bar{8}) &= 3, \quad \circ(\bar{9}) = 4, \quad \circ(\bar{10}) = 6, \quad \circ(\bar{11}) = 12, \quad \circ(\bar{12}) = \circ(\bar{0}) = 1.\end{aligned}$$

Si può osservare che tutti questi calcoli sono compendiati nella seguente formula:

$$\circ(\bar{k}) = \frac{12}{MCD(12, k)}$$

[dove $MCD(12, k)$ denota il massimo comun divisore tra 12, k]

N.B. Tale formula vale in ogni $(\mathbf{Z}_n, +)$:

$$\circ(\bar{k}) = \frac{n}{MCD(n, k)}, \quad \forall \bar{k} \in \mathbf{Z}_n.$$

Osservazione 1. Ogni sottogruppo ciclico è abeliano. Assumiamo infatti che G abbia notazione moltiplicativa e consideriamone un sottogruppo ciclico $\langle x \rangle$. Presi comunque $x^h, x^k \in \langle x \rangle$, risulta:

$$x^h x^k = x^{h+k} = x^{k+h} = x^k x^h.$$

Inoltre ogni sottogruppo ciclico $\langle x \rangle$ è finito o numerabile. Infatti, se $\circ(x) = t$, allora $|\langle x \rangle| = t$. Se invece $\circ(x) = \infty$, allora $\langle x \rangle$ è in corrispondenza biunivoca con \mathbf{Z} [e quindi con \mathbf{N}], tramite l'applicazione biiettiva

$$\varphi : \mathbf{Z} \rightarrow \langle x \rangle \text{ tale che } \varphi(h) = x^h, \quad \forall h \in \mathbf{Z},$$

[che φ sia suriettiva è ovvio; l'iniettività di φ è stata verificata nella dimostrazione della **Prop. 2**].

Veniamo ora alla definizione di *gruppo ciclico*, cioè gruppo che coincide con un proprio sottogruppo ciclico.

Definizione 2. Un gruppo G è detto *gruppo ciclico* se esiste $x \in G$ tale che $G = \langle x \rangle$. Ogni

siffatto elemento x è detto generatore del gruppo ciclico.

È subito evidente che esistono gruppi non ciclici. Infatti ogni gruppo non abeliano [ad esempio ogni S_n , con $n \geq 3$] in base all'**Osserv. 1** non è ciclico. [Ma potremmo allora chiederci: esistono gruppi abeliani non ciclici? Risponderemo a tale domanda alla fine del paragrafo].

Cerchiamo ora invece i gruppi ciclici. Le due proposizioni che seguono forniscono in proposito risposte assolutamente esaurienti.

Proposizione 3. Il gruppo $(\mathbf{Z}, +)$ è un gruppo ciclico. A meno di isomorfismi è l'unico gruppo ciclico infinito.

Dim. Risulta: $\langle 1 \rangle = \{h1, \forall h \in \mathbf{Z}\} = \mathbf{Z}$. Dunque $(\mathbf{Z}, +)$ è ciclico [si osservi che anche $\langle -1 \rangle = \mathbf{Z}$].

Sia ora (G, \cdot) un gruppo ciclico infinito, con $G = \langle x \rangle$ e $\circ(x) = \infty$. Dobbiamo verificare che $(G, \cdot) \cong (\mathbf{Z}, +)$. Si consideri l'applicazione

$$\varphi : \mathbf{Z} \rightarrow G = \langle x \rangle \text{ tale che } \varphi(h) = x^h, \forall h \in \mathbf{Z}.$$

Abbiamo già osservato (in **Osserv. 1**) che φ è biettiva. Si ha poi:

$$\varphi(h+k) = x^{h+k} = x^h x^k = \varphi(h) \cdot \varphi(k).$$

Dunque φ è un omomorfismo biettivo, cioè un isomorfismo di gruppi.

Proposizione 4. I gruppi ciclici finiti di un dato ordine sono tutti isomorfi tra loro. Per ogni $n > 0$ esiste un gruppo ciclico di ordine n (unico a meno di isomorfismi).

Dim. Siano G, H due gruppi ciclici finiti di ordine n . Assumiamo (ma non è restrittivo) che entrambi siano dotati di struttura moltiplicativa.

Sia dunque: $G = \langle x \rangle$, $H = \langle y \rangle$, con $\circ(x) = \circ(y) = n$. L'applicazione

$$\varphi : G \rightarrow H \text{ tale che } \varphi(x^t) = y^t, \forall t = 0, \dots, n-1,$$

è certamente biettiva. Basta quindi verificare che si tratta di un omomorfismo, cioè che

$$\varphi(x^t x^s) = \varphi(x^t) \varphi(x^s), \forall x^t, x^s \in G \text{ [con } 0 \leq t, s < n].$$

Se $t+s \equiv r \pmod{n}$, con $0 \leq r < n$, si ha (in base a **Prop. 1**) $x^{t+s} = x^r$ e $y^{t+s} = y^r$. Ne segue:

$$\varphi(x^t x^s) = \varphi(x^{t+s}) = \varphi(x^r) = y^r, \quad \varphi(x^t) \varphi(x^s) = y^t y^s = y^{t+s} = y^r.$$

Dunque φ è un isomorfismo.

Per completare la dimostrazione basta verificare che esiste sempre un gruppo ciclico finito di ordine n , $\forall n \geq 1$. Infatti, ogni gruppo di ordine 1 è ovviamente ciclico, mentre, per $n \geq 2$, si consideri il gruppo $(\mathbf{Z}_n, +)$: si tratta di un gruppo ciclico [infatti $\langle \bar{1} \rangle = \{h\bar{1}, \forall h \in \mathbf{Z}\} = \{\bar{0}, \bar{1}, \dots, \bar{n-1}\} = \mathbf{Z}_n$].

Osservazione 2. Esistono anche esempi di gruppi ciclici di ogni ordine n , dotati di struttura moltiplicativa: si tratta dei cosiddetti *gruppi delle radici n-sime dell'unità*.

Per radice n -sima dell'unità intendiamo ogni numero complesso $z = a + ib$ tale che $z^n = 1$. Si verifica subito che le radici n -sime dell'unità formano un gruppo rispetto al prodotto [se infatti $z^n = 1$ e $w^n = 1$, allora $(zw^{-1})^n = z^n(w^{-1})^n = 1$].

Per $n = 2$ le radici seconde dell'unità sono i due numeri reali 1, -1 , formanti il gruppo ciclico $\langle -1 \rangle$. Per ogni $n \geq 3$, le radici n -sime dell'unità sono i seguenti n numeri complessi (espressi in forma trigonometrica):

$$\zeta_{n,k} = \cos \frac{2k\pi}{n} + i \sin \frac{2k\pi}{n}, \quad \forall k = 0, 1, \dots, n-1.$$

Nel piano di Gauss tali numeri complessi sono visualizzati come i vertici del poligono regolare n -latero inscritto nella circonferenza goniometrica [cioè di centro $(0,0)$ e raggio 1] ed avente un vertice nel punto $(1, 0)$. Per ulteriori dettagli sui numeri complessi e sulla loro rappresentazione nel piano di Gauss, cfr. l'**Appendice** in coda a questo paragrafo.

Si osserva subito che il gruppo delle radici n -sime dell'unità è ciclico, generato ad esempio da $\zeta_{n,1} = \cos \frac{2\pi}{n} + i \sin \frac{2\pi}{n}$. Infatti si può verificare che $\zeta_{n,k} = (\zeta_{n,1})^k$, $\forall k = 0, 1, \dots, n-1$. Il numero complesso $\zeta_{n,1}$ corrisponde al primo vertice del poligono che si incontra dopo il vertice $(1, 0)$ percorrendo la circonferenza in verso antiorario.

Esaminiamo i gruppi delle radici quarte e terze dell'unità.

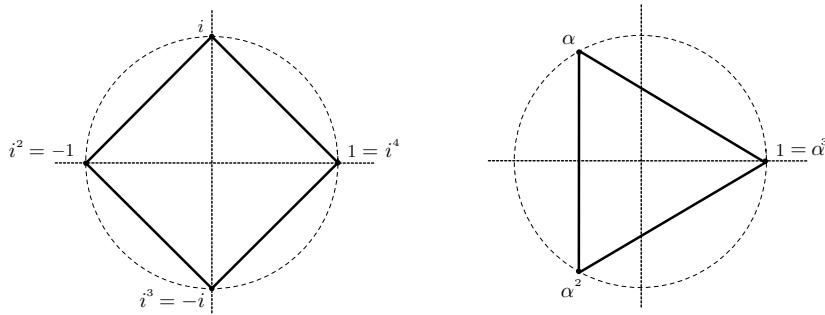
Una radice quarta dell'unità è il numero complesso $i (= \cos \frac{2\pi}{4} + i \sin \frac{2\pi}{4})$ e le altre tre radici quarte dell'unità sono i numeri complessi

$$i^2 = -1, \quad i^3 = -i, \quad i^4 = 1.$$

Una radice terza dell'unità è invece il numero complesso $\alpha = \cos \frac{2\pi}{3} + i \sin \frac{2\pi}{3} = -\frac{1}{2} + i \frac{\sqrt{3}}{2}$. Le altre due radici terze dell'unità sono

$$\alpha^2 = -\frac{1}{2} - i \frac{\sqrt{3}}{2}, \quad \alpha^3 = 1.$$

Nelle due figure che seguono visualizziamo le radici quarte e terze nel piano di Gauss.



Ma possiamo anche costruire i gruppi ciclici di ordine n in modo formale, o - come si dice - *tramite simboli e relazioni*. Basta considerare un simbolo x ed assegngargli la relazione $x^n = 1$. Ne segue che le potenze intere di x formano il gruppo ciclico di ordine n (dotato di struttura moltiplicativa)

$$C_n := \{1, x, x^2, \dots, x^{n-1}\},$$

che possiamo indicare anche $\langle x \mid x^n = 1 \rangle$.

Ad esempio il gruppo $C_4 = \langle x \mid x^4 = 1 \rangle$ (ciclico di ordine 4) ha la seguente tavola moltiplicativa:

\cdot	1	x	x^2	x^3
1	1	x	x^2	x^3
x	x	x^2	x^3	1
x^2	x^2	x^3	1	x
x^3	x^3	1	x	x^2

In base alla **Prop. 4**, i tre gruppi $\langle i \rangle$, \mathbf{Z}_4 e C_4 , essendo ciclici di ordine 4, sono isomorfi tra loro. Un isomorfismo tra due di essi si ottiene semplicemente associando un generatore dell'uno ad un generatore dell'altro. Ad esempio, un isomorfismo φ da $\langle i \rangle$ a \mathbf{Z}_4 è definito a partire da

$$\varphi(i) = \bar{1},$$

da cui $\varphi(-1) = \varphi(i^2) = \varphi(i) + \varphi(i) = \bar{1} + \bar{1} = \bar{2}$ e, allo stesso modo, $\varphi(-i) = \bar{3}$, $\varphi(1) = \bar{0}$. Analogamente, un isomorfismo ψ da \mathbf{Z}_4 a C_4 è definito ponendo

$$\psi(\bar{1}) = x,$$

da cui $\psi(\bar{2}) = \psi(\bar{1}) \cdot \psi(\bar{1}) = x^2$ e, analogamente, $\psi(\bar{3}) = x^3$, $\psi(\bar{0}) = 1$.

A proposito dei gruppi ciclici ci porremo le seguenti domande:

- (A) Quanti e quali sono i generatori di un gruppo ciclico?
- (B) Come sono fatti e quanti sono i sottogruppi di un gruppo ciclico?
- (C) Esistono gruppi abeliani, ma non ciclici?

Relativamente al problema (A), cominciamo ad esaminare il gruppo $(\mathbf{Z}, +)$, che, come sappiamo, a meno di isomorfismi è l'unico gruppo ciclico infinito.

Tale gruppo possiede, oltre al generatore 1, anche un altro generatore: -1 [infatti, $\forall n \in \mathbf{Z}$, $n = -n(-1)$. Dunque $\langle -1 \rangle = \mathbf{Z}$]. Gli altri interi non sono generatori di \mathbf{Z} . Ad esempio $\langle 2 \rangle = \langle -2 \rangle = 2\mathbf{Z}$ (interi pari), $\langle 3 \rangle = \langle -3 \rangle = 3\mathbf{Z}$ (multipli di 3), ecc..

Se G è un gruppo ciclico finito di ordine n , i suoi generatori sono tutti e soli gli elementi di G di periodo n . Per individuarli serve la seguente formula (che però non dimostriamo):

Sia (G, \cdot) un gruppo [arbitrario] e sia $x \in G$ un elemento di periodo finito. Risulta, $\forall t \in \mathbf{Z}$:

$$\circ(x^t) = \frac{\circ(x)}{\text{MCD}(t, \circ(x))}$$

[ed in notazione additiva, $\circ(tx) = \frac{\circ(x)}{\text{MCD}(t, \circ(x))}$. Si noti che tale formula è già stata presentata in coda al precedente esempio 3, relativamente ai gruppi \mathbf{Z}_n].

Segue subito da questa formula che, se $C_n = \langle x | x^n = 1 \rangle$, si ha:

x^t è un generatore di $C_n \iff \circ(x^t) = n \iff \text{MCD}(t, n) = 1 \iff t, n$ sono coprimi.

Ne segue che per ottenere i generatori di C_n basta determinare i $\varphi(n)$ interi t coprimi con n e compresi tra 1 ed n e considerare le potenze x^t corrispondenti.

Ad esempio, i generatori di $C_{12} = \langle x | x^{12} = 1 \rangle$ sono quattro: x, x^5, x^7, x^{11} . Analogamente, in $(\mathbf{Z}_{12}, +)$, i generatori (come già visto nel precedente esempio 3) sono $\bar{1}, \bar{5}, \bar{7}, \bar{11}$.

Nel gruppo ciclico $\langle i \rangle$ delle radici quarte dell'unità esistono due generatori: i e $-i$. Analogamente, in $(\mathbf{Z}_4, +)$ i due generatori sono $\bar{1}, \bar{3}$.

Veniamo ora al problema (B). Proviamo per prima cosa che tutti i sottogruppi di un gruppo ciclico sono gruppi ciclici.

Proposizione 5. *I sottogruppi di un gruppo ciclico sono gruppi ciclici.*

Dim. Distingueremo il caso del gruppo ciclico infinito $(\mathbf{Z}, +)$ dal caso dei gruppi ciclici finiti $C_n = \langle x | x^n = 1 \rangle$, anche se la dimostrazione è praticamente la stessa.

In $(\mathbf{Z}, +)$ consideriamo un sottogruppo H non nullo. Denotiamo con t il minimo intero positivo contenuto in H . Verifichiamo che $H = t\mathbf{Z}$.

Ovviamente $t\mathbf{Z} \subseteq H$. Viceversa, sia $h \in H$. Dividiamo h per t e sia $h = tq + r$, con $0 \leq r < t$. Poiché $r = h - tq \in H$, per la minimalità di t segue che $r = 0$. Dunque $h = qt \in t\mathbf{Z}$ e pertanto $H \subseteq t\mathbf{Z}$.

Sia ora H un sottogruppo di $C_n = \langle x | x^n = 1 \rangle$, con $H \neq \{1\}$. Sia $x^t \in H$, con t esponente minimo positivo (per cui ciò avvenga). Ovviamente $\langle x^t \rangle \subseteq H$. Viceversa, sia $x^h \in H$. Dividiamo h per t e sia $h = tq + r$, con $0 \leq r < t$. Poiché $x^r = x^{h-tq} = x^h(x^t)^{-q} \in H$, per la minimalità di t segue che $r = 0$. Dunque $x^h = (x^t)^q \in \langle x^t \rangle$ e pertanto $H \subseteq \langle x^t \rangle$.

Consideriamo ora un gruppo ciclico finito. Abbiamo appena dimostrato che tutti i suoi sottogruppi sono ciclici. Vale questo ulteriore risultato (che non dimostriamo).

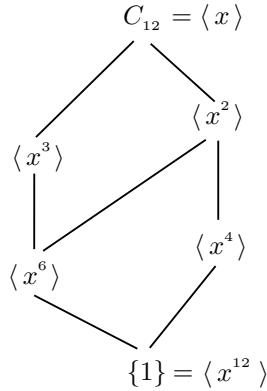
Proposizione 6. *Sia $C_n = \langle x | x^n = 1 \rangle$ un gruppo ciclico di ordine n . Per ogni divisore positivo d di n , esiste in C_n un unico sottogruppo di ordine d . Tale sottogruppo (che è ovviamente ciclico) è generato da $x^{n/d}$.*

Vedremo nel prossimo paragrafo che l'ordine di un sottogruppo di un gruppo finito è un divisore dell'ordine del gruppo (è il teorema di Lagrange). Questo risultato, unito alla Prop. 6, ci consente di concludere che in C_n non ci sono altri sottogruppi oltre a quelli indicati dalla Prop. 6.

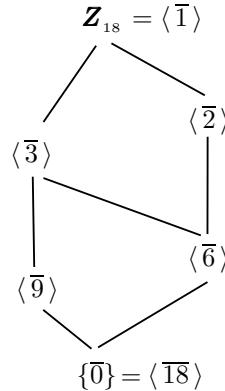
Possiamo ad esempio calcolare i sottogruppi di $C_{12} = \langle x | x^{12} = 1 \rangle$. Poiché i divisori positivi di $n = 12$ sono $d = 1, 2, 3, 4, 6, 12$, C_{12} possiede sei sottogruppi, cioè $\langle x^{12/d} \rangle$, con $d = 1, 2, 3, 4, 6, 12$. Si tratta quindi dei sottogruppi (anch'essi ciclici):

$$\langle x^{12} \rangle = \{1\}, \quad \langle x^6 \rangle, \quad \langle x^4 \rangle, \quad \langle x^3 \rangle, \quad \langle x^2 \rangle, \quad \langle x \rangle = C_{12}$$

[aventi rispettivamente ordine 1, 2, 3, 4, 6, 12]. Si osservi poi che $\langle x^2 \rangle$ (di ordine 6) contiene come sottogruppi $\langle x^4 \rangle$ (di ordine 3) e $\langle x^6 \rangle$ (di ordine 2); inoltre $\langle x^3 \rangle$ (di ordine 4) contiene $\langle x^6 \rangle$ (di ordine 2). Possiamo rappresentare i sottogruppi di C_{12} in questo grafico (detto *reticolo dei sottogruppi di C_{12}*). In esso i segmenti rappresentano le inclusioni tra sottogruppi (dal basso verso l'alto).



Un altro esempio: rappresentiamo, senza commenti, il reticolo dei sottogruppi di Z_{18} . In tal caso i sottogruppi sono i seguenti: $\langle \frac{18}{d} \bar{1} \rangle$, con $d = 1, 2, 3, 6, 9, 18$ e si ha:



Veniamo infine al problema (C): determinare esempi di gruppi abeliani non ciclici.

Un esempio di gruppo abeliano infinito e non ciclico è (Q^\cdot, \cdot) . Infatti $\forall q = \frac{a}{b} \in Q^\cdot$, possiamo verificare che il sottogruppo ciclico

$$\langle q \rangle = \left\{ \frac{a^t}{b^t}, \quad \forall t \in \mathbb{Z} \right\}$$

è contenuto propriamente in Q^\cdot . Ad esempio, scelto un primo p coprimo con gli interi a, b , non può risultare $p = \frac{a^t}{b^t}$, in quanto [in base all'unicità della fattorizzazione di un intero come prodotto di primi]:

$$\text{se } t > 0, \quad p b^t \neq a^t; \quad \text{se } t < 0, \quad p a^{-t} \neq b^{-t}.$$

Relativamente ai gruppi abeliani finiti, per ottenerne uno non ciclico bisogna cercarne uno i cui elementi abbiano periodi tutti inferiori all'ordine del gruppo. L'esempio più semplice si trova tra i gruppi di ordine 4 ed è chiamato *gruppo di Klein* (o anche *gruppo dei quattro*) e spesso denotato con la lettera V [V è l'iniziale di vier (quattro in tedesco)].

Tale gruppo può essere definito astrattamente in questo modo: V è generato da due simboli, che denotiamo a, b , soddisfacenti alle seguenti relazioni:

$$a^2 = 1, \quad b^2 = 1, \quad ab = ba.$$

Potremo quindi scrivere: $V = \langle a, b \mid a^2 = b^2 = 1, ba = ab \rangle$. Dalle relazioni tra i due simboli segue subito che: $a^{-1} = a$, $b^{-1} = b$. Inoltre

$$(ab)^2 = (ab)(ab) = a(ba)b = a(ab)b = a^2 b^2 = 1 \text{ e quindi } (ab)^{-1} = ab.$$

Tutte le possibili espressioni ottenute come prodotti finiti di potenze intere di a, b , si riducono ai quattro elementi $1, a, b, ab$. Dunque $\mathbf{V} = \{1, a, b, ab\}$. Si osservi che i tre elementi $\neq 1$ hanno periodo 2: dunque \mathbf{V} non è ciclico. La tavola moltiplicativa di tale gruppo è la seguente:

.	1	a	b	ab
1	1	a	b	ab
a	a	1	ab	b
b	b	ab	1	a
ab	ab	b	a	1

Si osservi che esempi concreti di gruppi di Klein si possono trovare all'interno di gruppi già studiati. Ad esempio in \mathbf{S}_4 sono gruppi di Klein (tra gli altri) i seguenti due sottogruppi

$$\{(1), (12), (34), (12)(34)\}, \quad \{(1), (12)(34), (13)(24), (14)(23)\}$$

[verificarne la chiusura rispetto al prodotto]. Ce ne sono altri due:

$$\{(1), (13), (24), (13)(24)\}, \quad \{(1), (14), (23), (14)(23)\}.$$

ESERCIZI PROPOSTI

5.1.1. Determinare il periodo dell'elemento x^{320} del gruppo ciclico $C_{15} = \langle x \mid x^{15} = 1 \rangle$. Indicare tutti i generatori del sottogruppo $\langle x^{320} \rangle$.

5.1.2. Determinare il reticolo dei sottogruppi di C_{20} .

5.1.3. Determinare i sottogruppi ciclici di \mathbf{S}_4 .

5.1.4. Verificare che $\mathbf{U}(\mathbf{Z}_9)$ è un gruppo ciclico e determinarne tutti i generatori.

5.1.5. Verificare che un gruppo di ordine 4 non possiede elementi di periodo 3. Dedurne che un gruppo di ordine 4 o è ciclico o è un gruppo di Klein [cioè che, a meno di isomorfismi, esistono due soli gruppi di ordine 4: C_4 e \mathbf{V}].

5.1.6. Sia G un gruppo di ordine 6.

(i) Verificare che G non può possedere cinque elementi di periodo 2.

(ii) Verificare che G non può possedere tre elementi di periodo 3.

5.1.7. Sono assegnati due simboli x, y legati dalle seguenti relazioni (moltiplicative):

$$x^4 = 1, \quad y^2 = 1, \quad yx = x^3y.$$

Verificare che gli elementi generati da tali simboli sono otto e scriverne la tavola moltiplicativa. Verificare che formano un gruppo [che è chiamato *gruppo diedrale di ordine 8*].

5.1.8. (i) Sia G un gruppo abeliano. Verificare che l'insieme H degli elementi di periodo finito di G è un sottogruppo di G .

(ii) Se invece G non è abeliano, H può non essere un sottogruppo di G . Per dimostrare tale affermazione si utilizzino i seguenti dati:

$$G = \mathbf{GL}_2(\mathbf{R}), \quad A = \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix} \in G.$$

Verificare che $A, {}^t A \in H$ mentre il prodotto $A {}^t A \notin H$.

APPENDICE: RICHIAMI SUI NUMERI COMPLESSI

L'equazione $X^2 + 1 = 0$ non ha soluzioni nel campo \mathbf{R} [infatti un quadrato non è mai negativo]. Cerchiamo allora un insieme contenente \mathbf{R} e contenente una soluzione (almeno) della precedente

equazione. Otterremo l'insieme \mathbf{C} dei numeri complessi.

Definizione 1. Sia i un simbolo verificante la condizione $i^2 = -1$. L'insieme delle espressioni formali $a + ib$, $\forall a, b \in \mathbf{R}$ è detto *insieme dei numeri complessi* e viene denotato \mathbf{C} . Per ogni numero complesso $z = a + ib \in \mathbf{C}$,

a è detta *parte reale di z* , denotata $\text{Re}(z)$; b è detta *parte immaginaria di z* , denotata $\text{Im}(z)$. Imponendo la regola $0i = 0$, si ottiene che, $\forall r \in \mathbf{R}$, $r = r + 0i \in \mathbf{C}$. Dunque $\mathbf{R} \subset \mathbf{C}$.

L'insieme \mathbf{C} può essere identificato con il piano \mathbf{R}^2 [che prende il nome di *piano di Gauss* (o *piano di Argand-Gauss*), tramite la biiezione

$$z = a + ib \in \mathbf{C} \longrightarrow (a, b) \in \mathbf{R}^2.$$

In particolare \mathbf{R} (come sottoinsieme di \mathbf{C}) coincide con l'*asse x* di \mathbf{R}^2 .

Definizione 2. Si definiscono su \mathbf{C} le due seguenti operazioni di *addizione* e *moltiplicazione*:

$$+ : \mathbf{C} \times \mathbf{C} \rightarrow \mathbf{C} \text{ tale che } (a_1 + ib_1) + (a_2 + ib_2) = (a_1 + a_2) + i(b_1 + b_2),$$

$$\cdot : \mathbf{C} \times \mathbf{C} \rightarrow \mathbf{C} \text{ tale che } (a_1 + ib_1) \cdot (a_2 + ib_2) = (a_1 a_2 - b_1 b_2) + i(a_1 b_2 + a_2 b_1),$$

$$\forall a_1 + ib_1, a_2 + ib_2 \in \mathbf{C}.$$

Osservazione 1. Nell'identificazione tra \mathbf{C} ed \mathbf{R}^2 , la somma di due numeri complessi corrisponde alla somma di vettori di \mathbf{R}^2 .

Il prodotto di due numeri complessi coincide con l'usuale moltiplicazione "polinomiale" delle due espressioni formali $a_1 + ib_1$, $a_2 + ib_2$, con l'ulteriore regola: $i^2 = -1$.

Proposizione 1. $(\mathbf{C}, +, \cdot)$ è un campo ed un \mathbf{R} -spazio vettoriale di dimensione 2.

Dim. [Lasciata per esercizio]. Si noti che: $0 = 0 + i0$ è l'elemento neutro della somma; $1 = 1 + 0i$ è l'elemento neutro del prodotto; $-z = -a - ib$ è l'opposto di $z = a + ib$; per ogni numero complesso $z = a + ib \neq 0$, $z^{-1} = \frac{a}{a^2 + b^2} - i \frac{b}{a^2 + b^2}$ è l'inverso di z . Infine, $\{1, i\}$ è una base di \mathbf{C} come \mathbf{R} -spazio vettoriale.

Definizione 3. Per ogni numero complesso $z = a + ib$,

$\bar{z} := a - ib$ è detto *coniugato di z* ;

$z\bar{z} = a^2 + b^2$ è detto *norma di z* ; si tratta di un numero reale ≥ 0 , denotato $\mathcal{N}(z)$;

$|z| := \sqrt{z\bar{z}} = \sqrt{a^2 + b^2}$ è detto *modulo di z* ; si tratta di un numero reale ≥ 0 .

Osservazione 2. Nell'identificazione tra \mathbf{C} ed \mathbf{R}^2 , \bar{z} è il simmetrico di z rispetto all'*asse x* . In base al teorema di Pitagora, $|z|$ è la distanza tra 0 e z [cioè tra l'origine $O = (0, 0)$ ed il punto (a, b) , se $z = a + ib$].

Proposizione 2. Risulta, $\forall z, z_1, z_2 \in \mathbf{C}$:

$$(i) \quad \bar{\bar{z}} = z; \quad \overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}; \quad \overline{z_1 \cdot z_2} = \overline{z_1} \cdot \overline{z_2}; \quad z = \bar{z} \iff z \in \mathbf{R};$$

$$(ii) \quad |z_1 z_2| = |z_1| \cdot |z_2|;$$

$$(iii) \quad |z_1 + z_2| \leq |z_1| + |z_2|.$$

Dim. La (i) e la (ii) sono del tutto ovvie; per la (iii) basta ricordare che in un triangolo la lunghezza di un lato è minore o uguale alla somma delle lunghezze degli altri due.

Ci occuperemo ora della rappresentazione trigonometrica dei numeri complessi.

Ad ogni numero complesso non nullo $z = a + ib \in \mathbf{C}^*$ resta associata una coppia (ρ, ϑ_0) di numeri reali, così definiti:

- $\rho = |z|$ è il *modulo* di z [già definito in **Def. 3**];
- $\vartheta_0 \in [0, 2\pi)$ è la misura in radianti dell'angolo (orientato in verso antiorario) di vertice $O = (0, 0)$ tra l'asse x (positivamente orientato) e la semiretta Oz . È detto *argomento principale di z* . Dalla trigonometria

$$\begin{cases} a = \rho \cos \vartheta_0 \\ b = \rho \sin \vartheta_0. \end{cases}$$

Osservazione 3. Sia $z = a + ib \in \mathbf{C}^*$. Abbiamo già osservato che

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

Relativamente a ϑ_0 , si ha:

$$\vartheta_0 = \begin{cases} \arccos \frac{a}{\sqrt{a^2 + b^2}}, & \text{se } b \geq 0 \\ 2\pi - \arccos \frac{a}{\sqrt{a^2 + b^2}}, & \text{se } b < 0. \end{cases}$$

Infatti, $\forall \vartheta_0 \in [0, 2\pi)$ risulta: $a = \rho \cos \vartheta_0$, cioè $\cos \vartheta_0 = \frac{a}{\rho}$. Ne segue:

- se $\vartheta_0 \in [0, \pi]$, cioè se $b \geq 0$, allora $\vartheta_0 = \arccos \frac{a}{\rho} = \arccos \frac{a}{\sqrt{a^2 + b^2}}$;
- se $\vartheta_0 \in (\pi, 2\pi)$, cioè se $b < 0$, allora [usando la seconda determinazione della funzione arccos cioè la funzione $f(x) = 2\pi - \arccos(x)$] si ottiene: $\vartheta_0 = 2\pi - \arccos \frac{a}{\rho} = 2\pi - \arccos \frac{a}{\sqrt{a^2 + b^2}}$.

Abbiamo così costruito un'applicazione

$$\Phi : \mathbf{C}^* \rightarrow \mathbf{R}^+ \times [0, 2\pi) \quad \text{tale che } \Phi(z) = (\rho, \vartheta_0), \quad \forall z \in \mathbf{C}^*,$$

con ρ, ϑ_0 definiti come nella precedente osservazione.

Tale applicazione Φ è biettiva. Per dimostrarlo ne costruiremo l'inversa. Poniamo

$$\Psi : \mathbf{R}^+ \times [0, 2\pi) \rightarrow \mathbf{C}^* \quad \text{tale che } \Psi(\rho, \vartheta_0) = \rho(\cos \vartheta_0 + i \sin \vartheta_0), \quad \forall (\rho, \vartheta_0) \in \mathbf{R}^+ \times [0, 2\pi).$$

Si tratta ora di verificare che risulta:

$$\Psi \circ \Phi = \mathbf{1}_{\mathbf{C}^*}, \quad \Phi \circ \Psi = \mathbf{1}_{\mathbf{R}^+ \times [0, 2\pi)}.$$

[Le semplici verifiche sono lasciate al lettore].

Abbiamo dunque ottenuto che, $\forall z \in \mathbf{C}^*$ esiste un'unica coppia $(\rho, \vartheta_0) \in \mathbf{R}^+ \times [0, 2\pi)$ tale che

$$z = \rho(\cos \vartheta_0 + i \sin \vartheta_0).$$

Tale espressione è detta *espressione trigonometrica di z* .

Osservazione 4. La biiezione $\Psi : \mathbf{R}^+ \times [0, 2\pi) \rightarrow \mathbf{C}^*$ può essere estesa ad un'applicazione [suriettiva ma non iniettiva]

$$\psi : \mathbf{R}^+ \times \mathbf{R} \rightarrow \mathbf{C}^* \quad \text{tale che } \psi(\rho, \vartheta) = \rho(\cos \vartheta + i \sin \vartheta), \quad \forall (\rho, \vartheta) \in \mathbf{R}^+ \times \mathbf{R}.$$

Ne segue che ogni $z \in \mathbf{C}^*$ si scrive nella forma

$$z = \rho(\cos \vartheta + i \sin \vartheta),$$

ancora detta *espressione trigonometrica di z* ; ϑ è detto *argomento di z* . Tale espressione non è però unica per z .

Ricordato che le funzioni \cos e \sin sono funzioni periodiche (di periodo 2π), si verifica facilmente che, $\forall z_1 = \rho_1(\cos \vartheta_1 + i \sin \vartheta_1)$, $z_2 = \rho_2(\cos \vartheta_2 + i \sin \vartheta_2) \in \mathbf{C}^*$, si ha:

$$z_1 = z_2 \iff \rho_1 = \rho_2 \text{ e } \vartheta_2 = \vartheta_1 + 2k\pi, \quad \exists k \in \mathbf{Z}.$$

Vogliamo ora calcolare il prodotto di due numeri complessi espressi in forma trigonometrica. Vale il seguente risultato.

Proposizione 3. (i) Siano $z_1 = \rho_1(\cos \vartheta_1 + i \sin \vartheta_1)$, $z_2 = \rho_2(\cos \vartheta_2 + i \sin \vartheta_2) \in \mathbf{C}^*$. Risulta:

$$z_1 z_2 = \rho_1 \rho_2 (\cos(\vartheta_1 + \vartheta_2) + i \sin(\vartheta_1 + \vartheta_2)).$$

(ii) (*Formula di De Moivre*). Sia $z = \rho(\cos \vartheta + i \sin \vartheta) \in \mathbf{C}$. Per ogni $n \in \mathbf{Z}$, risulta:

$$z^n = \rho^n (\cos n\vartheta + i \sin n\vartheta)$$

Dim. (i) Si ha:

$$\begin{aligned} z_1 z_2 &= \rho_1 \rho_2 (\cos \vartheta_1 + i \sin \vartheta_1) (\cos \vartheta_2 + i \sin \vartheta_2) = \\ &= \rho_1 \rho_2 [(\cos \vartheta_1 \cos \vartheta_2 - \sin \vartheta_1 \sin \vartheta_2) + i(\sin \vartheta_1 \cos \vartheta_2 + \cos \vartheta_1 \sin \vartheta_2)] \\ &= \rho_1 \rho_2 (\cos(\vartheta_1 + \vartheta_2) + i \sin(\vartheta_1 + \vartheta_2)). \end{aligned}$$

(ii) Se $n > 0$, segue subito da (i). Se $n = 0$, $\rho^0 (\cos 0 + i \sin 0) = 1 = z^0$. Sia $n < 0$. Si osservi che, posto $w := \rho^{-1} (\cos(-\vartheta) + i \sin(-\vartheta))$, risulta da (i) che $w = z^{-1}$. Allora

$$z^n = (z^{-1})^{|n|} = (\rho^{-1})^{|n|} (\cos(|n|(-\vartheta)) + i \sin(|n|(-\vartheta))) = \rho^n (\cos n\vartheta + i \sin n\vartheta).$$

Si noti che, posto $e^{i\vartheta} := \cos \vartheta + i \sin \vartheta$, $\forall \vartheta \in \mathbf{R}$ [nota come *identità di Eulero*], l'espressione trigonometrica $z = \rho(\cos \vartheta + i \sin \vartheta)$ diventa

$$z = \rho e^{i\vartheta},$$

nota come *espressione esponenziale di z*. La formula di De Moivre diventa:

$$z^n = \rho^n e^{in\vartheta}, \quad \forall n \in \mathbf{Z}.$$

Veniamo ora alla definizione di *radice n-sima di un numero complesso*.

Definizione 4. Sia $z \in \mathbf{C}$ e sia $n \in \mathbf{N}$. Si chiama *radice n-sima di z* ogni $\alpha \in \mathbf{C}$ tale che $\alpha^n = z$. Denotiamo con $\sqrt[n]{z} = \{\alpha \in \mathbf{C} : \alpha^n = z\}$ l'insieme delle radici n-sime di z. Si tratta delle soluzioni dell'equazione $X^n = z$.

Osservazione 5. Sia $z = r(\cos t + i \sin t) \neq 0$, con $r > 0$ e $t \in [0, 2\pi)$. Denotiamo con $\alpha = \rho(\cos \vartheta + i \sin \vartheta)$ un arbitrario numero complesso $\neq 0$, al momento non noto. Poiché, in base alla formula di de Moivre, $\alpha^n = \rho^n (\cos n\vartheta + i \sin n\vartheta)$, allora:

$$\alpha^n = z \iff \rho^n = r \quad \text{e} \quad n\vartheta = t + 2k\pi, \quad \exists k \in \mathbf{Z} \iff \rho = \sqrt[n]{r} \quad \text{e} \quad \vartheta = \frac{t+2k\pi}{n}, \quad \exists k \in \mathbf{Z}.$$

Essendo le funzioni cos e sin periodiche di periodo 2π , allora

$$\sqrt[n]{z} = \left\{ \sqrt[n]{r} \left(\cos \frac{t+2k\pi}{n} + i \sin \frac{t+2k\pi}{n} \right), \quad \forall k \in \mathbf{Z} \right\}.$$

Si può verificare che, $\forall h, k \in \{0, 1, \dots, n-1\}$, con $h \neq k$, gli argomenti $\frac{t+2h\pi}{n}$ e $\frac{t+2k\pi}{n}$ non differiscono per multipli interi di 2π . Viceversa, $\forall h \in \mathbf{Z}$, esiste un unico $r \in \{0, 1, \dots, n-1\}$, tale che $\frac{t+2h\pi}{n} = \frac{t+2r\pi}{n} + 2q\pi$ [basta dividere h per n: si ottiene $h = nq+r$, con $0 \leq r < n$]. Dunque si ha:

$$\sqrt[n]{z} = \left\{ \sqrt[n]{r} \left(\cos \frac{t+2k\pi}{n} + i \sin \frac{t+2k\pi}{n} \right), \quad \forall k = 0, 1, \dots, n-1 \right\}.$$

Concludiamo che $\sqrt[n]{z}$ è formato esattamente da n numeri complessi distinti, che hanno lo stesso modulo [cioè $\sqrt[n]{r}$] e dunque (pensati in \mathbf{R}^2) giacciono su una stessa circonferenza di centro l'origine 0 e raggio $\sqrt[n]{r}$. Sono ottenuti l'uno dall'altro con una rotazione antioraria di angolo multiplo di $\frac{2\pi}{n}$. Quindi sono i vertici di un poligono regolare n-latero, inscritto in una circonferenza di raggio $\sqrt[n]{r}$.

Se in particolare consideriamo il numero complesso $1 = \cos 0 + i \sin 0$, otteniamo l'insieme delle *radici n-sime dell'unità*:

$$\sqrt[n]{1} = \left\{ \cos \frac{2k\pi}{n} + i \sin \frac{2k\pi}{n}, \quad \forall k = 0, 1, \dots, n-1 \right\},$$

che è comunemente denotato \mathbf{C}_n . Per indicare le radici n-sime dell'unità si usa la seguente notazione:

$$\zeta_{n,k} := \cos \frac{2k\pi}{n} + i \sin \frac{2k\pi}{n}, \quad \forall n \geq 1, \quad \forall k = 0, 1, \dots, n-1.$$

In particolare, si pone:

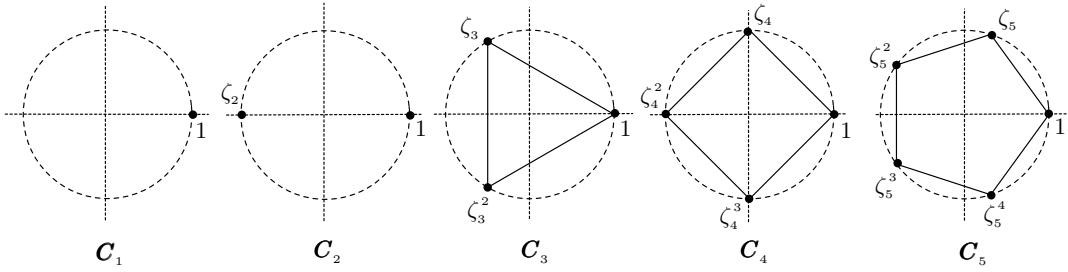
$$\zeta_n := \zeta_{n,1} = \cos \frac{2\pi}{n} + i \sin \frac{2\pi}{n}.$$

In base alla formula di De Moivre, risulta: $\zeta_n^k = \zeta_{n,k}$. Dunque

$$\mathbf{C}_n = \left\{ \zeta_n^k, \quad \forall k = 0, 1, \dots, n-1 \right\} = \left\{ 1, \zeta_n, \zeta_n^2, \dots, \zeta_n^{n-1} \right\}.$$

A titolo di esempio, scriviamo tutte le radici n-sime, con $1 \leq n \leq 5$. Risulta:

$\mathbf{C}_1 = \{1\}$, $\mathbf{C}_2 = \{1, \zeta_2\}$, $\mathbf{C}_3 = \{1, \zeta_3, \zeta_3^2\}$, $\mathbf{C}_4 = \{1, \zeta_4, \zeta_4^2, \zeta_4^3\}$, $\mathbf{C}_5 = \{1, \zeta_5, \zeta_5^2, \zeta_5^3, \zeta_5^4\}$,
con $\zeta_2 = -1$, $\zeta_3 = -\frac{1}{2} + i\frac{\sqrt{3}}{2}$, $\zeta_4 = i$, $\zeta_5 = \cos \frac{2\pi}{5} + i \sin \frac{2\pi}{5}$.



Concludiamo questa introduzione ai numeri complessi con un cenno al *teorema fondamentale dell'Algebra*.

Abbiamo osservato che il polinomio $X^2 + 1$ (irriducibile su \mathbf{R}) è riducibile su \mathbf{C} . Infatti si fattorizza nel prodotto $(X - i)(X + i)$. Inaspettatamente, lo stesso fatto vale per ogni polinomio di grado ≥ 2 in $\mathbf{C}[X]$. Questo è il contenuto del cosiddetto *teorema fondamentale dell'Algebra*.

Teorema 1 (*Teorema fondamentale dell'Algebra*). *I soli polinomi irriducibili in $\mathbf{C}[X]$ sono i polinomi di primo grado. Dunque ogni polinomio in $\mathbf{C}[X]$ di grado $n \geq 1$ si fattorizza nel prodotto di n fattori lineari.*

2. Teorema di Lagrange e gruppi quoziente

Nel precedente paragrafo abbiamo più volte accennato al seguente importante risultato, che intendiamo ora dimostrare e che è noto come *Teorema di Lagrange*.

Teorema 1. *Sia G un gruppo finito e sia H un suo sottogruppo. L'ordine di H è un divisore dell'ordine di G .*

Per dimostrare il teorema di Lagrange dobbiamo prima introdurre il concetto di *classe laterale destra [o sinistra] modulo un sottogruppo*.

Definizione 1. *Sia (G, \cdot) un gruppo e sia H un suo sottogruppo. Sia $a \in G$. Chiamiamo classe laterale destra di a modulo H [o, più semplicemente, laterale destro di a modulo H] l'insieme*

$$Ha = \{ha, \forall h \in H\}.$$

Osservazione 1. (i) Se G ha struttura additiva, il laterale destro di a modulo H è l'insieme

$$H + a = \{h + a, \forall h \in H\}.$$

(ii) Tra i laterali destri modulo H c'è H stesso [infatti $H = H1$]. Si noti anzi che risulta $H = Hh_0, \forall h_0 \in H$ [infatti l'inclusione $H \supseteq Hh_0$ è ovvia. Viceversa, ogni $h \in H$ si scrive nella forma $h = (hh_0^{-1})h_0$ e dunque $h \in Hh_0$].

(iii) Se $Ha \neq H$, Ha non può essere un sottogruppo di G , in quanto $1 \notin Ha$ [altrimenti si avrebbe $1 = ha$ (per qualche $h \in H$) e dunque $a = h^{-1} \in H$; ma allora $Ha = Hh^{-1} = H$, contro l'ipotesi].

(iv) $a \in Ha, \forall a \in G$ [infatti $a = 1a \in Ha$].

(v) Tutti i laterali destri modulo H hanno la stessa cardinalità. Per verificare tale fatto basta osservare che Ha e H sono in corrispondenza biunivoca, tramite l'applicazione

$$H \rightarrow Ha \text{ tale che } h \rightarrow ha, \forall h \in H.$$

[Che tale applicazione sia suriettiva è ovvio; che sia iniettiva discende dalla legge di cancellazione: $h_1a = h_2a \implies h_1 = h_2$]. Ne segue che $|Ha| = |H| = |Hb|, \forall a, b \in G$.

Il risultato che segue ci dice quando due laterali destri coincidono.

Proposizione 1. *Sia H un sottogruppo di (G, \cdot) . Risulta:*

$$Ha = Hb \iff ab^{-1} \in H.$$

Dim. (\implies). Se $Ha = Hb, a = 1a = hb, \exists h \in H$. Ne segue che $ab^{-1} = h \in H$.

(\impliedby). Sia $ab^{-1} = h_0 \in H$. Allora $a = h_0b$ e $b = h_0^{-1}a$. Per ogni $h \in H$ si ha:

$$ha = h h_0 b \in Hb; \quad hb = h h_0^{-1} a \in Ha.$$

Dunque $Ha \subseteq Hb$ e $Hb \subseteq Ha$, cioè $Ha = Hb$.

N.B. Da tale equivalenza segue subito che il laterale destro Ha è rappresentato da ogni elemento $a_1 \in Ha$ [infatti $a_1 \in Ha \implies a_1 a^{-1} \in H \implies Ha_1 = Ha$].

Proposizione 2. *Sia H un sottogruppo di (G, \cdot) . Risulta:*

$$Ha \cap Hb \neq \emptyset \implies Ha = Hb.$$

Dim. Sia $x \in Ha \cap Hb$. Allora $x = h_1a = h_2b$, $\exists h_1, h_2 \in H$. Segue che $h_2 = h_1ab^{-1}$ e quindi $ab^{-1} = h_1^{-1}h_2 \in H$. Dalla **Prop. 1**, $Ha = Hb$.

La precedente proposizione afferma che due laterali destri modulo H o coincidono o sono disgiunti. Poiché inoltre [in base all'**Osserv. 1(iv)**] ogni elemento di G appartiene ad un laterale destro modulo H , si conclude che i laterali destri modulo H formano una partizione di G .

Denoteremo con $\mathcal{L}_d(H)$ l'insieme di tutti i laterali destri modulo H a due a due distinti. La cardinalità di $\mathcal{L}_d(H)$ [cioè il numero dei laterali destri modulo H] è detto *indice di H in G* .

[A rigore dovremmo parlare di indice '*destro*' di H in G . Ma come vedremo nella successiva **Prop. 3**, non c'è bisogno di utilizzare questo aggettivo].

Osservazione 2. Come si traducono i precedenti risultati in notazione additiva?

Se H è un sottogruppo di $(G, +)$, abbiamo già osservato che il laterale destro di a modulo H è $H + a$. Inoltre:

$$H = H + 0 = H + h_0, \quad \forall h_0 \in H;$$

$H + a$ non è un gruppo, se $H + a \neq H$;

$H + a$ e $H + b$ sono in corrispondenza biunivoca;

$$H + a = H + b \iff a - b \in H;$$

$$(H + a) \cap (H + b) \neq \emptyset \implies H + a = H + b;$$

i laterali destri modulo H formano una partizione di G .

Esempi. (i) Calcoliamo i laterali destri modulo i due seguenti sottogruppi di S_3 :

$$H = \langle (23) \rangle, \quad H_1 = \langle (123) \rangle.$$

Ovviamente $H = \{(1), (23)\}$. Poiché $|H| = 2$ ed $|S_3| = 6$, c'è spazio in S_3 per due altri laterali destri modulo H . Poiché $(12) \notin H$ un altro laterale destro modulo H è

$$H(12) = \{(1)(12), (23)(12)\} = \{(12), (123)\}.$$

Poiché $(13) \notin H \cup H(12)$, il terzo laterale destro modulo H è

$$H(13) = \{(1)(13), (23)(13)\} = \{(13), (132)\}.$$

[Si noti che $H(12) = H(123)$ e che $H(13) = H(132)$].

Consideriamo l'altro sottogruppo H_1 . Risulta: $H_1 = \{(1), (123)(132)\}$. Poiché $|H_1| = 3$ ed $|S_3| = 6$, c'è spazio in S_3 per un solo altro laterale destro modulo H_1 . Poiché ad esempio $(12) \notin H_1$ tale laterale destro è

$$H_1(12) = \{(1)(12), (123)(12), (132)(12)\} = \{(12), (23), (13)\}.$$

[Si noti che $H_1(12) = H_1(13) = H_1(23)$].

(ii) Consideriamo in $(\mathbf{Z}, +)$ il sottogruppo $H = \langle 5 \rangle [= 5\mathbf{Z}]$. I suoi laterali destri modulo H sono i seguenti cinque:

$$5\mathbf{Z}, \quad 5\mathbf{Z} + 1, \quad 5\mathbf{Z} + 2, \quad 5\mathbf{Z} + 3, \quad 5\mathbf{Z} + 4.$$

[Ad esempio $5\mathbf{Z} + 2 = \{n \in \mathbf{Z} \mid n \equiv 2 \pmod{5}\} = \{\dots, -8, -3, 2, 7, 12, \dots\}$].

(iii) Consideriamo in (\mathbf{Q}^*, \cdot) il sottogruppo $H = \mathbf{Q}^*$. In tal caso i laterali destri modulo \mathbf{Q}^* sono infiniti. Basta verificare ad esempio che la famiglia di laterali destri

$$\{\mathbf{Q}^* \sqrt{p}, \quad \forall p \text{ numero primo (positivo)}\}$$

è una famiglia di laterali destri modulo \mathbf{Q}^* , a due a due distinti. Si ponga infatti $\mathbf{Q}^* \sqrt{p} = \mathbf{Q}^* \sqrt{q}$, con p, q interi primi (positivi): ne segue che $\sqrt{\frac{p}{q}} \in \mathbf{Q}^*$. Ciò è impossibile se $p \neq q$.

Veniamo ora alla dimostrazione del teorema di Lagrange (**Teor. 1**), enunciato all'inizio del paragrafo.

Dim. (Teor. 1). Essendo G finito, i laterali destri modulo H (a due a due distinti) sono in numero

finito [e tale numero è l'indice i di H in G]. Sia quindi

$$\mathfrak{L}_d(H) = \{Ha_1, Ha_2, \dots, Ha_i\}.$$

Poiché $\mathfrak{L}_d(H)$ è una partizione di G e poiché (in base a **Osserv. 1(v)**) $|Ha_t| = |H|$, $\forall t = 1, \dots, i$:

$$|G| = |Ha_1| + |Ha_2| + \dots + |Ha_i| = i|H|.$$

Dunque $|H|$ è un divisore di $|G|$.

N.B. Da $|G| = i|H|$ segue ovviamente che $i = \frac{|G|}{|H|}$.

Corollario 1. *Sia G un gruppo finito di ordine n . Per ogni $x \in G$, $\circ(x)$ è un divisore di n .*

Dim. Infatti $\circ(x) = |\langle x \rangle|$ è un divisore di $|G| = n$.

Corollario 2. *Sia G un gruppo finito di ordine p primo. G non possiede sottogruppi propri (cioè non banali). Inoltre G è ciclico (di ordine p).*

Dim. Sia H un sottogruppo di G . Poiché $|H|$ è un divisore di p , allora $|H| = 1$ oppure $|H| = p$, cioè $H = \{1\}$ oppure $H = G$.

Sia ora $x \in G$, $x \neq 1$. In base al precedente corollario, $\circ(x) = 1$ oppure p . Poiché $\circ(x) \neq 1$, allora $\circ(x) = p$. Dunque $\langle x \rangle = G$.

Si noti che il teorema di Lagrange non si inverte. Non è detto cioè che, per ogni divisore positivo t dell'ordine di G , esista in G un sottogruppo di ordine t . Ad esempio si potrebbe verificare che il gruppo alterno su 4 elementi A_4 [cfr. **Esempi 2(ii)** di Cap. 2.1] ha ordine 12 ma non possiede alcun sottogruppo di ordine 6 [cfr. www.mat.uniroma1.it/people/campanella, **Appunti di Algebra 1**, pag. 176].

Affrontiamo ora il seguente problema: è possibile definire una struttura di gruppo sull'insieme $\mathfrak{L}_d(H)$? Volendo definire un'operazione tra due laterali destri a partire dall'operazione di G , la definizione più naturale è la seguente:

$$(*) \quad Ha \cdot Hb = Hab, \quad \forall Ha, Hb \in \mathfrak{L}_d(H).$$

Dobbiamo però porci la seguente domanda: se $Ha = Ha_1$ e $Hb = Hb_1$, è vero che $Hab = Ha_1b_1$? La risposta a questa domanda è in generale negativa. Lo verifichiamo con questo esempio.

Riconsideriamo in S_3 il sottogruppo $H = \langle (23) \rangle$ [cfr. **Esempi (i)**]. Abbiamo già osservato che $H(12) = H(123)$ e che $H(13) = H(132)$. Si ha però

$$H(12) \cdot H(13) = H(123) \text{ mentre } H(123) \cdot H(132) = H$$

[e $H \neq H(123)$].

Dunque - come si usa dire - *l'operazione (*) non è in generale ben definita*. Ma in alcuni casi lo è. Quando? Per rispondere a questa domanda è necessario introdurre le *classi laterali sinistre* (per confrontarle poi con le destre). Ma si noti comunque già da ora che, quando l'operazione (*) è ben definita, gli assiomi di gruppo sono certamente verificati. Infatti risulta:

$$Ha \cdot H = Ha = H \cdot Ha; \quad (Ha \cdot Hb) \cdot Hc = Ha \cdot (Hb \cdot Hc); \quad Ha \cdot Ha^{-1} = H = Ha^{-1} \cdot Ha.$$

Definizione 2. *Sia H un sottogruppo di (G, \cdot) e sia $a \in G$. Chiamiamo classe laterale sinistra di a modulo H [o, più semplicemente, laterale sinistro di a modulo H] l'insieme*

$$aH = \{ah, \forall h \in H\}.$$

Si può facilmente verificare che

$$aH = bH \iff a^{-1}b \in H$$

e che anche l'insieme $\mathfrak{L}_s(H)$ dei laterali sinistri modulo H forma una partizione di G .

Usando i laterali sinistri si può ridimostrare (con identica dimostrazione) il teorema di Lagrange. Proveremo ora che la cardinalità dei laterali destri coincide con quella dei laterali sinistri, cioè che l'indice 'destro' di H in G coincide con l'indice 'sinistro'.

Proposizione 3. *Sia H un sottogruppo di G . Risulta: $|\mathfrak{L}_d(H)| = |\mathfrak{L}_s(H)|$.*

Dim. Si tratta di verificare che esiste una biiezione tra $\mathfrak{L}_d(H)$ e $\mathfrak{L}_s(H)$. Poniamo:

$$\Phi : \mathfrak{L}_d(H) \rightarrow \mathfrak{L}_s(H) \text{ tale che } Ha \rightarrow a^{-1}H, \quad \forall Ha \in \mathfrak{L}_d(H).$$

Osserviamo che Φ è ben definita, cioè che, se $Ha = Hb$, allora $\Phi(Ha) = \Phi(Hb)$. Infatti: $Ha = Hb \implies ab^{-1} \in H \implies ba^{-1} \in H \implies ba^{-1}H = H \implies a^{-1}H = b^{-1}H \implies \Phi(Ha) = \Phi(Hb)$. Queste implicazioni sono in realtà equivalenze. Lette da destra a sinistra provano che Φ è iniettiva. Infine Φ è suriettiva: infatti $\forall aH \in \mathfrak{L}_s(H)$, risulta $\Phi(Ha^{-1}) = aH$.

Per un gruppo abeliano G risulta:

$$aH = Ha, \quad \forall a \in G$$

[infatti $ah = ha, \forall h \in H$]. Per un gruppo non abeliano la situazione è invece più complicata, come cercheremo di chiarire con i due seguenti esempi.

Consideriamo di nuovo il gruppo S_3 e calcoliamo i laterali sinistri modulo $H = \langle (23) \rangle$. Oltre al laterale sinistro H abbiamo gli altri due laterali sinistri:

$$(12)H = \{(12)(1), (12)(23)\} = \{(12), (132)\}; \\ (13)H = \{(13)(1), (13)(23)\} = \{(13), (123)\}.$$

Come si può osservare, confrontando con calcoli svolti in precedenza,

$$(12)H \neq H(12), \quad (13)H \neq H(13).$$

Se invece consideriamo i laterali sinistri modulo il sottogruppo $H_1 = \langle (123) \rangle$, oltre al laterale H_1 abbiamo il laterale sinistro

$$(12)H_1 = \{(12)(1), (12)(123), (12)(132)\} = \{(12), (13), (23)\}.$$

Come si vede,

$$(12)H_1 = H_1(12).$$

Si noti che tale uguaglianza non avviene però '*elemento per elemento*'. Non è vero cioè che $(12)\sigma = \sigma(12), \forall \sigma \in H_1$. È invece vero che l'uguaglianza tra i due laterali vale *in modo complessivo*, nel senso che:

$$\forall \sigma \in H_1, \exists \tau \in H_1 \text{ tale che } (12)\sigma = \tau(12).$$

Definizione 3. *Si chiama sottogruppo normale ogni sottogruppo H di G tale che*

$$aH = Ha, \quad \forall a \in G.$$

Per indicare che H è normale in G si scrive di solito $H \trianglelefteq G$.

Nell'esempio precedente abbiamo quindi provato che $\langle (123) \rangle$ è un sottogruppo normale di S_3 , mentre $\langle (23) \rangle$ non lo è. Inoltre ogni sottogruppo di un gruppo abeliano è normale.

La seguente proposizione ci spiega perché i sottogruppi normali sono importanti.

Proposizione 4. *Sia (G, \cdot) un gruppo e sia H un suo sottogruppo. Sia $(*)$ l'operazione precedentemente definita su $\mathfrak{L}_d(H)$. Si ha:*

$$(*) \text{ è ben definita} \iff H \text{ è un sottogruppo normale di } G.$$

Dim. (\implies). Cominciamo col verificare che $aH \subseteq Ha, \forall a \in G$. Si tratta cioè di verificare che

$$ah \in Ha, \quad \forall a \in G, \forall h \in H.$$

Poiché $(*)$ è ben definita, si ha:

$$Hah = Ha \cdot Hh = Ha \cdot H = Ha.$$

Dunque $aha^{-1} \in H$ e pertanto $ah \in Ha$.

Proviamo ora l'inclusione opposta: $Ha \subseteq aH$, $\forall a \in G$. Per l'inclusione sopra dimostrata (applicata ad a^{-1}) si ha: $a^{-1}H \subseteq Ha^{-1}$.

Moltiplicando a destra per a : $a^{-1}Ha \subseteq Ha^{-1}a = H$. Dunque $a^{-1}Ha \subseteq H$. Moltiplicando tale inclusione a sinistra per a : $aa^{-1}Ha \subseteq aH$, cioè $Ha \subseteq aH$.

Abbiamo così provato che $aH = Ha$, $\forall a \in G$. Pertanto $H \trianglelefteq G$.

(\Leftarrow). Verifichiamo che (*) è ben definita, cioè che

$$Ha = Ha_1, \quad Hb = Hb_1 \implies Hab = Ha_1b_1.$$

Sia $aa_1^{-1} =: h_1 \in H$, $bb_1^{-1} =: h_2 \in H$. Allora

$$ab(a_1b_1)^{-1} = a(bb_1^{-1})a_1^{-1} = ah_2a_1^{-1}.$$

Poiché $H \trianglelefteq G$, $ah_2 = h_3a$, $\exists h_3 \in H$. Allora

$$ab(a_1b_1)^{-1} = ah_2a_1^{-1} = h_3a_1a_1^{-1} = h_3h_1 \in H.$$

Dunque $Hab = Ha_1b_1$.

In base alla **Prop. 4** e a quanto osservato prima della **Def. 2**, se $H \trianglelefteq G$ $(\mathfrak{L}_d(H), \cdot)$ è un gruppo. Vale la seguente definizione.

Definizione 4. Se H è un sottogruppo normale di G , il gruppo $(\mathfrak{L}_d(H), \cdot)$ viene denotato $(G/H, \cdot)$ e chiamato *gruppo quoziante di G modulo H* .

[Si noti che, essendo $H \trianglelefteq G$, $\mathfrak{L}_d(H) = \mathfrak{L}_s(H)$ e quindi gli elementi del gruppo quoziante possono anche essere scritti usando i laterali sinistri in luogo dei destri].

Se $H \trianglelefteq G$, è definita l'applicazione

$$\pi : G \rightarrow G/H \text{ tale che } \pi(x) = Hx, \quad \forall x \in G.$$

Tale applicazione è ovviamente suriettiva ed è un omomorfismo di gruppi. Infatti

$$\pi(xy) = Hxy = Hx \cdot Hy = \pi(x) \cdot \pi(y), \quad \forall x, y \in G.$$

È chiamata *proiezione canonica di G sul gruppo quoziante G/H* .

Come già osservato, se G è abeliano ogni sottogruppo H è normale e dunque è definito G/H . Tra i gruppi non abeliani, quali sottogruppi sono normali? Lo sono certamente i sottogruppi di indice 2 ed i nuclei di omomorfismi tra gruppi, come ora dimostriamo nelle due seguenti proposizioni.

Proposizione 5. Sia H un sottogruppo di G . Se H ha indice 2 in G , H è normale in G .

Dim. Poiché H ha indice 2, $|\mathfrak{L}_d(H)| = |\mathfrak{L}_s(H)| = 2$. Risulta, $\forall a \in G - H$:

$$\mathfrak{L}_d(H) = \{H, Ha\}, \quad \mathfrak{L}_s(H) = \{H, aH\}.$$

Dunque $G - H = Ha$ e $G - H = aH$. Pertanto $Ha = aH$ e quindi H è normale in G .

Proposizione 6. Il nucleo $\text{Ker}(f)$ di un omomorfismo di gruppi $f : G \rightarrow G'$ è un sottogruppo normale in G .

Dim. Verifichiamo la seguente inclusione:

$$(\bullet) \quad x(Ker(f))x^{-1} \subseteq Ker(f), \quad \forall x \in G.$$

Infatti, $\forall a \in Ker(f)$, si ha:

$$f(xax^{-1}) = f(x)f(a)f(x^{-1}) = f(x)1'f(x)^{-1} = 1'$$

e dunque $xax^{-1} \in Ker(f)$.

Da (\bullet) , moltiplicando a destra per x :

$$x(Ker(f)) \subseteq (Ker(f))x.$$

L'inclusione (\bullet) (applicata a x^{-1}) fornisce: $x^{-1}(Ker(f))x \subseteq Ker(f)$. Moltiplicando a sinistra per x si ottiene:

$$(Ker(f))x \subseteq x(Ker(f)).$$

Dunque $(Ker(f))x = x(Ker(f))$, $\forall x \in G$, cioè $Ker(f) \trianglelefteq G$.

Dalla **Prop. 6** segue facilmente il seguente importante risultato, noto come *teorema fondamentale di omomorfismo* (per i gruppi).

Teorema 1. Ogni omomorfismo di gruppi $f : G \rightarrow G'$ induce l'isomorfismo (di gruppi)

$$\varphi : G/Ker(f) \rightarrow Im(f),$$

così definito: $\varphi(xKer(f)) = f(x)$, $\forall xKer(f) \in G/Ker(f)$.

Dim. Il primo fatto da verificare è che l'applicazione φ è ben definita: se cioè $xKer(f) = yKer(f)$, allora $f(x) = f(y)$. Infatti:

$$xKer(f) = yKer(f) \implies y^{-1}x \in Ker(f) \implies f(y^{-1}x) = 1' \implies f(y)^{-1}f(x) = 1' \implies f(x) = f(y).$$

Ora verifichiamo che φ è biiettiva. Per provare che φ è iniettiva bisogna verificare che

$$\varphi(xKer(f)) = \varphi(yKer(f)) \implies xKer(f) = yKer(f).$$

Basta leggere le precedenti implicazioni in senso inverso. Infatti: $\varphi(xKer(f)) = \varphi(yKer(f)) \implies f(x) = f(y) \implies 1' = f(x)^{-1}f(y) = f(x^{-1}y) \implies x^{-1}y \in Ker(f) \implies xKer(f) = yKer(f)$.

Inoltre φ è suriettiva. Infatti, $\forall f(x) \in Im(f)$, si ha: $\varphi(xKer(f)) = f(x)$.

Infine verifichiamo che φ è un omomorfismo di gruppi. Infatti:

$$\varphi(xKer(f) \cdot yKer(f)) = \varphi(xyKer(f)) = f(xy) = f(x)f(y) = \varphi(xKer(f)) \cdot \varphi(yKer(f))$$

Il teorema fondamentale di omomorfismo ci consente di decomporre ogni omomorfismo di gruppi $f : G \rightarrow G'$ nella composizione di tre omomorfismi, cioè nella forma

$$f = i \circ \varphi \circ \pi,$$

dove $\pi : G \rightarrow G/Ker(f)$ è la proiezione canonica, $\varphi : G/Ker(f) \rightarrow Im(f)$ è l'isomorfismo del **Teor. 1** ed $i : Im(f) \rightarrow G'$ è l'applicazione canonica di inclusione [e si tratta di un monomorfismo]. Risulta infatti:

$$(i \circ \varphi \circ \pi)(x) = i \circ \varphi(\pi(x)) = i(\varphi(xKer(f))) = i(f(x)) = f(x), \quad \forall x \in G.$$

La precedente uguaglianza di applicazioni può essere visualizzata nel seguente diagramma di omomorfismi di gruppi.

$$\begin{array}{ccc} G & \xrightarrow{f} & G' \\ \pi \downarrow & & \uparrow i \\ G/Ker(f) & \xrightarrow[\cong]{\varphi} & Im(f) \end{array} \quad \begin{array}{ccc} x & \xrightarrow{f} & f(x) \\ \pi \downarrow & & \uparrow i \\ xKer(f) & \xrightarrow[\cong]{\varphi} & f(x) \end{array}$$

tale che, $\forall x \in G$,

Si usa dire che *tale diagramma è commutativo*, con ciò intendendo che per passare da G a G' è indifferente quale percorso si segua (cioè f oppure $i \circ \varphi \circ \pi$).

Si noti che nel **Cap. 1.3, Prop. 6** avevamo ottenuto la versione 'insiemistica' di questo risultato.

Un esempio. Sia $f : \mathbf{Z} \rightarrow \mathbf{Z}_n$ l'applicazione così definita:

$$f(k) = \bar{k} \in \mathbf{Z}_n, \quad \forall k \in \mathbf{Z}.$$

f è ovviamente un omomorfismo suriettivo da $(\mathbf{Z}, +)$ a $(\mathbf{Z}_n, +)$. Ha quindi immagine $Im(f) = \mathbf{Z}_n$, mentre il nucleo è

$$Ker(f) = \{k \in \mathbf{Z} \mid \bar{k} = \bar{0}\} = \{k \in \mathbf{Z} \mid k \equiv 0 \pmod{n}\} = n\mathbf{Z}.$$

Dal teorema fondamentale di omomorfismo,

$$\mathbf{Z}/n\mathbf{Z} \cong \mathbf{Z}_n$$

[e l'isomorfismo φ trasforma il laterale $k + n\mathbf{Z}$ nella classe resto \bar{k} , $\forall k \in \mathbf{Z}$].

Osservazione 3. Il teorema fondamentale di omomorfismo tra gruppi non è che il primo teorema fondamentale tra strutture algebriche. Sussistono anche analoghi teoremi fondamentali di omomorfismo tra anelli e tra spazi vettoriali. Rinviamo per questi approfondimenti a www.mat.uniroma1.it/people/campanella, **Appunti di Algebra 2**, Cap. 1, § 3].

ESERCIZI PROPOSTI

5.2.1. Determinare le relazioni di equivalenza su G associate alla partizione dei laterali destri ed a quella dei laterali sinistri modulo un sottogruppo H di G .

5.2.2. Determinare i sottogruppi ciclici di $\mathbf{U}(\mathbf{Z}_{15})$ e dire se tale gruppo è ciclico.

5.2.3. Verificare che, a meno di isomorfismi, esistono due soli gruppi di ordine 6: C_6 e S_3 .

5.2.4. Verificare che $\mathcal{U}(\mathbf{Z}_{16})/\langle \bar{7} \rangle$ è un gruppo ciclico di ordine 4.

5.2.5. Il sottogruppo ciclico $\langle (1 2 3 4) \rangle$ è normale in S_4 ?

5.2.6. Sia $\mathcal{H} = \{A \in \mathbf{GL}_n(K) \mid \det(A) = 1\}$. Verificare che è un sottogruppo normale di $\mathbf{GL}_n(K)$ e determinare il gruppo quoziante $\mathbf{GL}_n(K)/\mathcal{H}$.

5.2.7. Sia $\mathcal{H} = \{A \in \mathbf{GL}_2(\mathbf{Z}_5) \mid \det(A) = \bar{1} \text{ oppure } \det(A) = \bar{4}\}$. Verificare che \mathcal{H} è un sottogruppo normale di $\mathbf{GL}_2(\mathbf{Z}_5)$ e che il gruppo quoziante $\mathbf{GL}_2(\mathbf{Z}_5)/\mathcal{H}$ è ciclico. Indicarne poi un generatore.

5.2.8. Verificare che se G è un gruppo ciclico ed H è un suo sottogruppo, il gruppo quoziante G/H è ciclico.

5.2.9. Si consideri in $(\mathbf{Q}, +)$ il sottogruppo $(\mathbf{Z}, +)$ ed il gruppo quoziante $(\mathbf{Q}/\mathbf{Z}, +)$.

(i) Verificare che ogni elemento di \mathbf{Q}/\mathbf{Z} è rappresentabile con un razionale q tale che $0 \leq q < 1$.

(ii) Verificare che ogni elemento di \mathbf{Q}/\mathbf{Z} ha periodo finito e dedurne che \mathbf{Q}/\mathbf{Z} non è ciclico.

5.2.10. Nel gruppo moltiplicativo (\mathbf{C}^*, \cdot) si considerino i due sottoinsiemi

$$\mathbf{R}^{>0} \text{ (numeri reali positivi)}, \quad \mathcal{H} = \{z \in \mathbf{C}^* \mid |z| = 1\}.$$

(i) Verificare che $\mathbf{R}^{>0}$ ed \mathcal{H} sono sottogruppi di (\mathbf{C}^*, \cdot) .

(ii) Dimostrare che il gruppo quoziante $\mathbf{C}^*/_{\mathbf{R}^{>0}}$ è isomorfo a \mathcal{H} .

ESERCIZI PROPOSTI

Capitolo 1

1.1.1. Sono assegnati tre insiemi A, B, C .

- (i) Verificare che $A - (B - C) = (A - B) \cup (A \cap C)$.
- (ii) Verificare che $(A - B) - C = A - (B \cup C)$.
- (iii) Verificare che $(A - B) - C \subseteq A - (B - C)$ e che tale inclusione può essere propria.

Soluzione. (i) Si ha: $x \in A - (B - C) \iff x \in A \wedge (x \notin B - C)$.

Si osservi che $x \in B - C \iff x \in B \wedge x \notin C$. Negando tale equivalenza, si ottiene

$$x \notin B - C \iff x \notin B \vee x \in C.$$

Ne segue:

$$\begin{aligned} x \in A - (B - C) &\iff x \in A \wedge [x \notin B \vee x \in C] \iff \\ [x \in A \wedge x \notin B] \vee [x \in A \wedge x \in C] &\iff (x \in A - B) \vee (x \in A \cap C) \iff x \in (A - B) \cup (A \cap C). \end{aligned}$$

(ii) Si ha:

$$\begin{aligned} x \in (A - B) - C &\iff (x \in A - B) \wedge (x \notin C) \iff [x \in A \wedge x \notin B] \wedge (x \notin C) \iff \\ x \in A \wedge [x \notin B \wedge x \notin C] &\iff x \in A \wedge [x \notin B \cup C] \iff x \in A - (B \cup C). \end{aligned}$$

(iii) Da (ii) e dall'ovvia inclusione $B - C \subseteq B \cup C$ segue:

$$A - (B - C) \supseteq A - (B \cup C) = (A - B) - C.$$

Un esempio in cui tale inclusione è stretta può essere ottenuto ponendo $A = B = C \neq \emptyset$. Infatti si ha:

$$(A - A) - A = \emptyset - A = \emptyset, \text{ mentre } A - (A - A) = A - \emptyset = A \neq \emptyset.$$

* * *

1.1.2. Sono assegnati tre insiemi A, B, C .

- (i) Verificare che $(A \cup B) - C = (A - C) \cup (B - C)$ e che $(A \cap B) - C = (A - C) \cap (B - C)$.
- (ii) Verificare che $A \cap (B - C) = (A \cap B) \cap (A - C)$.
- (iii) Determinare un insieme T tale che $A \cup (B - C) = (A \cup B) \cap T$.

Soluzione. (i) Si ha:

$$\begin{aligned} x \in (A \cup B) - C &\iff (x \in A \cup B) \wedge x \notin C \iff [x \in A \vee x \in B] \wedge (x \notin C) \iff \\ [x \in A \wedge x \notin C] \vee [x \in B \wedge x \notin C] &\iff (x \in A - C) \vee (x \in B - C) \iff x \in (A - C) \cup (B - C). \end{aligned}$$

Si ha:

$$\begin{aligned} x \in (A \cap B) - C &\iff [x \in A \wedge x \in B] \wedge x \notin C \iff [x \in A \wedge x \notin C] \wedge [x \in B \wedge x \notin C] \iff \\ (x \in A - C) \wedge (x \in B - C) &\iff x \in (A - C) \cap (B - C). \end{aligned}$$

(ii) Si ha:

$$\begin{aligned} x \in A \cap (B - C) &\iff (x \in A) \wedge [x \in B \wedge x \notin C] \iff [x \in A \wedge x \in B] \wedge [x \in A \wedge x \notin C] \iff \\ x \in (A \cap B) \cap (A - C). & \end{aligned}$$

(iii) Si denoti con X un insieme contenente $A \cup B \cup C$. Risulta:

$$\begin{aligned} x \in A \cup (B - C) &\iff x \in A \vee [x \in B \wedge x \notin C] \iff [x \in A \vee x \in B] \wedge [x \in A \vee x \notin C] \iff \\ [x \in A \cup B] \wedge [x \in A \cup (X - C)] &\iff x \in (A \cup B) \cap T \text{ con } T = A \cup (X - C). \end{aligned}$$

Si può anche osservare che $T = X - (C - A)$. Infatti:

$$\begin{aligned} x \in X - (C - A) &\iff x \in X \wedge x \notin C - A \iff x \notin C - A \iff \\ \iff (x \in A) \vee (x \notin C) &\iff x \in A \cup (X - C)]. \end{aligned}$$

* * *

1.1.3. Tra i numeri naturali compresi tra 100 e 999, contare quelli che hanno esattamente due cifre uguali tra loro.

Soluzione. Denotiamo con X l'insieme dei naturali cercati. Tale insieme X si ripartisce nei tre seguenti sottoinsiemi (a due a due disgiunti):

$$\begin{aligned} X_1 &= \{abb, \text{ con } 1 \leq a \leq 9, 0 \leq b \leq 9, b \neq a\}, \\ X_2 &= \{bab, \text{ con } 1 \leq b \leq 9, 0 \leq a \leq 9, a \neq b\}, \\ X_3 &= \{bba, \text{ con } 1 \leq b \leq 9, 0 \leq a \leq 9, a \neq b\}. \end{aligned}$$

Ciascuno dei tre insiemi ha cardinalità 81. Infatti tutti e tre sono ottenuti scegliendo a e b in 9 modi indipendenti; dunque, in base al principio del prodotto, $|X_1| = |X_2| = |X_3| = 9 \cdot 9 = 81$.

Dal principio generalizzato della somma,

$$|X| = |X_1| + |X_2| + |X_3| = 81 \cdot 3 = 243.$$

* * *

1.1.4. Trovare il numero dei naturali compresi tra 100 e 999 formati da cifre non nulle e a due a due distinte.

Soluzione. Denotiamo con X l'insieme dei naturali cercati. Tale insieme si identifica con l'insieme di terne

$$\left\{ (a, b, c) \in \mathbf{N}^3, \text{ con } 1 \leq a \leq 9, 1 \leq b \leq 9, 1 \leq c \leq 9 \text{ e con } a \neq b \neq c \neq a \right\}.$$

Per $i = 1, \dots, 9$, consideriamo in X il sottoinsieme

$$X_i = \left\{ (i, b, c), \text{ con } 1 \leq b \leq 9, b \neq i \text{ e con } 1 \leq c \leq 9, c \neq i, c \neq b \right\}.$$

Poiché b può essere scelto in 8 modi diversi e c in 7 modi diversi (indipendenti dai primi), in base al principio del prodotto

$$|X_i| = 8 \cdot 7 = 56.$$

Essendo $\{X_1, X_2, \dots, X_9\}$ una partizione di X , si ha, per il principio generalizzato della somma,

$$|X| = |X_1| + |X_2| + \dots + |X_9| = 9 \cdot 56 = 504.$$

* * *

1.1.5. Contare i naturali tra 1 e 1000 che non sono divisibili né per 4, né per 5, né per 6.

Soluzione. Poniamo $I = \{n \in \mathbf{N} : 1 \leq n \leq 1000\}$. Denotiamo poi con A_4, A_5, A_6 gli insiemi di naturali di I che sono rispettivamente divisibili per 4, 5, 6. Dunque

$$A_4 = \{4, 8, 12, \dots\} = \{4t, \forall t = 1, \dots, [\frac{1000}{4}]\},$$

dove $[\frac{1000}{4}] = 250$ indica la parte intera del numero razionale $\frac{1000}{4}$. Analogamente:

$$A_5 = \{5, 10, 15, \dots\} = \{5t, \forall t = 1, \dots, [\frac{1000}{5}]\} = \{5t, \forall t = 1, \dots, 200\},$$

$$A_6 = \{6, 12, 18, \dots\} = \{6t, \forall t = 1, \dots, [\frac{1000}{6}]\} = \{6t, \forall t = 1, \dots, 166\}.$$

Si ha quindi: $|A_4| = 250$, $|A_5| = 200$, $|A_6| = 166$.

L'insieme X cercato è

$$X = (I - A_4) \cap (I - A_5) \cap (I - A_6) = I - (A_4 \cup A_5 \cup A_6).$$

Per ottenere $|X|$ basterà allora calcolare $|A_4 \cup A_5 \cup A_6|$. Per fare ciò ricorriamo al principio di inclusione - esclusione.

Si osservi che $A_4 \cap A_5$ è l'insieme degli $n \in I$ che sono divisibili sia per 4 che per 5, cioè che sono divisibili per il *mcm* di 4, 5 [che è 20]. Dunque

$$A_4 \cap A_5 = \{20t, \forall t = 1, \dots, [\frac{1000}{20}]\} = \{20t, \forall t = 1, \dots, 50\}.$$

Analogamente,

$$A_4 \cap A_6 = \{12t, \forall t = 1, \dots, [\frac{1000}{12}]\} = \{12t, \forall t = 1, \dots, 83\},$$

$$A_5 \cap A_6 = \{30t, \forall t = 1, \dots, [\frac{1000}{30}]\} = \{30t, \forall t = 1, \dots, 33\}.$$

Infine, osservato che il *mcm* tra 4, 5, 6 è 60, allora

$$A_4 \cap A_5 \cap A_6 = \{60t, \forall t = 1, \dots, [\frac{1000}{60}]\} = \{60t, \forall t = 1, \dots, 16\}.$$

Possiamo ora applicare il principio di inclusione - esclusione.

$$\begin{aligned} |A_4 \cup A_5 \cup A_6| &= |A_4| + |A_5| + |A_6| - |A_4 \cap A_5| - |A_4 \cap A_6| - |A_5 \cap A_6| + |A_4 \cap A_5 \cap A_6| = \\ &= 250 + 200 + 166 - 50 - 83 - 33 + 16 = 466. \end{aligned}$$

Dunque

$$|X| = |I - (A_4 \cup A_5 \cup A_6)| = 1000 - 466 = 534.$$

* * *

1.1.6. Sia $a \in \mathbf{R}$. Dimostrare che, $\forall n \geq 1$:

$$a^n - 1 = (a - 1)(a^{n-1} + a^{n-2} + \dots + a + 1).$$

Soluzione. Procediamo per induzione su $n \geq 1$.

Base induttiva. Risulta: $a^1 - 1 = (a - 1) \cdot 1$ [ovvio].

Passo induttivo. Sia $n \geq 2$ e sia $a^{n-1} - 1 = (a - 1)(a^{n-2} + \dots + a + 1)$. Bisogna verificare che $a^n - 1 = (a - 1)(a^{n-1} + a^{n-2} + \dots + a + 1)$. Infatti:

$$\begin{aligned} a^n - 1 &= a^n - a + a - 1 = a(a^{n-1} - 1) + (a - 1) = a(a - 1)(a^{n-2} + \dots + a + 1) + (a - 1) = \\ &= (a - 1)[a(a^{n-2} + \dots + a + 1) + 1] = (a - 1)(a^{n-1} + a^{n-2} + \dots + a + 1). \end{aligned}$$

* * *

1.1.7. Determinare tutte le possibili partizioni di un insieme X di cardinalità 3.

Soluzione. Si ponga $X = \{1, 2, 3\}$. X ammette ovviamente le due partizioni banali

$$\mathfrak{U}_1 = \{\{1, 2, 3\}\}, \quad \mathfrak{U}_2 = \{\{1\}, \{2\}, \{3\}\}.$$

Le altre partizioni sono necessariamente costituite da un sottoinsieme di cardinalità 1 e da uno di cardinalità 2 (disgiunti). Dunque sono le seguenti tre:

$$\mathfrak{U}_3 = \{\{1\}, \{2, 3\}\}, \quad \mathfrak{U}_4 = \{\{2\}, \{1, 3\}\}, \quad \mathfrak{U}_5 = \{\{3\}, \{1, 2\}\}.$$

* * *

1.2.1. Sia $f : A \rightarrow B$ un'applicazione. Siano A' e B' sottoinsiemi non vuoti rispettivamente di A e di B . Verificare che:

- (i) $f^{-1}(f(A')) \supseteq A'$. Se f è iniettiva, $f^{-1}(f(A')) = A'$.
- (ii) $f(f^{-1}(B')) \subseteq B'$. Se f è suriettiva, $f(f^{-1}(B')) = B'$.

Soluzione. (i) Risulta, $\forall a \in A$: $a \in f^{-1}(f(A')) \iff f(a) \in f(A')$. Se quindi $a \in A'$, allora $f(a) \in f(A')$ e quindi $a \in f^{-1}(f(A'))$.

Se poi f è iniettiva, basta verificare che $f^{-1}(f(A')) \subseteq A'$. Se infatti $a \in f^{-1}(f(A'))$, allora $f(a) = f(a')$, $\exists a' \in A'$. Ma poiché f è iniettiva, $a = a' \in A'$.

(ii) Risulta: $f(f^{-1}(B')) = \{f(a), \forall a \in A : f(a) \in B'\} = \text{Im } f \cap B' \subseteq B'$. Se poi f è suriettiva, allora $\text{Im } f = B$ e quindi $f(f^{-1}(B')) = B \cap B' = B'$.

* * *

1.2.2. Determinare due insiemi finiti A, B e due applicazioni $f : A \rightarrow B$, $g : B \rightarrow A$ tali che g ed f non sono biiettive, ma $g \circ f = \mathbf{1}_A$.

Soluzione. Poniamo ad esempio $A = \{a, b\}$ e $B = \{1, 2, 3\}$.

$$\text{Se } f : \begin{cases} a \rightarrow 1 \\ b \rightarrow 3 \end{cases} \text{ e } g : \begin{cases} 1 \rightarrow a \\ 2 \rightarrow a \\ 3 \rightarrow b, \end{cases} \text{ allora } g \circ f : \begin{cases} a \rightarrow a \\ b \rightarrow b. \end{cases}$$

* * *

1.2.3. Dati gli insiemi $A = \{a, b, c\}$ e $B = \{1, 2\}$, quante sono e come si possono scrivere le applicazioni suriettive da A a B ? E le applicazioni iniettive da B ad A ?

Soluzione. Le applicazioni da A a B possono identificarsi con terne di elementi di B [intendendo che i tre elementi di ogni terna sono le immagini rispettivamente di a, b, c]. Le possibili terne sono 8:

$$(1, 1, 1), (1, 2, 1), (1, 1, 2), (1, 2, 2), (2, 1, 1), (2, 2, 1), (2, 1, 2), (2, 2, 2).$$

Solo la prima e l'ultima non danno luogo ad applicazioni suriettive. Pertanto le applicazioni suriettive sono le restanti 6.

Le applicazioni da B ad A sono $2^4 = 16$ e si identificano con le coppie di elementi di A , cioè

$$\begin{aligned} & (a, a), (a, b), (a, c), (a, d), (b, a), (b, b), (b, c), (b, d), \\ & (c, a), (c, b), (c, c), (c, d), (d, a), (d, b), (d, c), (d, d). \end{aligned}$$

Di esse 12 sono formate da elementi distinti e quindi corrispondono alle applicazioni iniettive cercate. [Si ricordi del resto che le applicazioni iniettive sono $4 \cdot \dots \cdot (4 - 2 + 1) = 4 \cdot 3 = 12$].

* * *

1.2.4. Sia $f : \mathbf{R} \rightarrow \mathbf{R}$ tale che $f(x) = \sin(x)$, $\forall x \in \mathbf{R}$. Determinare $f^{-1}([- \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}])$, dove $[- \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}]$ è l'intervallo chiuso di estremi $-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}$.

Soluzione. Risulta:

$$\sin^{-1}([- \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}]) = \bigcup_{k \in \mathbf{Z}} [\frac{\pi}{4} - k, \frac{\pi}{4} + k].$$

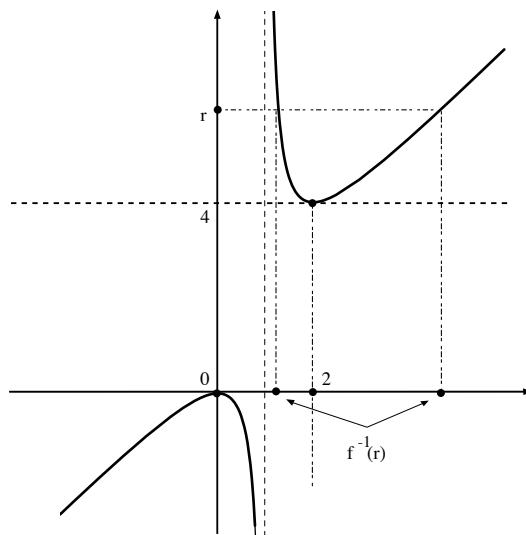
* * *

1.2.5. Sia $f : \mathbf{R} - \{1\} \rightarrow \mathbf{R}$ l'applicazione tale che $f(x) = \frac{x^2}{x-1}$, $\forall x \in \mathbf{R}$, $x \neq 1$. Per ogni $r \in \mathbf{R}$, determinare $f^{-1}(r)$.

Soluzione. Poiché $f^{-1}(r) = \{x \in \mathbf{R} \mid f(x) = r\}$, graficamente si risolve cercando le intersezioni tra il grafico della funzione e la retta $y = r$ (parallela all'asse x).

Nel caso in esame, $f(x) = r$ per $x \neq 1$ si scrive nella forma $x^2 = r(x-1)$, da cui $x^2 - rx + r = 0$. Tale equazione ammette soluzioni per $r \geq 4$ e $r \leq 0$ [infatti il discriminante è $\Delta = r(r-4)$]. Dunque

$$\begin{aligned} f^{-1}(r) &= \emptyset, \text{ se } 0 < r < 4; \quad f^{-1}(0) = \{0\}; \quad f^{-1}(4) = \{2\}; \\ f^{-1}(r) &= \left\{ \frac{1}{2}(r - \sqrt{r(r-4)}) \right\}, \text{ se } r < 0 \text{ o } r > 4. \end{aligned}$$



* * *

1.2.6. Verificare, per induzione su $n \geq 0$, che per ogni $x, y \in \mathbf{R}$ risulta:

$$(x+y)^n = \sum_{k=0}^n \binom{n}{k} x^{n-k} y^k.$$

Soluzione. Per induzione su $n \geq 0$.

Base induttiva. Risulta: $(x+y)^0 = \sum_{k=0}^0 \binom{0}{k} x^{0-k} y^k = 1$. Infatti: $(x+y)^0 = 1$, $\sum_{k=0}^0 \binom{0}{k} x^{0-k} y^k = \binom{0}{0} x^0 y^0 = 1$.

Passo induttivo. Sia $n \geq 1$. Per ipotesi induttiva, sia $(x+y)^{n-1} = \sum_{k=0}^{n-1} \binom{n-1}{k} x^{n-1-k} y^k$. Bisogna verificare che

$$(x+y)^n = \sum_{k=0}^n \binom{n}{k} x^{n-k} y^k$$

Infatti si ha:

$$\begin{aligned} (x+y)^n &= (x+y)(x+y)^{n-1} = (x+y) \left(\sum_{k=0}^{n-1} \binom{n-1}{k} x^{n-1-k} y^k \right) = \\ &= \sum_{k=0}^{n-1} \binom{n-1}{k} x^{n-k} y^k + \sum_{k=0}^{n-1} \binom{n-1}{k} x^{n-1-k} y^{k+1} = \quad [\text{ponendo } h = k+1] \\ &= \sum_{k=0}^{n-1} \binom{n-1}{k} x^{n-k} y^k + \sum_{h=1}^n \binom{n-1}{h-1} x^{n-h} y^h = \quad [\text{ponendo } k = h] \\ &= \binom{n-1}{0} x^n + \sum_{k=1}^{n-1} \binom{n-1}{k} x^{n-k} y^k + \sum_{k=1}^{n-1} \binom{n-1}{k-1} x^{n-k} y^k + \binom{n-1}{n-1} x^0 y^n = \\ &= x^n + \sum_{k=1}^{n-1} [\binom{n-1}{k} + \binom{n-1}{k-1}] x^{n-k} y^k + y^n = \\ &= x^n + \sum_{k=1}^{n-1} \binom{n}{k} x^{n-k} y^k + y^n = \sum_{k=0}^n \binom{n}{k} x^{n-k} y^k. \end{aligned}$$

* * *

1.2.7. Per ogni $n \geq 1$, si ponga $S_n := 1 + 2 + 3 + \dots + n = \sum_{k=1}^n k$. Verificare che risulta:

$$S_n = \binom{n+1}{2}, \quad \forall n \geq 1.$$

Soluzione. Per induzione su $n \geq 1$.

Base induttiva. Risulta: $S_1 = \binom{1+1}{2}$ [infatti $S_1 = 1$, $\binom{1+1}{2} = \binom{2}{2} = 1$].

Passo induttivo. Sia $n \geq 1$ e sia $S_n = \binom{n+1}{2}$. Bisogna verificare che $S_{n+1} = \binom{n+2}{2}$.

Infatti $S_{n+1} = S_n + (n+1) = \binom{n+1}{2} + (n+1) = \binom{n+1}{2} + \binom{n+1}{1} = \binom{n+2}{2}$ [cfr. **Prop. 2.6**].

Nota. Tale risultato può essere dimostrato anche senza induzione. Infatti:

$$\begin{aligned} 2S_n &= (1 + 2 + \dots + (n-1) + n) + (n + (n-1) + \dots + 2 + 1) = \\ &= [1 + n] + [2 + (n-1)] + [3 + (n-2)] + \dots + [(n-1) + 2] + [n + 1] = \\ &= \sum_{k=1}^n [k + (n-k+1)] = \sum_{k=1}^n (n+1) = n(n+1). \end{aligned}$$

Dunque $S_n = \frac{n(n+1)}{2} = \binom{n+1}{2}$.

* * *

1.2.8. Per ogni $n \geq 1$, si ponga

$$\Sigma_n := 1 + 3 + 5 + \dots + (2n-1) \quad [\text{somma dei primi } n \text{ numeri dispari}].$$

Verificare che risulta: $\Sigma_n = n^2$, $\forall n \geq 1$.

Soluzione. Presentiamo due dimostrazioni: la prima usa l'induzione l'esercizio precedente; la seconda procede direttamente per induzione.

(1). Risulta:

$$\Sigma_n = S_{2n} - \sum_{k=0}^n 2k = S_{2n} - 2S_n = \binom{2n+1}{2} - 2\binom{n+1}{2} = \frac{(2n+1)2n}{2} - 2\frac{(n+1)n}{2} = n^2$$

(2). Per induzione su $n \geq 1$.

Base induttiva. Risulta: $\Sigma_1 = 1^2$ [ovvio].

Passo induttivo. Sia $n \geq 1$ e sia $\Sigma_n = n^2$. Dobbiamo verificare che $\Sigma_{n+1} = (n+1)^2$. Infatti:

$$\Sigma_{n+1} = \Sigma_n + (2n+1) = n^2 + (2n+1) = (n+1)^2.$$

* * *

1.3.1. Sia $A = \{a, b, c\}$ un insieme (di cardinalità 3). Sia ρ la relazione su A avente grafico

$$\mathfrak{R} = \{(a, a), (a, b), (a, c), (b, b), (b, c), (c, c)\}.$$

Verificare se ρ è una relazione d'ordine totale su A .

Soluzione. La relazione ρ si scrive in forma cartesiana nel seguente modo:

ρ	a	b	c
a	1	1	1
b	0	1	1
c	0	0	1

Poiché la diagonale Δ di A è contenuta in \mathfrak{R} , ρ è riflessiva. Poiché, se $x \neq y$, $x\rho y$ e $y\rho x$ non sono mai simultaneamente verificate, ρ è antisimmetrica. Poiché inoltre, se $x \neq y$, $x \not\rho y$ e $y \not\rho x$ non sono mai simultaneamente verificate, ρ è totale.

Resta da verificare che ρ è transitiva. Si tratta di verificare che, $\forall x, y, z \in A$ (a due a due distinti), risulta:

$$x\rho y, y\rho z \implies x\rho z.$$

Si hanno per le terne distinte (x, y, z) le seguenti sei possibilità

x	y	z	$x\rho y$	$y\rho z$	$x\rho z$
a	b	c	sì	sì	sì
a	c	b	sì	no	/
b	a	c	no	/	/
b	c	a	sì	no	/
c	a	b	no	/	/
c	b	a	no	/	/

Come si osserva, l'unico caso in cui $x\rho y$ e $y\rho z$ si ha per $(x, y, z) = (a, b, c)$ ed in tal caso $x\rho z$. Pertanto ρ è transitiva e quindi è una relazione d'ordine totale su A .

* * *

1.3.2. Dimostrare il principio generalizzato del prodotto: *se $k \geq 2$ e A_1, A_2, \dots, A_k sono insiemi finiti, si ha:*

$$|A_1 \times A_2 \times \dots \times A_k| = |A_1| \cdot |A_2| \cdot \dots \cdot |A_k|.$$

Soluzione. Si procede per induzione su $k \geq 2$.

Base induttiva. Risulta: $|A_1 \times A_2| = |A_1| \cdot |A_2|$ [è esattamente il principio del prodotto].

Passo induttivo. Sia ora $n \geq 2$ e si assuma vero l'asserto per ogni prodotto cartesiano con n fattori. Si consideri $A_1 \times A_2 \times \dots \times A_{n+1}$. Risulta subito che $A_1 \times A_2 \times \dots \times A_{n+1}$ è in corrispondenza biunivoca con $A_1 \times (A_2 \times \dots \times A_{n+1})$ [tramite la biiezione $(a_1, a_2, \dots, a_{n+1}) \rightarrow (a_1, (a_2, \dots, a_{n+1}))$]. Allora, usando l'ipotesi induttiva relativamente a $A_2 \times \dots \times A_{n+1}$:

$$|A_1 \times A_2 \times \dots \times A_{n+1}| = |A_1| \cdot |A_2 \times \dots \times A_{n+1}| = |A_1| \cdot (|A_2| \cdot \dots \cdot |A_{n+1}|) = |A_1| \cdot \dots \cdot |A_{n+1}|.$$

Il passo induttivo è così dimostrato. Ne segue il principio generalizzato del prodotto.

* * *

1.3.3. Determinare una biiezione tra \mathbf{N} e \mathbf{Z} [ciò dimostra che $|\mathbf{Z}| = |\mathbf{N}|$, cioè che \mathbf{Z} è numerabile].

Soluzione. Si consideri l'applicazione $f : \mathbf{N} \rightarrow \mathbf{Z}$ tale che:

$$\begin{aligned}
 0 &\rightarrow 0 \\
 1 &\rightarrow -1 \\
 2 &\rightarrow 1 \\
 3 &\rightarrow -2 \\
 4 &\rightarrow 2 \\
 5 &\rightarrow -3 \\
 6 &\rightarrow 3 \\
 \dots
 \end{aligned}$$

Tale applicazione è quindi così definita:

$$f(n) = \frac{n}{2} \text{ se } n \text{ è pari;} \quad f(n) = -\frac{n+1}{2} \text{ se } n \text{ è dispari.}$$

Per provare che f è biettiva, ne costruiamo l'inversa $g : \mathbf{Z} \rightarrow \mathbf{N}$. Dai dati precedenti osserviamo che g deve operare come segue:

$$\begin{aligned}
 0 &\rightarrow 0 \\
 -1 &\rightarrow 1 \\
 1 &\rightarrow 2 \\
 -2 &\rightarrow 3 \\
 2 &\rightarrow 4 \\
 -3 &\rightarrow 5 \\
 3 &\rightarrow 6 \\
 \dots
 \end{aligned}$$

Quindi g è così definita:

$$g(n) = 2n \text{ se } n \geq 0; \quad g(n) = -1 - 2n \text{ se } n < 0.$$

Si tratta ora di verificare che $g \circ f = \mathbf{1}_{\mathbf{N}}$ e che $f \circ g = \mathbf{1}_{\mathbf{Z}}$. Ci limitiamo alla prima verifica, lasciando la seconda al lettore. Risulta:

$$\begin{aligned}
 \text{se } n \text{ è pari: } (g \circ f)(n) &= g\left(\frac{n}{2}\right) = 2 \cdot \frac{n}{2} = n; \\
 \text{se } n \text{ è dispari: } (g \circ f)(n) &= g\left(-\frac{n+1}{2}\right) = -1 - 2\left(-\frac{n+1}{2}\right) = -1 + n + 1 = n.
 \end{aligned}$$

Dunque $(g \circ f)(n) = n, \forall n \in \mathbf{N}$.

* * *

1.4.1. Sia $\sigma = (1 3 2 4)(5 6) \in \mathbf{S}_6$. Determinare il minimo intero $k \geq 1$ tale che $\sigma^k = \mathbf{1}_x$.

Soluzione. Si ha:

$$\sigma^2 = (1 2)(3 4), \quad \sigma^3 = \sigma^2 \sigma = (1 4 2 3)(5 6), \quad \sigma^4 = \sigma^2 \sigma^2 = \mathbf{1}_x.$$

Pertanto $k = 4$.

* * *

1.4.2. Stesso esercizio con $\sigma = (1 3 2)(4 5) \in \mathbf{S}_5$.

Soluzione. Si ha:

$$\sigma^2 = (1 2 3), \quad \sigma^3 = \sigma^2 \sigma = (4 5), \quad \sigma^4 = \sigma^2 \sigma^2 = (1 3 2), \quad \sigma^5 = \sigma^2 \sigma^3 = (1 2 3)(4 5), \quad \sigma^6 = \sigma^3 \sigma^3 = \mathbf{1}_x.$$

Pertanto $k = 6$.

N.B. Il numero k cercato è in entrambi i casi il minimo comune multiplo delle lunghezze dei cicli disgiunti di σ .

* * *

1.4.3. Scrivere tutte le permutazioni di \mathbf{S}_5 che contengono il ciclo $(1 2)$.

Soluzione. In \mathbf{S}_5 le permutazioni che contengono $(1 2)$ si ottengono aggiungendo il ciclo $(1 2)$ alle permutazioni di $\{3, 4, 5\}$. Dunque sono le sei permutazioni:

$$(1 2), (1 2)(3 4), (1 2)(3 5), (1 2)(4 5), (1 2)(3 4 5), (1 2)(3 5 4).$$

* * *

1.4.4. Assegnate le permutazioni $\sigma_1 = (1\ 3\ 4\ 2)$, $\sigma_2 = (2\ 5)(3\ 4) \in S_5$, determinare la permutazione $\tau \in S_5$ tale che $\sigma_2 = \tau \sigma_1$.

Soluzione. Da $\sigma_2 = \tau \sigma_1$ segue subito che $\tau = \sigma_2 \sigma_1^{-1}$. Pertanto

$$\tau = (2\ 5)(3\ 4)(1\ 3\ 4\ 2)^{-1} = (2\ 5)(3\ 4)(1\ 2\ 4\ 3) = (1\ 2\ 5\ 4).$$

* * *

1.4.5. Scrivere la "tavola pitagorica" di S_3 , cioè la tavola 6×6 formata da tutti i prodotti $\sigma_i \sigma_j$, al variare di σ_i, σ_j in S_3 .

Soluzione.

	(1)	(12)	(13)	(23)	(123)	(132)
(1)	(1)	(12)	(13)	(23)	(123)	(132)
(12)	(12)	(1)	(123)	(132)	(13)	(23)
(13)	(13)	(132)	(1)	(123)	(23)	(12)
(23)	(23)	(123)	(132)	(1)	(12)	(13)
(123)	(123)	(23)	(12)	(13)	(132)	(1)
(132)	(132)	(13)	(23)	(12)	(1)	(123)

* * *

1.4.6. (i) Verificare che se $\sigma \in S_n$ è una permutazione di classe dispari, non esiste alcuna permutazione $\alpha \in S_n$ tale che $\alpha^2 = \sigma$.

(ii) Determinare $\alpha \in S_6$ tale che $\alpha^2 = (1\ 2\ 3)(4\ 5\ 6)$.

(iii) Spiegare perché non esiste $\alpha \in S_6$ tale che $\alpha^2 = (1\ 2)(3\ 4\ 5\ 6)$.

Soluzione. (i) Se $\alpha = \gamma_1 \dots \gamma_t$ (prodotto di cicli a due a due disgiunti), allora $\alpha^2 = \gamma_1^2 \dots \gamma_t^2$. Se γ_i è prodotto di t_i trasposizioni, γ_i^2 è prodotto di $2t_i$ trasposizioni e quindi è di classe pari. Ne segue che α^2 è di classe pari e dunque $\alpha^2 \neq \sigma$.

(ii) Basta considerare il 6-ciclo $\alpha = (1\ 4\ 2\ 5\ 3\ 6)$. Risulta: $\alpha^2 = (1\ 2\ 3)(4\ 5\ 6)$.

(iii) La permutazione $\sigma = (1\ 2)(3\ 4\ 5\ 6)$ è di classe pari e quindi non si può applicare la (i). Assumiamo per assurdo che sia

$$\alpha^2 = \sigma, \text{ con } \alpha = \gamma_1 \dots \gamma_t \text{ (prodotto di cicli disgiunti)}.$$

Allora $\sigma = \gamma_1^2 \dots \gamma_t^2$ e le permutazioni γ_i^2 sono a due a due disgiunte. Poiché σ è prodotto di un 2-ciclo e di un 4-ciclo disgiunti, una delle permutazioni γ_i^2 coincide con tale 2-ciclo e quindi è di classe dispari: assurdo.

* * *

1.4.7. (i) Verificare che $\forall \sigma \in S_n \exists k \in N, k \geq 1$ tale che $\sigma^k = \mathbf{1}_x$.

(ii) Verificare che $\forall \sigma, \tau \in S_n$ l'equazione (di primo grado) $\sigma X = \tau$ ammette una ed una sola soluzione (in S_n).

Soluzione. (i) Poiché S_n è un insieme finito, $\exists h, k \geq 1, h \neq k$ tali che $\sigma^h = \sigma^k$. Assumiamo $h < k$. Allora $\sigma^h \sigma^{-h} = \sigma^k \sigma^{-h}$, cioè $\mathbf{1}_x = \sigma^0 = \sigma^{k-h}$, con $k - h > 0$.

(ii) Sia $x = \sigma^{-1}\tau$. x è soluzione dell'equazione $\sigma X = \tau$. Infatti:

$$\sigma x = \sigma(\sigma^{-1}\tau) = (\sigma\sigma^{-1})\tau = \mathbf{1}_x \tau = \tau.$$

Se poi $y \in S_n$ è un'altra soluzione dell'equazione $\sigma X = \tau$, allora $\sigma y = \tau = \sigma x$. Ne segue che $\sigma^{-1}(\sigma y) = \sigma^{-1}(\sigma x)$, da cui $y = x$.

* * *

1.4.8. Determinare per quali $\sigma \in S_4$ l'equazione $X^2 = \sigma$ è risolubile.

Soluzione. Se σ è di classe dispari, l'equazione $X^2 = \sigma$ non è mai risolubile, in quanto, $\forall \alpha \in S_4$, α^2 è sempre di classe pari. Verifichiamo tale affermazione: se $\alpha = \gamma_1 \dots \gamma_t$ (prodotto di cicli disgiunti), allora

$$\alpha^2 = \gamma_1^2 \dots \gamma_t^2$$

e ogni γ_i^2 è prodotto di un numero pari di trasposizioni.

Studieremo quindi la risolubilità di $X^2 = \sigma$, per ogni permutazione σ di classe pari. Le permutazioni di classe pari di S_4 hanno una delle seguenti strutture cicliche:

$$(-), \quad (- -)(- -), \quad (- - -)$$

[cioè sono l'1-ciclo (1), o coppie di due 2-cicli disgiunti $(ab)(cd)$ o 3-cicli (abc)].

- L'equazione $X^2 = (1)$ è ovviamente risolubile. Infatti ha soluzione (1). Ma ha anche altre soluzioni: tutti i 2-cicli $(- -)$ e tutti i prodotti di due 2-cicli disgiunti $(- -)(- -)$.

- L'equazione $X^2 = (ab)(cd)$ è risolubile. Infatti, come subito si verifica, ammette le due soluzioni $(acbd)$ e $(adb c)$.

- L'equazione $X^2 = (abc)$ è risolubile. Infatti ha soluzione (acb) , come subito si verifica.

Concludendo, l'equazione $X^2 = \sigma$ è risolubile $\iff \sigma$ è una permutazione di classe pari.

* * *

1.4.9. In S_5 sono assegnati un 3-ciclo σ ed un 2-ciclo τ , disgiunti tra loro. Sia H l'insieme formato dalle permutazioni di S_5 ottenibili come prodotti finiti di σ e di τ . Determinare le permutazioni di H e verificare se H è un sottogruppo di S_5 .

Soluzione. Poiché σ e τ sono cicli disgiunti, commutano. Ogni permutazione di H è quindi della forma $\sigma^h \tau^k$, per ogni $h, k \geq 0$. Inoltre si ha: $\sigma^3 = \tau^2 = (1)$. Pertanto:

$$H = \{(1), \sigma, \sigma^2, \tau, \sigma\tau, \sigma^2\tau\}.$$

Verifichiamo che H è un sottogruppo di S_5 . Evidentemente $(1) \in H$; inoltre $HH \subseteq H$ [in quanto H è l'insieme di tutti i prodotti finiti dei suoi elementi]; infine verifichiamo che $H^{-1} \subseteq H$. Infatti

$$\sigma^{-1} = \sigma^2, \quad (\sigma^2)^{-1} = \sigma, \quad \tau^{-1} = \tau, \quad (\sigma\tau)^{-1} = \sigma^2\tau, \quad (\sigma^2\tau)^{-1} = \sigma\tau.$$

N.B. Cosa avviene se σ e τ non sono disgiunti?

Si hanno due possibilità: o σ e τ hanno due interi in comune o ne hanno uno solo. Nel primo caso H è contenuto in un S_3 , mentre nel secondo caso in un S_4 .

Nel primo caso prendiamo ad esempio in considerazione le permutazioni $\sigma = (123)$, $\tau = (13)$. Si verifica in tal caso che

$$H = \{(1), \sigma, \sigma^2, \tau, \sigma\tau, \tau\sigma\}.$$

Nel secondo caso prendiamo invece in considerazione le permutazioni $\sigma = (123)$, $\tau = (14)$. Si verifica in tal caso che H contiene ad esempio le seguenti 13 permutazioni (a due a due distinte)

$$(1), \sigma, \sigma^2, \tau, \sigma\tau, \tau\sigma, \sigma^2\tau, \tau\sigma^2, (\sigma\tau)^2, (\tau\sigma)^2, (\sigma^2\tau)^2, (\tau\sigma^2)^2, \sigma\tau\sigma.$$

Non abbiamo ancora gli strumenti tecnici per affermare che H ha in tal caso ordine 24 ed è isomorfo a S_4 .

* * *

1.4.10. (i) Quanti sono i 3-cicli di S_6 ?

(ii) Quante sono le permutazioni di S_6 che sono prodotto di due 3-cicli disgiunti?

Soluzione. (i) Le applicazioni iniettive da $\{1, 2, 3\}$ a $\{1, 2, 3, 4, 5, 6\}$ corrispondono biunivocamente alle terne di interi distinti compresi tra 1 e 6. Quante sono tali applicazioni?

Sia f una siffatta applicazione. L'intero $f(1)$ può essere scelto in 6 modi diversi, l'intero $f(2)$ in 5 modi diversi e l'intero $f(3)$ in 4 modi diversi. Complessivamente si hanno $6 \cdot 5 \cdot 4 = 120$ applicazioni iniettive, cioè 120 terne di interi distinti. Ogni 3-ciclo corrisponde a tre di tali terne. Infatti ogni 3-ciclo (abc) si scrive in tre modi:

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

Dunque i 3-cicli di S_6 sono $\frac{120}{3} = 40$.

(ii) Per ottenere tutte le permutazioni di S_6 che siano prodotto di due 3-cicli disgiunti, cominciamo con l'osservare che i due 3-cicli commutano e che l'elemento 1 si trova necessariamente in uno dei due 3-cicli. Assumeremo quindi che le permutazioni cercate si scrivano nella forma

$$(1 a b)(c d e).$$

Determiniamo i 3-cicli $(1 a b)$. Se poniamo $2 \leq a < b$, otteniamo le seguenti coppie (a, b) : $(2, 3), (2, 4), (2, 5), (2, 6), (3, 4), (3, 5), (3, 6), (4, 5), (4, 6), (5, 6)$. Conseguentemente abbiamo i dieci 3-cicli:

$$(123), (124), (125), (126), (134), (135), (136), (145), (146), (156).$$

Ad essi dobbiamo aggiungere i rispettivi inversi $(1 b a)$, cioè

$$(132), (142), (152), (162), (143), (154), (163), (154), (164), (165).$$

Dunque abbiamo ottenuto 20 3-cicli di tipo $(1 a b)$. Ognuno di essi va ora moltiplicato per i possibili 3-cicli "supplementari", che sono due e cioè: $(c d e)$ e $(c e d)$. Si hanno complessivamente 40 permutazioni del tipo cercato.

* * *

ESERCIZI PROPOSTI

Capitolo 2

2.1.1. Sia (G, \cdot) un gruppo e siano $a, b \in G$. Verificare che, se $ab = 1_G$, allora $ba = 1_G$ [cioè se b è "inverso a destra" di a , è anche "inverso a sinistra" di a].

Soluzione. Si ha:

$$ab = 1_G \implies a^{-1}(ab) = a^{-1}1_G = a^{-1} \implies (a^{-1}a)b = a^{-1} \implies 1_G b = a^{-1} \implies b = a^{-1}.$$

Ne segue che $ba = a^{-1}a = 1_G$.

* * *

2.1.2. Si consideri in \mathbf{S}_5 il sottoinsieme

$$\Sigma = \Sigma_{\{1, 2\}} := \{\sigma \in \mathbf{S}_5 : \sigma(\{1, 2\}) = \{1, 2\}\}.$$

Verificare che Σ è un sottogruppo di \mathbf{S}_5 e scriverne gli elementi.

Soluzione. Σ è l'insieme delle permutazioni di \mathbf{S}_5 che trasformano il sottoinsieme $\{1, 2\}$ in sé e quindi necessariamente anche $\{3, 4, 5\}$ in sé.

Ovviamente $\mathbf{1}_x$ fissa $\{1, 2\}$ (elemento per elemento) e dunque $\mathbf{1}_x \in \Sigma$. Se poi $\sigma, \tau \in \Sigma$, anche $\sigma\tau \in \Sigma$. Se infine $\sigma \in \Sigma$, anche $\sigma^{-1} \in \Sigma$. Dunque $\Sigma \leq \mathbf{S}_5$.

Le permutazioni che fissano l'insieme $\{1, 2\}$ sono $(1), (12)$. Quelle che fissano l'insieme $\{3, 4, 5\}$ sono

$$(1), (34), (35), (45), (345), (354).$$

Dunque le permutazioni di Σ sono le seguenti dodici permutazioni di \mathbf{S}_5 :

$$(1), (34), (35), (45), (345), (354), (12), (12)(34), (12)(35), (12)(45), (12)(345), (12)(354).$$

* * *

2.1.3. Sia K un campo ed n un intero ≥ 2 . In K^n si consideri il sottoinsieme

$$W = \{(c_1, c_2, \dots, c_n) \in K^n : c_1 = 0\}.$$

Verificare se W è un sottospazio vettoriale di K^n e se è un sottoanello di K^n .

Soluzione. $(W, +)$ è un gruppo abeliano. Verifichiamolo. Innanzitutto la n -pla nulla $\underline{0}$ è elemento di W . Se poi $(0, c_2, \dots, c_n), (0, d_2, \dots, d_n) \in W$, si ha:

$$(0, c_2, \dots, c_n) + (0, d_2, \dots, d_n) = (0, c_2 + d_2, \dots, c_n + d_n) \in W.$$

Infine $(0, c_2, \dots, c_n)$ ha per opposto $(0, -c_2, \dots, -c_n) \in W$.

La moltiplicazione per uno scalare manda gli elementi di W in W . Infatti, $\forall a \in K$:

$$a(0, c_2, \dots, c_n) = (0, ac_2, \dots, ac_n) \in W.$$

Gli assiomi di spazio vettoriale sono ovviamente verificati su W (perché lo sono in K^n) e dunque W è un sottospazio vettoriale di K^n .

Presi comunque $(0, c_2, \dots, c_n), (0, d_2, \dots, d_n) \in W$, si ha:

$$(0, c_2, \dots, c_n) \cdot (0, d_2, \dots, d_n) = (0, c_2d_2, \dots, c_nd_n) \in W.$$

Lae leggi distributive tra somma e prodotto valgono in W (perché valgono in K^n) e dunque W è un sottoanello di K^n .

N.B. Si osservi che W ammette elemento unità $(0, 1, \dots, 1)$, ma che tale unità non è l'unità di K^n .

* * *

2.2.1. Assegnate le due matrici (dipendenti da un parametro $a \in \mathbf{R}$)

$$A = \begin{pmatrix} 1 & 0 & 1 \\ a & 0 & 1 \end{pmatrix} \in \mathfrak{M}_{2,3}(\mathbf{R}), \quad B = \begin{pmatrix} 1 & a \\ a & 0 \\ 1 & 2 \end{pmatrix} \in \mathfrak{M}_{3,2}(\mathbf{R}),$$

- (i) Determinare gli eventuali $a \in \mathbf{R}$ per cui AB è una matrice triangolare superiore.
(ii) Determinare gli eventuali $a \in \mathbf{R}$ per cui BA è una matrice simmetrica.

Soluzione. Sviluppando il prodotto righe per colonne si ottiene:

$$AB = \begin{pmatrix} 2 & a+2 \\ a+1 & a^2+2 \end{pmatrix}, \quad BA = \begin{pmatrix} 1+a^2 & 0 & 1+a \\ a & 0 & a \\ 1+2a & 0 & 3 \end{pmatrix}.$$

- (i) La matrice AB è triangolare superiore $\iff a+1=0 \iff a=-1$.

$$(ii) \text{ La matrice } BA \text{ è simmetrica} \iff \begin{cases} a=0 \\ 1+2a=1+a \\ a=0 \end{cases} \iff \begin{cases} a=0 \\ a=2a \\ a=0 \end{cases} \iff a=0.$$

* * *

2.2.2 Siano $A, B \in \mathfrak{M}_n(K)$ due matrici simmetriche. Verificare che

la matrice AB è simmetrica $\iff A$ e B commutano.

Soluzione. Per definizione, se A, B sono matrici simmetriche di $\mathfrak{M}_n(K)$, allora ${}^t A = A$ e ${}^t B = B$. Poiché ${}^t(AB) = {}^t B {}^t A$, si ha:

$$AB \text{ è simmetrica} \iff {}^t(AB) = AB \iff {}^t B {}^t A = AB \iff BA = AB.$$

* * *

2.2.3. Assegnata la matrice $A = \begin{pmatrix} 0 & 1 \\ 2 & 2 \\ -1 & 0 \end{pmatrix} \in \mathfrak{M}_{3,2}(\mathbf{R})$, verificare che le due matrici ${}^t A A$ e $A {}^t A$

sono simmetriche. È vero, più in generale, che per ogni $A \in \mathfrak{M}_{m,n}(K)$, le due matrici ${}^t A A$ e $A {}^t A$ sono simmetriche? È vero che, per ogni $A \in \mathfrak{M}_n(K)$, risulta: ${}^t A A = A {}^t A$?

Soluzione. Risulta:

$${}^t A A = \begin{pmatrix} 5 & 4 \\ 4 & 5 \end{pmatrix}, \quad A {}^t A = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 8 & -2 \\ 0 & -2 & 1 \end{pmatrix}.$$

Come si vede, le due matrici sono simmetriche.

Per ogni $A \in \mathfrak{M}_{m,n}(K)$ la matrice ${}^t A A$ è simmetrica. Infatti ${}^t({}^t A A) = {}^t A {}^t({}^t A) = {}^t A A$. Analogamente si verifica che $A {}^t A$ è simmetrica. Infatti ${}^t(A {}^t A) = ({}^t A) {}^t A = A {}^t A$.

Se $A \in \mathfrak{M}_n(K)$, entrambe le matrici ${}^t A A$ e $A {}^t A$ sono in $\mathfrak{M}_n(K)$. Ma non è detto che coincidano. Infatti, considerata ad esempio in $\mathfrak{M}_2(\mathbf{R})$ la matrice $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, si ha:

$${}^t A A = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \quad \text{mentre} \quad A {}^t A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

* * *

2.2.4. Verificare che le matrici simmetriche di $\mathfrak{M}_n(K)$ formano un sottospazio vettoriale di $\mathfrak{M}_n(K)$. Formano un sottoanello di $\mathfrak{M}_n(K)$?

Soluzione. Denotiamo con $\mathfrak{S}_n(K)$ le matrici simmetriche di $\mathfrak{M}_n(K)$. Per definizione, se $A, B \in \mathfrak{S}_n(K)$, allora ${}^t A = A$ e ${}^t B = B$. Si ha:

$${}^t(A + B) = {}^t A + {}^t B = A + B$$

e dunque in $\mathfrak{S}_n(K)$ viene indotta la somma di matrici di $\mathfrak{M}_n(K)$. Inoltre ${}^t(-A) = -({}^t A) = -A$ e dunque l'opposta di una matrice simmetrica è simmetrica. Infine la matrice nulla $\mathbf{0}$ è simmetrica. Pertanto $(\mathfrak{S}_n(K), +)$ è un gruppo additivo. Se poi $c \in K$ e $A \in \mathfrak{S}_n(K)$, allora anche $c A \in \mathfrak{S}_n(K)$

[infatti ${}^t(cA) = c{}^tA = cA$]. Pertanto la moltiplicazione per uno scalare viene indotta su $\mathfrak{S}_n(K)$. Gli assiomi di spazio vettoriale valgono in $\mathfrak{S}_n(K)$ perché valgono in $\mathfrak{M}_n(K)$.

Invece $\mathfrak{S}_n(K)$ non è un sottoanello di $\mathfrak{M}_n(K)$. Infatti il prodotto di due matrici simmetriche non è in generale una matrice simmetrica. Ad esempio

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \text{ (matrice non simmetrica).}$$

* * *

2.2.5. Si consideri in $\mathfrak{M}_2(\mathbf{R})$ la matrice $A = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$. Verificare che $A \in \mathbf{GL}_2(\mathbf{R})$, cioè che esiste $B \in \mathfrak{M}_2(\mathbf{R})$ tale che $AB = I_2 = BA$.

Soluzione. Si ponga $B = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathfrak{M}_2(\mathbf{R})$, con a, b, c, d elementi incogniti (in \mathbf{R}). Bisogna determinare a, b, c, d in modo che risultino $AB = I_2 = BA$. Si ha:

$$AB = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a+2c & b+2d \\ c & d \end{pmatrix}.$$

Imponendo la condizione $AB = I_2$ si ottiene

$$a+2c=1, \quad b+2d=0, \quad c=0, \quad d=1 \quad \text{e dunque } B = \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix}.$$

Resta da verificare la condizione $BA = I_2$. Infatti

$$BA = \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I_2.$$

* * *

2.3.1. Sia p un intero ≥ 2 . Dimostrare il seguente risultato:

p è primo \iff se p divide un prodotto, divide almeno un fattore [cioè $p | ab \implies p | a$ o $p | b$].

Per dimostrare tale risultato si proceda come segue:

(1) Usando l'identità di Bézout provare il seguente *Lemma di Euclide*:

siano $a, b, c \in \mathbf{Z}$. Se $a | bc$ e $MCD(a, b) = 1$, allora $a | c$.

(2) Usando il lemma di Euclide, dimostrare (\implies): se $p | ab$ e $p \nmid a$, allora $p | b$.

(3) Dimostrare (\iff): p ha solo fattori banali [cioè: $p = xy \implies x = \pm 1$ o $y = \pm 1$].

Soluzione. (1) Per ipotesi $bc = at$, $\exists t \in \mathbf{Z}$. Dall'identità di Bézout, $1 = ar + bs$, $\exists r, s \in \mathbf{Z}$. Ne segue, moltiplicando tale uguaglianza per c :

$$c = arc + bsc = arc + ats = a(rc + ts) \quad \text{e dunque } a | c.$$

(2) Poiché $p \nmid a$, allora a, p non hanno fattori primi comuni e quindi $MCD(a, p) = 1$. Dal lemma di Euclide, $p | b$.

(3) Se $p = xy$, allora $p | xy$. Per ipotesi, si ha: $p | x$ o $p | y$.

Se $p | x$, allora $x = pt$, $\exists t \in \mathbf{Z}$. Dunque $p = pty$, da cui $ty = 1$ e quindi $y = \pm 1$.

Se $p \nmid x$, allora $p | y$. Quindi $y = ps$, $\exists s \in \mathbf{Z}$, da cui $p = xps$ e quindi $1 = xs$, cioè $x = \pm 1$.

* * *

2.3.2. Scrivere la tavola moltiplicativa di \mathbf{Z}_6^\times . Verificare se $\{\bar{0}, \bar{2}, \bar{4}\}$ è un sottoanello di \mathbf{Z}_6 .

Soluzione. La tavola moltiplicativa di \mathbf{Z}_6^\times è:

\cdot	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{4}$	$\bar{5}$
$\bar{1}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{4}$	$\bar{5}$
$\bar{2}$	$\bar{2}$	$\bar{4}$	$\bar{0}$	$\bar{2}$	$\bar{4}$
$\bar{3}$	$\bar{3}$	$\bar{0}$	$\bar{3}$	$\bar{0}$	$\bar{3}$
$\bar{4}$	$\bar{4}$	$\bar{2}$	$\bar{0}$	$\bar{4}$	$\bar{2}$
$\bar{5}$	$\bar{5}$	$\bar{4}$	$\bar{3}$	$\bar{2}$	$\bar{1}$

Si ponga $B = \{\bar{0}, \bar{2}, \bar{4}\}$. Queste tre classi resto sono le sole in \mathbf{Z}_6 rappresentate da interi pari. Ovviamente la somma di due classi resto rappresentate da interi pari è rappresentata da un pari. Inoltre, in \mathbf{Z}_6 si ha: $-\bar{2} = \bar{4}$ e $-\bar{4} = \bar{2}$. Ne segue facilmente che $(B, +)$ è un gruppo abeliano. Poiché anche il prodotto di due classi resto rappresentate da interi pari è rappresentato da un intero pari, si conclude che B è un sottoanello di \mathbf{Z}_6 (privo però di unità).

* * *

2.3.3. Scrivere la tavola additiva di \mathbf{Z}_5 e la tavola moltiplicativa \mathbf{Z}_5^* . Esistono sottoanelli propri di \mathbf{Z}_5 ?

Soluzione. Risulta:

+	\bar{0}	\bar{1}	\bar{2}	\bar{3}	\bar{4}
\bar{0}	\bar{0}	\bar{1}	\bar{2}	\bar{3}	\bar{4}
\bar{1}	\bar{1}	\bar{2}	\bar{3}	\bar{4}	\bar{0}
\bar{2}	\bar{2}	\bar{3}	\bar{4}	\bar{0}	\bar{1}
\bar{3}	\bar{3}	\bar{4}	\bar{0}	\bar{1}	\bar{2}
\bar{4}	\bar{4}	\bar{0}	\bar{1}	\bar{2}	\bar{3}

\cdot	\bar{1}	\bar{2}	\bar{3}	\bar{4}
\bar{1}	\bar{1}	\bar{2}	\bar{3}	\bar{4}
\bar{2}	\bar{2}	\bar{4}	\bar{1}	\bar{3}
\bar{3}	\bar{3}	\bar{1}	\bar{4}	\bar{2}
\bar{4}	\bar{4}	\bar{3}	\bar{2}	\bar{1}

Si indichi con B un eventuale sottoanello proprio di \mathbf{Z}_5 (cioè $B \neq \{\bar{0}\}$ e $B \neq \mathbf{Z}_5$). Ovviamente $B \ni \bar{0}$. Se B contenesse $\bar{1}$, allora $B = \mathbf{Z}_5$ [infatti ogni classe resto è ottenibile come somma di più copie di $\bar{1}$]. Poiché $B \neq \{\bar{0}\}$, B contiene un elemento tra $\bar{2}, \bar{3}, \bar{4}$. Ma

$$\bar{2} + \bar{2} + \bar{2} = \bar{1}, \bar{3} + \bar{3} = \bar{1}, \bar{4} + \bar{4} + \bar{4} = \bar{1}.$$

Dunque necessariamente $B \ni \bar{1}$ (e quindi non è un sottoanello proprio). Pertanto \mathbf{Z}_5 non contiene sottoanelli propri (e neppure sottogruppi propri rispetto a $+$).

* * *

2.3.4. Dimostrare il Piccolo Teorema di Fermat, procedendo come segue:

- (1) Facendo uso del lemma di Euclide verificare che, se p è primo ed a è coprimo con p , gli interi $a, 2a, 3a, \dots, (p-1)a$ sono a due a due non congruenti $\text{mod } p$.
- (2) Usando (1) ed il lemma di Euclide, dimostrare il Piccolo teorema di Fermat, cioè:

$$\text{se } p \text{ è primo ed } a \text{ è coprimo con } p, a^{p-1} \equiv 1 \pmod{p}.$$

Soluzione. (1) Per assurdo, sia $ta \equiv sa \pmod{p}$, con $1 \leq t < s \leq p-1$. Allora $(s-t)a \equiv 0 \pmod{p}$, cioè $p \mid (s-t)a$. Poiché $MCD(a, p) = 1$, dal lemma di Euclide segue che $p \mid s-t$. D'altra parte $0 < s-t < p$ e dunque $s-t \notin p\mathbf{Z}$: assurdo.

(2) Da (1) $a, 2a, 3a, \dots, (p-1)a$ sono a due a due non congruenti $\text{mod } p$. Quindi, a meno dell'ordine, sono complessivamente congruenti a $1, 2, \dots, p-1 \pmod{p}$. Pertanto

$$1 \cdot 2 \cdot 3 \cdot \dots \cdot (p-1) \equiv a \cdot 2a \cdot 3a \cdot \dots \cdot (p-1)a \pmod{p}.$$

Ne segue che

$$(p-1)! \equiv (p-1)! a^{p-1} \pmod{p}, \text{ ovvero } p \mid (p-1)! (1 - a^{p-1}).$$

Poiché p è primo e $(p-1)!$ è coprimo con p , $MCD(p, (p-1)!) = 1$. Dal Lemma di Euclide segue che $p \mid (1 - a^{p-1})$, cioè che $0 \equiv 1 - a^{p-1} \pmod{p}$, ovvero $a^{p-1} \equiv 1 \pmod{p}$.

* * *

2.3.5. Calcolare le ultime due cifre del numero naturale 82^{81} .

Soluzione. Si tratta di risolvere la congruenza $82^{81} \equiv X \pmod{100}$.

Si ha: $82^{81} = 2^{81} \cdot 41^{81}$. Se $2^{81} \equiv a \pmod{100}$ e $41^{81} \equiv b \pmod{100}$, allora $82^{81} \equiv ab \pmod{100}$. Risolviamo separatamente le due equazioni

$$2^{81} \equiv a \pmod{100} \text{ e } 41^{81} \equiv b \pmod{100}.$$

Per la seconda equazione si può ricorrere al teorema di Eulero-Fermat, essendo $MCD(41, 100) = 1$. Si ha: $41^{\varphi(100)} \equiv 1 \pmod{100}$, cioè $41^{40} \equiv 1 \pmod{100}$. Allora

$$41^{81} = (41^{40})^2 \cdot 41 \equiv 1^2 \cdot 41 = 41 \pmod{100}.$$

Per risolvere la prima congruenza $2^{81} \equiv a \pmod{100}$ non si può usare il teorema di Eulero-Fermat ed occorrono invece più calcoli. Risulta: $2^{81} = 2(2^{40})^2$. Calcoliamo $2^{40} \pmod{100}$. Risulta:

$$\begin{aligned} 2^{10} &= 2^5 \cdot 2^5 = 1024 \equiv 24 \pmod{100}; \\ 2^{40} &= (2^{10})^4 \equiv (24)^4 = (24^2)^2 \equiv (76)^2 \equiv (-24)^2 \equiv 76 \pmod{100}. \end{aligned}$$

Ne segue: $2^{81} = 2(2^{40})^2 \equiv 2(76)^2 \equiv 2 \cdot 76 \equiv 52 \pmod{100}$.

Si conclude che $82^{81} \equiv 41 \cdot 52 = 2132 \equiv 32 \pmod{100}$. Le ultime due cifre di 82^{81} sono 3, 2.

* * *

2.3.6. Risolvere l'equazione $\overline{39}X = \overline{12}$ in \mathbf{Z}_{603} .

Soluzione. Studiamo l'equazione congruenziale

$$39X \equiv 12 \pmod{603}.$$

Tale equazione è compatibile: infatti $MCD(39, 603) = 3$ e $3 \mid 12$. Inoltre ammette tre soluzioni comprese tra 0 e 602. Le corrispondenti classi resto ($\pmod{603}$) sono le soluzioni dell'equazione in \mathbf{Z}_{603} assegnata.

Per calcolare una soluzione dell'equazione $39X \equiv 12 \pmod{603}$, dividiamo per 3 tutti i suoi dati numerici, ottenendo l'equazione congruenziale ad essa equivalente

$$13X \equiv 4 \pmod{201}.$$

Per ottenerne una soluzione bisogna calcolare l'inverso di $\overline{13}$ in \mathbf{Z}_{201} . A tale scopo usiamo un'identità di Bézout relativa a 13, 201. Si ha:

$$201 = 13 \cdot 15 + 6, \quad 13 = 6 \cdot 2 + 1.$$

Dunque:

$$[1] = [13] - [6] \cdot 2 = [13] - ([201] - [13] \cdot 15) \cdot 2 = -2 \cdot [201] + 31 \cdot [13]$$

Un'identità di Bézout relativa a 13, 201 è quindi

$$1 = 31 \cdot 13 - 2 \cdot 201.$$

Passando in \mathbf{Z}_{201} si ottiene

$$\overline{1} = \overline{31} \cdot \overline{13} - \overline{2} \cdot \overline{201} = \overline{31} \cdot \overline{13} + \overline{0} = \overline{31} \cdot \overline{13}.$$

Segue che $13 \cdot 31 \equiv 1 \pmod{201}$ e pertanto $X \equiv 4 \cdot 31 = 124 \pmod{201}$. L'equazione congruenziale $39X \equiv 12 \pmod{603}$ quindi le tre soluzioni tra 0 e 602:

$$124 + \frac{603}{3}h, \quad \text{per } h = 0, 1, 2, \quad \text{cioè: } 124, 325, 526.$$

Si conclude che l'equazione assegnata $\overline{39}X = \overline{12}$ in \mathbf{Z}_{603} ha le tre soluzioni

$$\overline{124}, \overline{325}, \overline{526} \in \mathbf{Z}_{603}.$$

* * *

2.3.7. Risolvere l'equazione congruenziale lineare $14X \equiv 10 \pmod{120}$.

Soluzione. Si ha: $MCD(14, 120) = 2$ e $2 \mid 10$. Pertanto l'equazione è compatibile ed ammette due soluzioni comprese tra 0 e 119.

Dividiamo per 2 i dati numerici dell'equazione precedente ed otteniamo

$$7X \equiv 5 \pmod{60}.$$

Calcoliamo l'inverso di 7 mod 60, tramite l'identità di Bézout relativa a 60, 7. Applichiamo l'algoritmo euclideo delle divisioni successive a tali numeri:

$$60 = 7 \cdot 8 + 4, \quad 7 = 4 \cdot 1 + 3, \quad 4 = 3 \cdot 1 + 1.$$

Ne segue:

$$[4] = [60] - [7] \cdot 8, \quad [3] = [7] - [4] \cdot 1, \quad [1] = [4] - [3] \cdot 1,$$

da cui:

$$[1] = ([60] - [7] \cdot 8) - ([7] - [4] \cdot 1) = [60] - [7] \cdot 9 + ([60] - [7] \cdot 8) = [60] \cdot 2 + [7] \cdot (-17).$$

Allora un'identità di Bézout cercata è $1 = 2 \cdot 60 + (-17) \cdot 7$. Ne segue che

$$1 \equiv (-17) \cdot 7 \pmod{60}.$$

Poiché $-17 \equiv 43 \pmod{60}$, l'inverso di $7 \pmod{60}$ è 43 . L'equazione $7X \equiv 5 \pmod{60}$ si riscrive nella forma $X \equiv 5 \cdot 43 \pmod{60}$. Poiché $5 \cdot 43 = 215 \equiv 35 \pmod{60}$, tale equazione ha (unica) soluzione 35 .

Una soluzione dell'equazione congruenziale assegnata è quindi 35 e l'altra è data da $35 + \frac{120}{2} = 95$.

* * *

2.4.1. Sia $\mathfrak{L} = \{A \in \mathfrak{M}_3(\mathbf{R}) \mid a_{12} = a_{13} = a_{21} = a_{31} = 0\}$. Verificare se \mathfrak{L} è un \mathbf{R} -sottospazio vettoriale di $\mathfrak{M}_3(\mathbf{R})$ e se è un sottoanello di $\mathfrak{M}_3(\mathbf{R})$.

Soluzione. Le matrici di \mathfrak{L} sono della forma

$$A = \begin{pmatrix} a & 0 & 0 \\ 0 & b & c \\ 0 & d & e \end{pmatrix}, \quad \forall a, b, c, d, e \in \mathbf{R}.$$

Poiché la somma e la moltiplicazione per uno scalare sono definite "componente per componente", segue subito che, $\forall A, B \in \mathfrak{L}$ e $\forall t \in \mathbf{R}$, anche $A - B \in \mathfrak{L}$ e $tA \in \mathfrak{L}$. Pertanto \mathfrak{L} è un \mathbf{R} -sottospazio vettoriale di $\mathfrak{M}_3(\mathbf{R})$.

Se $A, A' \in \mathfrak{L}$, anche $AA' \in \mathfrak{L}$. Infatti

$$AA' = \begin{pmatrix} a & 0 & 0 \\ 0 & b & c \\ 0 & d & e \end{pmatrix} \begin{pmatrix} a' & 0 & 0 \\ 0 & b' & c' \\ 0 & d' & e' \end{pmatrix} = \begin{pmatrix} aa' & 0 & 0 \\ 0 & \dots & \dots \\ 0 & \dots & \dots \end{pmatrix} \in \mathfrak{L}.$$

Si conclude che \mathfrak{L} è un anello. Poiché $I_3 \in \mathfrak{L}$, \mathfrak{L} è anche unitario.

* * *

2.4.2. Sia $n \geq 2$ e sia $f : \mathfrak{M}_n(K) \rightarrow \mathfrak{M}_{n-1}(K)$ l'applicazione che ad ogni matrice $A \in \mathfrak{M}_n(K)$ associa la matrice ottenuta da A privandola degli elementi dell'ultima riga e dell'ultima colonna.

Verificare che f è un omomorfismo di K -spazi vettoriali e determinarne nucleo ed immagine. È vero che f è un omomorfismo di anelli?

Soluzione. Denotiamo con A' la matrice ottenuta da A eliminandone gli elementi dell'ultima riga e dell'ultima colonna. Poiché la somma e la moltiplicazione per uno scalare sono definite "componente per componente", segue subito che, $\forall A, B \in \mathfrak{M}_n(K)$ e $\forall c \in K$,

$$(A+B)' = A' + B', \quad (cA)' = cA'.$$

Dunque f è un omomorfismo di K -spazi vettoriali. Si ha:

$$A \in \text{Ker}(f) \iff A' = \mathbf{0} \iff a_{ij} = 0, \quad \forall i, j < n.$$

Inoltre $\text{Im}(f) = \mathfrak{M}_{n-1}(K)$. Infatti, $\forall B \in \mathfrak{M}_{n-1}(K)$, se a tale matrice aggiungiamo un'ulteriore riga e un'ulteriore colonna di elementi scelti arbitrariamente in K , otteniamo una matrice $A \in \mathfrak{M}_n(K)$ tale che $f(A) = B$.

f non è un omomorfismo di anelli. Scelta ad esempio la matrice $A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \in \mathfrak{M}_2(\mathbf{R})$, risulta:

$$f(A^2) = (A^2)' = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}' = (2), \quad \text{mentre} \quad f(A)f(A) = A'A' = (1)(1) = (1).$$

* * *

2.4.3. Sia τ una fissata permutazione di \mathbf{S}_n . Sia $f : \mathbf{S}_n \rightarrow \mathbf{S}_n$ l'applicazione così definita:

$$f(\sigma) = \tau \sigma \tau^{-1}, \quad \forall \sigma \in \mathbf{S}_n.$$

Verificare che f è un automorfismo di \mathbf{S}_n [è detto *automorfismo di coniugio* (relativo a τ)].

Soluzione. Si ha, $\forall \sigma_1, \sigma_2 \in \mathbf{S}_n$:

$$f(\sigma_1 \sigma_2) = \tau (\sigma_1 \sigma_2) \tau^{-1} = \tau \sigma_1 (\tau^{-1} \tau) \sigma_2 \tau^{-1} = (\tau \sigma_1 \tau^{-1})(\tau \sigma_2 \tau^{-1}) = f(\sigma_1) f(\sigma_2).$$

Dunque f è un omomorfismo. Per provare che è un isomorfismo basta considerare l'applicazione

$$g : \mathbf{S}_n \rightarrow \mathbf{S}_n \text{ tale che } g(\sigma) = \tau^{-1} \sigma \tau, \quad \forall \sigma \in \mathbf{S}_n.$$

Si ha: $(f \circ g)(\sigma) = f(\tau^{-1} \sigma \tau) = \tau (\tau^{-1} \sigma \tau) \tau^{-1} = \sigma$. Dunque $g \circ f = \mathbf{1}_{\mathbf{S}_n}$. Analogamente si verifica che $f \circ g = \mathbf{1}_{\mathbf{S}_n}$. Quindi f è un omomorfismo biiettivo di \mathbf{S}_n in sé (cioè un automorfismo di \mathbf{S}_n).

* * *

2.4.4. Sia $n \geq 2$ e sia $\pi : \mathbf{Z} \rightarrow \mathbf{Z}_n$ l'applicazione così definita:

$$\pi(a) = \bar{a}, \quad \forall a \in \mathbf{Z}.$$

Verificare che π è un omomorfismo di anelli e determinare $Ker(\pi)$. L'omomorfismo π è detto *proiezione canonica* di \mathbf{Z} sull'anello quoziante \mathbf{Z}_n .

Soluzione. Si ha, $\forall a, b \in \mathbf{Z}$:

$$\pi(a+b) = \overline{a+b} = \bar{a} + \bar{b} = \pi(a) + \pi(b), \quad \pi(ab) = \overline{ab} = \bar{a}\bar{b} = \pi(a)\pi(b).$$

Quindi π è un omomorfismo di anelli. Inoltre π è suriettivo. Infatti, $\forall \bar{a} \in \mathbf{Z}_n$, $\bar{a} = \pi(a)$.

Calcoliamo $Ker(\pi)$. Sia $a \in \mathbf{Z}$. Si ha:

$$a \in Ker(\pi) \iff \bar{a} = \bar{0} \iff a \equiv_n 0 \iff a \in n\mathbf{Z}.$$

Pertanto $Ker(\pi) = n\mathbf{Z} = \{0, \pm n, \pm 2n, \dots\}$.

* * *

2.4.5. L'applicazione $f : \mathbf{R} \rightarrow \mathfrak{M}_2(\mathbf{R})$ tale che

$$\varphi(t) = \begin{pmatrix} t & 0 \\ 0 & t \end{pmatrix}, \quad \forall t \in \mathbf{R},$$

è un omomorfismo di spazi vettoriali o di anelli? In caso affermativo, calcolarne nucleo ed immagine.

Soluzione. Si tratta di verificare se:

$$\varphi(t+s) = \varphi(t) + \varphi(s), \quad \forall t, s \in \mathbf{R};$$

$$\varphi(t \cdot s) = \varphi(t)\varphi(s), \quad \forall t, s \in \mathbf{R};$$

$$\varphi(ct) = c\varphi(t), \quad \forall t, c \in \mathbf{R}.$$

Si ha:

$$\varphi(t+s) = \begin{pmatrix} t+s & 0 \\ 0 & t+s \end{pmatrix} = \begin{pmatrix} t & 0 \\ 0 & t \end{pmatrix} + \begin{pmatrix} s & 0 \\ 0 & s \end{pmatrix} = \varphi(t) + \varphi(s);$$

$$\varphi(t \cdot s) = \begin{pmatrix} ts & 0 \\ 0 & ts \end{pmatrix} = \begin{pmatrix} t & 0 \\ 0 & t \end{pmatrix} \begin{pmatrix} s & 0 \\ 0 & s \end{pmatrix} = \varphi(t)\varphi(s);$$

$$\varphi(ct) = \begin{pmatrix} ct & 0 \\ 0 & ct \end{pmatrix} = c \begin{pmatrix} t & 0 \\ 0 & t \end{pmatrix} = c\varphi(t).$$

Dunque φ è un omomorfismo di anelli e di spazi vettoriali.

Risulta:

$$Ker \varphi = \{t \in \mathbf{R} : \varphi(t) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}\} = \{0\}.$$

Dunque φ è un monomorfismo. Inoltre

$$Im \varphi = \left\{ \begin{pmatrix} t & 0 \\ 0 & t \end{pmatrix}, \quad \forall t \in \mathbf{R} \right\}$$

è un sottoanello ed un \mathbf{R} -sottospazio vettoriale di $\mathfrak{M}_2(\mathbf{R})$. È formato da tutte e sole le matrici scalari di $\mathfrak{M}_2(\mathbf{R})$.

* * *

2.4.6. Sia $f : V \rightarrow W$ un omomorfismo di K -spazi vettoriali.

(i) Se W_1 è un sottospazio vettoriale di W , verificare che $f^{-1}(W_1)$ è un sottospazio vettoriale di V .

(ii) Se V_1 è un sottospazio vettoriale di V , verificare che $f(V_1)$ è un sottospazio vettoriale di W .

Soluzione. (i) Per definizione di controimmagine,

$$f^{-1}(W_1) = \{\underline{v} \in V \mid f(\underline{v}) \in W_1\}.$$

Siano $a, b \in K$ e $v_1, v_2 \in f^{-1}(W_1)$ [dunque $f(v_1), f(v_2) \in W_1$]. Bisogna verificare che

$$a\underline{v}_1 + b\underline{v}_2 \in f^{-1}(W_1).$$

Infatti $f(a\underline{v}_1 + b\underline{v}_2) = af(\underline{v}_1) + bf(\underline{v}_2) \in W_1$.

(ii) Siano $\underline{w}_1, \underline{w}_2 \in f(V_1)$, con $\underline{w}_1 = f(\underline{v}_1)$, $\underline{w}_2 = f(\underline{v}_2)$ e $\underline{v}_1, \underline{v}_2 \in V_1$. Siano $a, b \in K$. Bisogna verificare che

$$a \underline{w}_1 + b \underline{w}_2 \in f(V_1).$$

Poiché $a\underline{v}_1 + b\underline{v}_2 \in V_1$, allora:

$$a \underline{w}_1 + b \underline{w}_2 = a f(\underline{v}_1) + b f(\underline{v}_2) = f(a\underline{v}_1 + b\underline{v}_2) \in f(V_1).$$

N.B. Si noti che le stesse conclusioni (i) e (ii) si hanno anche se f è un omomorfismo tra gruppi o tra anelli: *l'immagine e la controimmagine di sottoanelli (o sottogruppi) sono sottoanelli (o sottogruppi)*.

* * *

2.4.7. Sia $\partial : K[X] \rightarrow K[X]$ l'applicazione di derivazione, così definita: $\forall P = \sum_{i=0}^n a_i X^i \in K[X]$,

$$\partial P = a_1 + 2a_2 X + 3a_3 X^2 + \dots + n a_n X^{n-1}$$

[se $K = \mathbf{R}$, ∂P è l'usuale derivata prima del polinomio P].

(i) Verificare che ∂ è un omomorfismo di K -spazi vettoriali ma non un omomorfismo di anelli.

(ii) Posto $K = \mathbf{R}$, calcolare $Im \partial$.

(iii) Posto $K = \mathbf{R}$, calcolare $Ker \partial$.

Soluzione. (i) Posto $c \in K$, $Q = \sum_{j=0}^m b_j X^j \in K[X]$ e P definito come nel testo, si ha:

$$\partial(P + Q) = (a_1 + b_1) + 2(a_2 + b_2)X + 3(a_3 + b_3)X^2 + \dots = \partial(P) + \partial(Q),$$

$$\partial(cP) = c a_1 + 2ca_2 X + 3ca_3 X^2 + \dots = c\partial(P).$$

Ne segue che ∂ è un omomorfismo di K -spazi vettoriali. Invece si ha ad esempio:

$$\partial(X^2) = 2X, \quad \partial(X)\partial(X) = 1 \cdot 1 = 1.$$

Ne segue che ∂ non è un omomorfismo di anelli.

(ii) Risulta subito che $Im \partial = \mathbf{R}[X]$. Infatti, $\forall P = \sum_{i=0}^n a_i X^i \in \mathbf{R}[X]$, si ha:

$$\partial(a_0 X + \frac{1}{2} a_1 X^2 + \frac{1}{3} a_2 X^3 + \dots + \frac{1}{n+1} a_n X^{n+1}) = P.$$

(iii) Si ha subito che $Ker \partial = \mathbf{R}$. Infatti ogni polinomio costante ha derivata nulla, mentre per ogni polinomio P di grado $d \geq 1$ $\partial(P)$ ha grado $d - 1 \geq 0$. Dunque $P \notin Ker \partial$.

N.B. Le considerazioni svolte in (ii) e (iii) non valgono se K è un campo \mathbf{Z}_p . Ad esempio in $\mathbf{Z}_2[X]$ $\partial(X^2) = \overline{2}X = \overline{0}$ e, sempre in $\mathbf{Z}_2[X]$, $X \notin Im(\partial)$.

* * *

2.4.8. In $\mathfrak{M}_2(\mathbf{R})$ si consideri il sottoinsieme

$$\mathcal{H} = \{A \in \mathfrak{M}_2(\mathbf{R}) \mid (A)_{11} = 0\}$$

[si ricorda che $(A)_{11}$ denota l'elemento a_{11} della matrice A]. Verificare che \mathcal{H} è un \mathbf{R} -sottospazio vettoriale di $\mathfrak{M}_2(\mathbf{R})$ e che è nucleo di un opportuno omomorfismo $\varphi : \mathfrak{M}_2(\mathbf{R}) \rightarrow \mathbf{R}$.

Soluzione. Si consideri l'applicazione $\varphi : \mathfrak{M}_2(\mathbf{R}) \rightarrow \mathbf{R}$ tale che $\varphi(A) = (A)_{11}$, $\forall A \in \mathfrak{M}_2(\mathbf{R})$.

Si ha, $\forall A, B \in \mathfrak{M}_2(\mathbf{R})$ e $\forall c \in \mathbf{R}$:

$$\varphi(A + B) = (A + B)_{11} = (A)_{11} + (B)_{11} = \varphi(A) + \varphi(B),$$

$$\varphi(cA) = (cA)_{11} = c(A)_{11} = c\varphi(A).$$

Pertanto φ è un omomorfismo di \mathbf{R} -spazi vettoriali. Si ha poi:

$$Ker \varphi = \{A \in \mathfrak{M}_2(\mathbf{R}) \mid \varphi(A) = 0\} = \mathcal{H}.$$

A questo punto non è necessario eseguire la verifica diretta del fatto che \mathcal{H} è un \mathbf{R} -sottospazio vettoriale di $\mathfrak{M}_2(\mathbf{R})$, in quanto il nucleo di un omomorfismo è sempre un sottospazio vettoriale.

* * *

2.5.1. Assegnati in \mathbf{R}^3 i due vettori $\underline{x} = (1, 0, -1)$ ed $\underline{y} = (0, 1, 1)$, verificare che sono linearmente indipendenti e determinare un vettore $\underline{z} \in \mathbf{R}^3$ tale che $\{\underline{x}, \underline{y}, \underline{z}\}$ sia una base di \mathbf{R}^3 .

Soluzione. Per provare che $\underline{x}, \underline{y}$ sono linearmente indipendenti bisogna dimostrare che

$$a\underline{x} + b\underline{y} = \underline{0} \implies a = b = 0$$

cioè che il sistema di equazioni lineari

$$\begin{cases} a = 0 \\ b = 0 \\ -a + b = 0 \end{cases}$$

non ammette soluzioni $\neq (0, 0)$. Ciò è evidente: l'unica soluzione è $(a, b) = (0, 0)$.

Un vettore \underline{z} tale che $\{\underline{x}, \underline{y}, \underline{z}\}$ sia una base di \mathbf{R}^3 è ogni vettore $\underline{z} \notin \langle \underline{x}, \underline{y} \rangle$. Per ottenerne uno possiamo ad esempio considerare i tre vettori $\underline{e}_1, \underline{e}_2, \underline{e}_3$ della base canonica di \mathbf{R}^3 e scegliere ad esempio il primo che non sia contenuto in $\langle \underline{x}, \underline{y} \rangle$.

Risulta: $\underline{e}_1 \in \langle \underline{x}, \underline{y} \rangle \iff \underline{e}_1 = a\underline{x} + b\underline{y}, \exists a, b \in \mathbf{R}$. Ciò si traduce nel sistema di equazioni lineari

$$\begin{cases} a = 1 \\ b = 0 \\ -a + b = 0, \end{cases}$$

che non ammette soluzioni. Dunque $\underline{e}_1 \notin \langle \underline{x}, \underline{y} \rangle$ e quindi $\{\underline{x}, \underline{y}, \underline{e}_1\}$ è una base di \mathbf{R}^3 .

* * *

2.5.2. Assegnati in \mathbf{R}^3 i vettori

$$\underline{x} = (1, 0, -1), \underline{y} = (0, 1, 1), \underline{z} = (1, 1, 0), \underline{w} = (-1, 0, 2),$$

verificare se è possibile estrarre da essi una base di \mathbf{R}^3 .

Soluzione. I primi due vettori $\underline{x}, \underline{y}$ sono linearmente indipendenti [cfr. esercizio precedente]. Verifichiamo se $\underline{z} \in \langle \underline{x}, \underline{y} \rangle$, cioè se esistono $a, b \in \mathbf{R}$ tali che $\underline{z} = a\underline{x} + b\underline{y}$. Ciò si traduce nel sistema di equazioni lineari

$$\begin{cases} a = 1 \\ b = 1 \\ -a + b = 0, \end{cases}$$

che ha soluzione $(a, b) = (1, 1)$. Dunque \underline{z} può essere eliminato dalla lista dei vettori assegnati. Verifichiamo ora se $\underline{w} \in \langle \underline{x}, \underline{y} \rangle$. Ciò si traduce nel sistema di equazioni lineari

$$\begin{cases} a = -1 \\ b = 0 \\ -a + b = 2, \end{cases}$$

che non ha soluzioni. Segue che $\underline{w} \notin \langle \underline{x}, \underline{y} \rangle$ e quindi $\{\underline{x}, \underline{y}, \underline{w}\}$ è una base di \mathbf{R}^3 estratta dai quattro vettori assegnati.

* * *

2.5.3. Sono assegnate in $\mathfrak{M}_2(\mathbf{R})$ le seguenti quattro matrici

$$A_1 = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, A_2 = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}, A_3 = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}, A_4 = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}.$$

(i) Verificare che tali matrici sono linearmente dipendenti ed esprimere l'ultima in funzione delle prime tre.

(ii) Posto $W = \langle A_1, A_2, A_3 \rangle$, determinare $\dim(W)$ e indicare una base di W .

(iii) Verificare se la matrice elementare $E^{11} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ appartiene a W e, in caso affermativo, scriverne le coordinate rispetto alla base di W ottenuta in (ii).

Soluzione. (i) Per rispondere a (i) basta determinare $a, b, c \in \mathbf{R}$ tali che

$$a A_1 + b A_2 + c A_3 = A_4.$$

Si ha:

$$a A_1 + b A_2 + c A_3 = \begin{pmatrix} a & a \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} b & 0 \\ b & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ c & c \end{pmatrix} = \begin{pmatrix} a+b & a \\ b+c & c \end{pmatrix}.$$

Quindi

$$a A_1 + b A_2 + c A_3 = A_4 \iff \begin{pmatrix} a+b & a \\ b+c & c \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} \iff \begin{cases} a+b=0 \\ a=1 \\ b+c=0 \\ c=1. \end{cases}$$

Tale sistema ammette, come facilmente si verifica, soluzione $(a, b, c) = (1, -1, 1)$. Pertanto

$$A_1 - A_2 + A_3 = A_4.$$

(ii) Per rispondere a (ii) basta verificare che i tre vettori A_1, A_2, A_3 sono linearmente indipendenti, perché in tal caso $\dim(W) = 3$ ed i tre vettori formano una base di W .

Si ponga

$$a A_1 + b A_2 + c A_3 = \mathbf{0} \quad [\text{matrice nulla di } \mathfrak{M}_2(\mathbf{R})].$$

Si ottiene

$$\begin{pmatrix} a+b & a \\ b+c & c \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad \text{cioè} \quad \begin{cases} a+b=0 \\ a=0 \\ b+c=0 \\ c=0. \end{cases}$$

Tale sistema lineare ammette come unica soluzione $(a, b, c) = (0, 0, 0)$. Pertanto A_1, A_2, A_3 sono linearmente indipendenti.

(iii) Poniamo $E^{11} = a A_1 + b A_2 + c A_3$ e verifichiamo se il corrispondente sistema lineare ammette soluzione. Si ottiene:

$$\begin{pmatrix} a+b & a \\ b+c & c \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \text{cioè} \quad \begin{cases} a+b=1 \\ a=0 \\ b+c=0 \\ c=0. \end{cases}$$

Segue subito che $b = 0$ e $b = 1$. Dunque il sistema non ha soluzioni e pertanto $E^{11} \notin W$.

* * *

2.5.4. Sia $\mathcal{L} = \{A \in \mathfrak{M}_3(\mathbf{R}) \mid a_{12} = a_{13} = a_{21} = a_{31} = 0\}$. Si tratta di un \mathbf{R} -sottospazio vettoriale di $\mathfrak{M}_3(\mathbf{R})$ (cfr. Eserc. 2.4.3). Determinare una base e la dimensione di \mathcal{L} .

Soluzione. La generica matrice A di \mathcal{L} è

$$A = \begin{pmatrix} a & 0 & 0 \\ 0 & b & c \\ 0 & d & e \end{pmatrix}.$$

Tale matrice dipende dai cinque parametri $a, b, c, d, e \in \mathbf{R}$. È evidente che le cinque matrici elementari

$$E^{1,1}, E^{2,2}, E^{2,3}, E^{3,2}, E^{3,3}$$

sono un sistema di generatori di \mathcal{L} . Infatti $A = a E^{1,1} + b E^{2,2} + c E^{2,3} + d E^{3,2} + e E^{3,3}$. Inoltre le cinque matrici sono linearmente indipendenti, in quanto fanno parte di una base (la base canonica di $\mathfrak{M}_3(\mathbf{R})$). Segue che formano una base di \mathcal{L} e dunque $\dim(\mathcal{L}) = 5$.

* * *

2.5.5. Sia $\{\underline{e}_1, \underline{e}_2\}$ una base di un \mathbf{R} -spazio vettoriale V di dimensione 2. Siano

$$\underline{f}_1 = \underline{e}_1 - 2 \underline{e}_2, \quad \underline{f}_2 = 2 \underline{e}_1 + \underline{e}_2 \in V.$$

Verificare che $\{\underline{f}_1, \underline{f}_2\}$ è una base di V . Assegnato in V il vettore $\underline{v} = \underline{e}_1 - \underline{e}_2$, esprimere rispetto alla base $\{\underline{f}_1, \underline{f}_2\}$.

Soluzione. Poiché $\dim(V) = 2$, per provare che $\{\underline{f}_1, \underline{f}_2\}$ è una base di V basta verificare che i vettori $\underline{f}_1, \underline{f}_2$ sono linearmente indipendenti. Sia $a\underline{f}_1 + b\underline{f}_2 = \underline{0}$. Allora

$$a(\underline{e}_1 - 2\underline{e}_2) + b(2\underline{e}_1 + \underline{e}_2) = (a+2b)\underline{e}_1 - (2a-b)\underline{e}_2 = \underline{0}.$$

Ne segue che

$$\begin{cases} a+2b=0 \\ 2a-b=0. \end{cases}$$

Tale sistema di equazioni lineari ha soltanto la soluzione $(a,b) = (0,0)$ [infatti, dalla seconda equazione, $b = 2a$; sostituendo nella prima, si ottiene $5a = 0$. Allora $a = 0$ e quindi $b = 0$]. Pertanto $\underline{f}_1, \underline{f}_2$ sono linearmente indipendenti.

Possiamo ora esprimere i vettori $\underline{e}_1, \underline{e}_2$ in funzione di $\{\underline{f}_1, \underline{f}_2\}$. Si ha:

$$2\underline{f}_1 - \underline{f}_2 = 2\underline{e}_1 - 4\underline{e}_2 - 2\underline{e}_1 - \underline{e}_2 = -5\underline{e}_2.$$

Pertanto $\underline{e}_2 = -\frac{2}{5}\underline{f}_1 + \frac{1}{5}\underline{f}_2$. Ne segue:

$$\underline{e}_1 = \underline{f}_1 + 2\underline{e}_2 = \underline{f}_1 - \frac{4}{5}\underline{f}_1 + \frac{2}{5}\underline{f}_2 = \frac{1}{5}\underline{f}_1 + \frac{2}{5}\underline{f}_2.$$

Abbiamo ottenuto

$$\underline{e}_1 = \frac{1}{5}(\underline{f}_1 + 2\underline{f}_2), \quad \underline{e}_2 = \frac{1}{5}(-2\underline{f}_1 + \underline{f}_2).$$

Si ha quindi

$$\underline{v} = \underline{e}_1 - \underline{e}_2 = \frac{1}{5}(\underline{f}_1 + 2\underline{f}_2) - \frac{1}{5}(-2\underline{f}_1 + \underline{f}_2) = \frac{3}{5}\underline{f}_1 + \frac{1}{5}\underline{f}_2.$$

Possiamo dire che \underline{v} ha coordinate $(1, -1)$ in base $\{\underline{e}_1, \underline{e}_2\}$ e coordinate $(\frac{3}{5}, \frac{1}{5})$ in base $\{\underline{f}_1, \underline{f}_2\}$.

* * *

2.5.6. Assegnata una base $\{\underline{e}_1, \underline{e}_2, \underline{e}_3, \underline{e}_4\}$ di un \mathbf{R} -spazio vettoriale V avente dimensione 4, si considerino i due sottospazi vettoriali

$$V_1 = \langle \underline{e}_1 - \underline{e}_2, \underline{e}_1 - \underline{e}_3 \rangle, \quad V_2 = \langle \underline{e}_1 - \underline{e}_4, \underline{e}_3 - \underline{e}_4 \rangle.$$

Verificare se i due sottospazi vettoriali sono supplementari.

Soluzione. Se due sottospazi vettoriali V_1, V_2 sono supplementari, deve risultare $V_1 \cap V_2 = \{\underline{0}\}$. I vettori di $V_1 \cap V_2$ sono tutti e soli quelli per cui $\exists a, b, c, d \in \mathbf{R}$ tali che

$$a(\underline{e}_1 - \underline{e}_2) + b(\underline{e}_1 - \underline{e}_3) = c(\underline{e}_1 - \underline{e}_4) + d(\underline{e}_3 - \underline{e}_4),$$

cioè

$$(a+b-c)\underline{e}_1 - a\underline{e}_2 - (b+d)\underline{e}_3 + (c+d)\underline{e}_4 = \underline{0}.$$

Tale uguaglianza vettoriale si traduce nel sistema di equazioni

$$\begin{cases} a+b-c=0 \\ -a=0 \\ -(b+d)=0 \\ c+d=0, \end{cases} \quad \text{che si riscrive nella forma} \quad \begin{cases} a=0 \\ c=b \\ d=-b \\ c+d=0. \end{cases}$$

Tale sistema ammette soluzioni $(0, b, b, -b)$, $\forall b \in \mathbf{R}$, e, in particolare, la soluzione $(0, 1, 1, -1)$. Segue che

$$\underline{e}_1 - \underline{e}_3 = (\underline{e}_1 - \underline{e}_4) - (\underline{e}_3 - \underline{e}_4) \in V_1 \cap V_2$$

[ciò che del resto poteva essere notato sin da principio]. Concludiamo quindi che i due sottospazi non sono supplementari.

N.B.1. Si può anche procedere in altro modo.

Per prima cosa possiamo verificare che $\dim(V_1) = \dim(V_2) = 2$. Poi, posto

$$i := \dim(V_1 \cap V_2), \quad u := \dim(V_1 + V_2),$$

essendo (in base alla formula di Grassmann) $u+i=4$, basta verificare che $u < 4$ [da cui $i > 0$].

Per provare che $u < 4$ basta verificare che sono linearmente indipendenti i quattro vettori

$$\underline{e}_1 - \underline{e}_2, \underline{e}_1 - \underline{e}_3, \underline{e}_1 - \underline{e}_4, \underline{e}_3 - \underline{e}_4.$$

N.B.2. Si può verificare che $\dim(V_1 + V_2) = 3$. A tale scopo basta verificare che i primi tre vettori della lista precedente sono linearmente indipendenti]. Ne segue che $\dim(V_1 \cap V_2) = 1$. Una base di $V_1 \cap V_2$ è data dal vettore $\underline{e}_1 - \underline{e}_3$.

* * *

2.5.7. Una matrice $A \in \mathfrak{M}_n(K)$ è detta *antisimmetrica* se risulta $a_{ij} = -a_{ji}$, $\forall i, j = 1, \dots, n$ [cioè se $A = -A^t$]. Verificare che le matrici antisimmetriche di $\mathfrak{M}_3(\mathbf{R})$ formano un sottospazio vettoriale di dimensione 3 in $\mathfrak{M}_3(\mathbf{R})$ e indicarne una base.

Soluzione. Osserviamo che, se A è una matrice antisimmetrica, $a_{ii} = -a_{ii}$ e dunque $2a_{ii} = 0$. Se quindi $K \neq \mathbf{Z}_2$, allora $a_{ii} = 0$, $\forall i = 1, \dots, n$.

Denotiamo con \mathfrak{S} le matrici antisimmetriche di $\mathfrak{M}_3(\mathbf{R})$. La generica matrice di \mathfrak{S} è della forma

$$A = \begin{pmatrix} 0 & a & b \\ -a & 0 & c \\ -b & -c & 0 \end{pmatrix}, \quad \forall a, b, c \in \mathbf{R}.$$

Si verifica subito che, se $A, B \in \mathfrak{S}$ e $c, d \in \mathbf{R}$ anche $cA + dB \in \mathfrak{S}$ [infatti $-^t(cA + dB) = c(-^tA) + d(-^tB) = cA + dB$]. Dunque \mathfrak{S} è un sottospazio vettoriale di $\mathfrak{M}_3(\mathbf{R})$.

Risulta:

$$\begin{pmatrix} 0 & a & b \\ -a & 0 & c \\ -b & -c & 0 \end{pmatrix} = a \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + b \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} + c \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}.$$

Dunque le tre matrici antisimmetriche

$$H_1 = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad H_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \quad H_3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}$$

formano un sistema di generatori di \mathfrak{S} .

Verifichiamo che tali matrici sono linearmente indipendenti (e quindi formano una base di \mathfrak{S}). Sia

$$aH_1 + bH_2 + cH_3 = \mathbf{0}.$$

Poiché $H_1 = E^{12} - E^{21}$, $H_2 = E^{13} - E^{31}$, $H_3 = E^{23} - E^{32}$, la precedente relazione si riscrive nella forma

$$aE^{12} - aE^{21} + bE^{13} - bE^{31} + cE^{23} - cE^{32} = \mathbf{0}.$$

Essendo i nove vettori E^{ij} linearmente indipendenti, si conclude che lo sono anche i sei vettori sopra considerati. Si conclude che $a = b = c = 0$.

* * *

2.5.8. Scelto $n \in \mathbf{N}$, si consideri in $K[X]$ il sottoinsieme

$$W = W_n = \{P \in K[X] \mid \deg(P) \leq n\}.$$

Verificare che W è un K -sottospazio vettoriale di $K[X]$. Indicarne la dimensione ed una base.

Soluzione. Se $P, Q \in W$ e $a, b \in K$, si ha:

$$\deg(aP + bQ) \leq \max(\deg(P), \deg(Q)) \leq n.$$

Dunque $aP + bQ \in W$ e pertanto W è un K -sottospazio vettoriale di $K[X]$.

Consideriamo ora gli $n+1$ monomi $1, X, X^2, \dots, X^n$. Preso comunque $P = \sum_{i=0}^n a_i X^i \in W$, risulta:

$$P = a_0 \cdot 1 + a_1 \cdot X + a_2 \cdot X^2 + \dots + a_n \cdot X^n.$$

Pertanto $\{1, X, X^2, \dots, X^n\}$ è un sistema di generatori di W . Tali monomi sono linearmente indipendenti. Infatti

$$\sum_{i=0}^n a_i X^i = 0 \implies a_0 = a_1 = \dots = a_n = 0$$

[in quanto un polinomio è nullo \iff tutti i suoi coefficienti sono nulli].

Abbiamo così provato che $\dim(W_n) = n + 1$.

N.B. Rispetto alla base "monomiale" $\{1, X, X^2, \dots, X^n\}$ le coordinate di un polinomio coincidono con i suoi coefficienti. Per questo motivo la base monomiale è una base speciale di W_n e viene detta *base canonica* di W_n .

* * *

ESERCIZI PROPOSTI

Capitolo 3

3.1.1. Risolvere con l'algoritmo di Gauss il seguente $SL(3, 5, \mathbf{R})$

$$\begin{cases} x_3 + 2x_5 = 2 \\ x_4 - x_5 = 3 \\ 2x_1 + x_3 = 1. \end{cases}$$

Dedurne una base del sottospazio vettoriale Σ_0 delle soluzioni del SLO associato.

Soluzione. Il sistema ha matrice completa

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

Con $\mathbf{I}[(1^a) \leftrightarrow (3^a)]$ si ottiene la matrice

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

Osservato che la seconda colonna è nulla, operiamo lo scambio di variabili $x_2 \leftrightarrow x_4$. Ciò significa introdurre un nuovo set di variabili y_1, \dots, y_5 tali che

$$y_1 = x_1, \quad y_2 = x_4, \quad y_3 = x_3, \quad y_4 = x_2, \quad y_5 = x_5.$$

La matrice diventa

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

Operiamo ora con $\mathbf{II}[(1^a) \rightarrow \frac{1}{2}(1^a)]$. Si ottiene la matrice

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

Otteniamo quindi il sistema a scala

$$\begin{cases} y_1 + \frac{1}{2}y_3 = \frac{1}{2} \\ y_2 - y_5 = 3 \\ y_3 + 2y_5 = 2. \end{cases}$$

Per risolverlo poniamo $y_4 = t$, $y_5 = s$. Otteniamo successivamente:

$$y_3 = 2 - 2s, \quad y_2 = 3 + s, \quad y_1 = s - \frac{1}{2}.$$

Ne segue:

$$\begin{cases} x_1 = y_1 = s - \frac{1}{2} \\ x_2 = y_4 = t \\ x_3 = y_3 = 2 - 2s \\ x_4 = y_2 = 3 + s \\ x_5 = y_5 = s. \end{cases}$$

Il SL assegnato ha ∞^2 soluzioni, descritte dall'insieme

$$\Sigma = \{(s - \frac{1}{2}, t, 2 - 2s, 3 + s, s), \quad \forall t, s \in \mathbf{R}\}.$$

Una soluzione particolare del sistema è $\underline{x} = (-\frac{1}{2}, 0, 2, 3, 0)$ (ottenuta ponendo $t = s = 0$). Allora

$$\Sigma_0 = -\underline{x} + \Sigma = \{(s, t, -2s, s, s), \quad \forall t, s \in \mathbf{R}\}.$$

Poiché $(s, t, -2s, s, s) = s(1, 0, -2, 1, 1) + t(0, 1, 0, 0, 0)$, si conclude che Σ_0 ha come base i due vettori $(1, 0, -2, 1, 1)$ e $(0, 1, 0, 0, 0)$, cioè $\underline{e}_1 - 2\underline{e}_3 + \underline{e}_4 + \underline{e}_5$ ed \underline{e}_2 .

3.1.2. Al variare del parametro $a \in \mathbf{R}$, risolvere il seguente $SLO(3, 2, \mathbf{R})$

$$\begin{cases} x + ay = 0 \\ 2x + 2y = 0 \\ ax = 0. \end{cases}$$

Soluzione. Il SLO assegnato ha matrice dei coefficienti

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

[Rimarchiamo che per i SLO è inutile considerare la matrice completa del sistema, in quanto non sussiste problema di compatibilità e d'altra parte la colonna dei termini noti, essendo nulla, resta sempre inalterata con operazioni elementari di riga].

Con $\text{III}[(2^a) \rightarrow (2^a) - 2(1^a)]$ e $\text{III}[(3^a) \rightarrow (3^a) - a(1^a)]$, si ottiene la matrice

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

Si hanno due eventualità: (i) $2 - 2a \neq 0$, (ii) $2 - 2a = 0$.

Nel caso (i), $a \neq 1$. Con $\text{II}[(2^a) \rightarrow \frac{1}{2-2a}(2^a)]$ si ottiene la matrice $\begin{pmatrix} 1 & a \\ 0 & 1 \\ 0 & -a^2 \end{pmatrix}$.

Con $\text{III}[(3^a) \rightarrow (3^a) + a^2(2^a)]$ si ottiene la matrice $\begin{pmatrix} 1 & a \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$. Si elimina la terza riga e si ottiene

la matrice $\begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}$, che corrisponde al $SL0(2, 2, \mathbf{R})$ a scala $\begin{cases} x + ay = 0 \\ y = 0. \end{cases}$ Tale SLO ammette l'unica soluzione $(x, y) = (0, 0)$.

Nel caso (ii), $a = 1$ e la matrice del sistema è $A = \begin{pmatrix} 1 & 1 \\ 0 & 0 \\ 0 & -1 \end{pmatrix}$. Si elimina la seconda riga e si ottiene la matrice $A = \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix}$. Con $\text{II}[(2^a) \rightarrow -(2^a)]$ si ottiene la matrice $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ che corrisponde al $SL0(2, 2, \mathbf{R})$ a scala $\begin{cases} x = 0 \\ y = 0. \end{cases}$ Tale SLO ammette l'unica soluzione $(x, y) = (0, 0)$.

Dunque in ogni caso il SLO è privo di autosoluzioni.

* * *

3.1.3. Al variare del parametro $a \in \mathbf{R}$, risolvere il seguente $SLO(2, 2, \mathbf{R})$

$$\begin{cases} 2x + (a+2)y = 0 \\ (a+1)x + (a^2+2)y = 0. \end{cases}$$

Soluzione. Il SLO ha matrice dei coefficienti

$$A = \begin{pmatrix} 2 & a+2 \\ a+1 & a^2+2 \end{pmatrix}.$$

Con $\text{II}[(1^a) \rightarrow \frac{1}{2}(1^a)]$ si ottiene la matrice

$$\begin{pmatrix} 1 & \frac{a+2}{2} \\ a+1 & a^2+2 \end{pmatrix}.$$

Con $\text{III}[(2^a) \rightarrow (2^a) - (a+1)(1^a)]$ si ottiene la matrice

$$\begin{pmatrix} 1 & \frac{a+2}{2} \\ 0 & a^2+2-(a+1)\frac{a+2}{2} \end{pmatrix} = \begin{pmatrix} 1 & \frac{a+2}{2} \\ 0 & \frac{a^2}{2}-\frac{3}{2}a+1 \end{pmatrix} = \begin{pmatrix} 1 & \frac{a+2}{2} \\ 0 & \frac{1}{2}(a-1)(a-2) \end{pmatrix}.$$

Risulta: se $a \neq 1, 2$, il SLO ha soltanto la soluzione banale $(x, y) = (0, 0)$.

Se invece $a = 1$ o $a = 2$, la seconda riga è nulla e può quindi essere eliminata. Il SLO si riduce a

$$\{ x + (\frac{a+2}{2})y = 0 \quad [\text{con } a = 1, 2]. \}$$

Se $a = 1$ il sistema ha le ∞^1 soluzioni $(x, y) = (-\frac{3}{2}t, t)$, $\forall t \in \mathbf{R}$. Se invece $a = 2$ il sistema ha le ∞^1 soluzioni $(x, y) = (-2t, t)$, $\forall t \in \mathbf{R}$.

* * *

3.1.4. Al variare dei parametri non nulli $a, b \in \mathbf{R}$, risolvere il seguente $SLO(3, 3, \mathbf{R})$

$$\begin{cases} a y + b z = 0 \\ -a x + z = 0 \\ -b x - y = 0. \end{cases}$$

Soluzione. Il SLO assegnato ha matrice dei coefficienti

$$A = \begin{pmatrix} 0 & a & b \\ -a & 0 & 1 \\ -b & -1 & 0 \end{pmatrix}.$$

Con **I** $[(1^a) \leftrightarrow (2^a)]$ si ottiene la matrice $\begin{pmatrix} -a & 0 & 1 \\ 0 & a & b \\ -b & -1 & 0 \end{pmatrix}$.

Con **II** $[(1^a) \rightarrow -\frac{1}{a}(1^a)]$ si ottiene la matrice $\begin{pmatrix} 1 & 0 & -\frac{1}{a} \\ 0 & a & b \\ -b & -1 & 0 \end{pmatrix}$.

Con **III** $[(3^a) \rightarrow (3^a) + b(1^a)]$ si ottiene la matrice $\begin{pmatrix} 1 & 0 & -\frac{1}{a} \\ 0 & a & b \\ 0 & -1 & -\frac{b}{a} \end{pmatrix}$.

Con **II** $[(2^a) \rightarrow \frac{1}{a}(2^a)]$ si ottiene la matrice $\begin{pmatrix} 1 & 0 & -\frac{1}{a} \\ 0 & 1 & \frac{b}{a} \\ 0 & -1 & -\frac{b}{a} \end{pmatrix}$.

Con **III** $[(3^a) \rightarrow (3^a) + (2^a)]$ si ottiene la matrice $\begin{pmatrix} 1 & 0 & -\frac{1}{a} \\ 0 & 1 & \frac{b}{a} \\ 0 & 0 & 0 \end{pmatrix}$.

Eliminiamo la terza riga (nulla) ed otteniamo la matrice del seguente sistema lineare a scala:

$$\begin{cases} x - \frac{1}{a}z = 0 \\ y + \frac{b}{a}z = 0. \end{cases}$$

Poniamo $z = t$ ed otteniamo successivamente $y = -t\frac{b}{a}$, $x = t\frac{1}{a}$. Pertanto l'insieme delle soluzioni del SLO assegnato è $\{(\frac{t}{a}, -\frac{bt}{a}, t), \forall t \in \mathbf{R}\}$. Si tratta del sottospazio vettoriale di \mathbf{R}^3 generato dal vettore $(\frac{1}{a}, -\frac{b}{a}, 1)$ [ovvero dal vettore $(1, -b, a)$].

* * *

3.1.5 Risolvere con l'algoritmo di Gauss il seguente $SL(2, 2, \mathbf{Z}_5)$

$$\begin{cases} \overline{3}x + y = \overline{1} \\ x - y = \overline{0}. \end{cases}$$

Soluzione. Il sistema ha matrice completa [si osservi che $\overline{-1} = \overline{4}$]:

$$M = \begin{pmatrix} \overline{3} & \overline{1} & \overline{1} \\ \overline{1} & \overline{4} & \overline{0} \end{pmatrix}.$$

Con **II** $[(1^a) \rightarrow \frac{1}{3}(1^a)]$ si ottiene la matrice [si osservi che $\frac{1}{3} = \overline{2}$]:

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

Con **III** $[(2^a) \rightarrow (2^a) - (1^a)]$ si ottiene la matrice

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

Con **II** $[(2^a) \rightarrow \frac{1}{2}(2^a)]$ si ottiene la matrice [si osservi che $\frac{1}{2} = \overline{3}$]:

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

Si è ottenuto il SL a scala $\begin{cases} x + \bar{2}y = \bar{2} \\ y = \bar{4}, \end{cases}$ il quale ammette l'unica soluzione $(x, y) = (\bar{4}, \bar{4})$.

* * *

3.2.1. Eseguire il calcolo del determinante delle due seguenti matrici, sfruttando opportunamente le proprietà dei determinanti (per semplificare il calcolo)

$$A = \begin{pmatrix} 999 & 1000 & 1001 \\ 1 & 2 & 3 \\ 0 & -1 & 2 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 50 & 0 \\ 2 & 100 & -1 \\ 3 & 75 & 2 \end{pmatrix}.$$

Soluzione. Si osserva subito che

$$A^{(1)} = 1000(1 \ 1 \ 1) + (-1 \ 0 \ 1).$$

Conviene allora calcolare $\det(A)$ sostituendo la prima riga di A con la combinazione lineare di matrici-riga scritta sopra. Si ha:

$$\det(A) = 1000 \begin{vmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 0 & -1 & 2 \end{vmatrix} + \begin{vmatrix} -1 & 0 & 1 \\ 1 & 2 & 3 \\ 0 & -1 & 2 \end{vmatrix} = 1000 \cdot 4 + (-8) = 3992.$$

Si osserva subito che $B_{(2)} = 25 \begin{pmatrix} 2 \\ 4 \\ 3 \end{pmatrix}$. Pertanto si ha:

$$\det(B) = 25 \begin{vmatrix} 1 & 2 & 0 \\ 2 & 4 & -1 \\ 3 & 3 & 2 \end{vmatrix} = 25 \cdot (-3) = -75.$$

* * *

3.2.2 Per ogni $\alpha \in \mathbf{R}$, si considerino le due matrici

$$A = \begin{pmatrix} \cos \alpha & \sin \alpha \\ \sin \alpha & -\cos \alpha \end{pmatrix}, \quad B = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \in \mathfrak{M}_2(\mathbf{R}).$$

Verificare che sono entrambe invertibili e che $A^{-1} = {}^t A$, $B^{-1} = {}^t B$.

Soluzione. Risulta:

$$\det(A) = -\cos^2 \alpha - \sin^2 \alpha = -1, \quad \det(B) = \cos^2 \alpha + \sin^2 \alpha = 1.$$

Essendo $\det(A), \det(B) \neq 0$, A e B sono invertibili. Calcoliamone le matrici inverse. Si ha:

$$A^{-1} = \frac{1}{\det(A)} \begin{pmatrix} -\cos \alpha & -\sin \alpha \\ -\sin \alpha & \cos \alpha \end{pmatrix} = - \begin{pmatrix} -\cos \alpha & -\sin \alpha \\ -\sin \alpha & \cos \alpha \end{pmatrix} = \begin{pmatrix} \cos \alpha & \sin \alpha \\ \sin \alpha & -\cos \alpha \end{pmatrix} = A = {}^t A$$

[l'ultima uguaglianza perché A è simmetrica].

$$B^{-1} = \frac{1}{\det(B)} \begin{pmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{pmatrix} = \begin{pmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{pmatrix} = {}^t B.$$

N.B. Si potrebbe verificare che le matrici A, B sopra considerate sono le uniche matrici invertibili di ordine 2, la cui inversa coincida con la trasposta.

* * *

3.2.3 Assegnata la matrice quadrata

$$A = \begin{pmatrix} 1 & 2 & -1 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix} \in \mathfrak{M}_3(\mathbf{R}),$$

(i) Calcolare la matrice dei complementi algebrici \mathcal{C}_A e verificare che $\det(\mathcal{C}_A) = (\det(A))^2$.

(ii) Dimostrare che, $\forall A \in \mathbf{GL}_n(K)$, risulta: $\det(\mathcal{C}_A) = (\det(A))^{n-1}$.

Soluzione. (i) Risulta:

$$\mathcal{C}_A = \begin{pmatrix} 1 & 1 & -1 \\ -2 & 2 & 2 \\ 3 & -1 & 1 \end{pmatrix}.$$

Si verifica subito che $\det(A) = 4$, $\det(\mathcal{C}_A) = 16 = 4^2$.

(ii) È noto (cfr. **Prop. 2.2** di questo capitolo) che ${}^t\mathcal{C}_A A = \det(A) I_n$. È altresì noto che $\det(cA) = c^n \det(A)$, $\forall A \in \mathfrak{M}_n(K)$, $\forall c \in K$

Utilizzando le proprietà dei determinanti (tra cui il teorema di Binet) si ha:

$$(\det(A))^n = \det(\det(A) I_n) = \det({}^t\mathcal{C}_A A) = \det({}^t\mathcal{C}_A) \det(A) = \det(\mathcal{C}_A) \det(A),$$

cioè

$$(\det(A))^n = \det(\mathcal{C}_A) \det(A).$$

Cancellando $\det(A)$ [che è $\neq 0$], si ottiene $\det(\mathcal{C}_A) = (\det(A))^{n-1}$, come richiesto.

* * *

3.2.4. Consideriamo le matrici di $\mathfrak{M}_2(\mathbf{Z}_2)$. Quante sono? Quante e quali di esse sono quelle invertibili? Il gruppo che esse formano è commutativo?

Soluzione. Le matrici $A \in \mathfrak{M}_2(\mathbf{Z}_2)$ sono $2^4 = 16$. Infatti ognuno de quattro elementi di ogni matrice A può essere scelto in due modi diversi [$\bar{0}$ oppure $\bar{1}$] indipendentemente dagli altri elementi.

Le matrici invertibili sono quelle formate esattamente da tre elementi $= \bar{1}$ oppure da due soli elementi $= \bar{1}$, ma disposti non sulla stessa riga o colonna. Sono quindi le seguenti $4 + 2$ matrici

$$\left(\begin{array}{cc} \bar{0} & \bar{1} \\ \bar{1} & \bar{1} \end{array} \right), \left(\begin{array}{cc} \bar{1} & \bar{0} \\ \bar{1} & \bar{1} \end{array} \right), \left(\begin{array}{cc} \bar{1} & \bar{1} \\ \bar{0} & \bar{1} \end{array} \right), \left(\begin{array}{cc} \bar{1} & \bar{1} \\ \bar{1} & \bar{0} \end{array} \right), \left(\begin{array}{cc} \bar{0} & \bar{1} \\ \bar{1} & \bar{0} \end{array} \right), \left(\begin{array}{cc} \bar{1} & \bar{0} \\ \bar{0} & \bar{1} \end{array} \right).$$

Tali matrici formano un gruppo rispetto al prodotto, denotato $\mathbf{GL}_2(\mathbf{Z}_2)$. Per verificare se tale gruppo è commutativo non c'è da fare che la verifica diretta.

Consideriamo ad esempio le due matrici $\left(\begin{array}{cc} \bar{0} & \bar{1} \\ \bar{1} & \bar{0} \end{array} \right)$, $\left(\begin{array}{cc} \bar{0} & \bar{1} \\ \bar{1} & \bar{1} \end{array} \right)$. Si ha:

$$\left(\begin{array}{cc} \bar{0} & \bar{1} \\ \bar{1} & \bar{0} \end{array} \right) \left(\begin{array}{cc} \bar{0} & \bar{1} \\ \bar{1} & \bar{1} \end{array} \right) = \left(\begin{array}{cc} \bar{1} & \bar{1} \\ \bar{0} & \bar{1} \end{array} \right), \quad \left(\begin{array}{cc} \bar{0} & \bar{1} \\ \bar{1} & \bar{1} \end{array} \right) \left(\begin{array}{cc} \bar{0} & \bar{1} \\ \bar{1} & \bar{0} \end{array} \right) = \left(\begin{array}{cc} \bar{1} & \bar{0} \\ \bar{1} & \bar{1} \end{array} \right).$$

Dunque il gruppo $\mathbf{GL}_2(\mathbf{Z}_2)$ non è commutativo.

* * *

3.2.5. Verificare che ogni matrice antisimmetrica reale di ordine dispari ha determinante nullo (cfr. **Eserc. 2.5.7**).

Soluzione. Si noti che, $\forall A \in \mathfrak{M}_n(K)$, $\det(-A) = (-1)^n \det(A)$. Se A è antisimmetrica [cioè $A = -{}^t A$], si ha:

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

Se poi n è dispari (e la matrice è antisimmetrica), allora $\det(A) = -\det(A)$ e quindi $\det(A) = 0$.

* * *

3.2.6. È assegnata in $\mathfrak{M}_4(\mathbf{Q})$ la matrice

$$A = \begin{pmatrix} a+1 & a-1 & 0 & a \\ 0 & 1 & 1 & 1 \\ b-1 & b & b+1 & 0 \\ 1 & -1 & 0 & 0 \end{pmatrix}, \text{ con } a, b \in \mathbf{Z}.$$

Verificare, senza eseguire il calcolo diretto del determinante ma usando solo le proprietà dei determinanti, che $\det(A)$ è un multiplo intero di ab .

Soluzione. Gli interi a, b si trovano rispettivamente nella prima e terza riga di A . Si ha:

$$A^{(1)} = a(1 \ 1 \ 0 \ 1) + (1 \ -1 \ 0 \ 0),$$

$$A^{(3)} = b(1 \ 1 \ 1 \ 0) + (-1 \ 0 \ 1 \ 0).$$

Allora:

$$\det(A) = a \begin{vmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ b-1 & b & b+1 & 0 \\ 1 & -1 & 0 & 0 \end{vmatrix} + \begin{vmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ b-1 & b & b+1 & 0 \\ 1 & -1 & 0 & 0 \end{vmatrix}.$$

Il secondo determinante è nullo perché la prima e la quarta riga (della rispettiva matrice) coincidono. Dunque, operando su $A^{(3)}$, si ha:

$$\det(A) = a \begin{vmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ b-1 & b & b+1 & 0 \\ 1 & -1 & 0 & 0 \end{vmatrix} = ab \begin{vmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 1 & -1 & 0 & 0 \end{vmatrix} + a \begin{vmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ -1 & 0 & 1 & 0 \\ 1 & -1 & 0 & 0 \end{vmatrix}.$$

Di nuovo, il secondo determinante è nullo perché la somma della prima e terza riga (della rispettiva matrice) coincide con la seconda. Si conclude che

$$\det(A) = ab \begin{vmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 1 & -1 & 0 & 0 \end{vmatrix} \in ab\mathbf{Z}.$$

N.B. Se si vuole eseguire il calcolo di tale determinante [ad esempio con lo sviluppo di Laplace rispetto all'ultima colonna], si ottiene $\det(A) = 3ab$.

* * *

3.2.7. Sia $\det : \mathbf{GL}_n(K) \rightarrow K^\cdot$ l'applicazione che associa ad ogni matrice quadrata invertibile A il rispettivo determinante $\det(A)$.

Verificare che \det è un epimorfismo del gruppo $(\mathbf{GL}_n(K), \cdot)$ sul gruppo (K^\cdot, \cdot) e calcolarne il nucleo.

Soluzione. Per ogni $c \in K^\cdot$, sia D la matrice diagonale di ordine n , avente per diagonale la n -pla $(c, 1, 1, \dots, 1)$. Poiché $\det(D) = c$, l'applicazione \det è suriettiva.

Per verificare che \det è un omomorfismo, basta osservare che $\det(AB) = \det(A)\det(B)$ (teorema di Binet). Pertanto \det è un epimorfismo.

Risulta:

$$\text{Ker}(\det) = \{A \in \mathbf{GL}_n(K) \mid \det(A) = 1\}.$$

[Abbiamo così provato che le matrici il cui determinante è 1 formano un sottogruppo di $\mathbf{GL}_n(K)$].

* * *

3.2.8. Sia $H = \{A \in \mathbf{GL}_n(K) \mid \det(A) = \pm 1\}$. Verificare che H è nucleo di un omomorfismo di gruppi $\varphi : \mathbf{GL}_n(K) \rightarrow K^\cdot$. Determinare poi l'immagine $\text{Im}(\varphi)$.

Soluzione. Si ponga

$$\varphi(A) = \det(A)^2, \quad \forall A \in \mathbf{GL}_n(K).$$

In base al teorema di Binet, risulta, $\forall A, B \in \mathbf{GL}_n(K)$:

$$\varphi(AB) = \det(AB)^2 = (\det(A)\det(B))^2 = \det(A)^2\det(B)^2 = \varphi(A)\varphi(B).$$

Pertanto φ è un omomorfismo di gruppi. Si ha:

$$\text{Ker}(\varphi) = \{A \in \mathbf{GL}_n(K) \mid \det(A)^2 = 1\} = \{A \in \mathbf{GL}_n(K) \mid \det(A) = \pm 1\} = H.$$

Calcoliamo ora $\text{Im}(\varphi)$. Consideriamo in K^\cdot il sottoinsieme

$$K_1 := \{c \in K^\cdot \mid c = a^2, \exists a \in K\}$$

[cioè l'insieme dei $c \in K^\cdot$ che sono quadrati di elementi di K]. È evidente che $\text{Im}(\varphi) \subseteq K_1$ [infatti ogni $\det(A)^2 \in K_1$]. Viceversa, se $c = a^2$, allora $a \neq 0$ e la matrice diagonale A avente sulla diagonale gli elementi $a, 1, 1, \dots, 1$ è invertibile e verifica la condizione:

$$\det(A)^2 = a^2 = c.$$

Dunque $c \in \text{Im}(\varphi)$.

N.B. Abbiamo in particolare provato che K_1 è un sottogruppo del gruppo moltiplicativo (K^\cdot, \cdot) . Ad esempio, se $K = \mathbf{R}$, $K_1 = \mathbf{R}^+$.

* * *

3.3.1. Sia $A = \begin{pmatrix} 1 & -1 & 2 \\ 0 & -2 & 1 \\ 0 & -2 & 1 \\ -1 & -5 & 1 \\ 0 & 4 & -2 \end{pmatrix} \in \mathfrak{M}_{5,3}(\mathbf{R})$. Calcolare $rg(A)$, come massimo numero di colonne linearmente indipendenti di A .

Soluzione. Verifichiamo che le prime due colonne di A sono linearmente indipendenti. Sia

$$x A_{(1)} + y A_{(2)} = \mathbf{0}.$$

Tale relazione tra colonne si trasforma nel $SLO(5, 2, \mathbf{R})$

$$\begin{cases} x - y = 0 \\ -2y = 0 \\ -2y = 0 \\ -x - 5y = 0 \\ 4y = 0, \end{cases}$$

che ammette soltanto la soluzione banale $(x, y) = (0, 0)$. Poniamo ora

$$x A_{(1)} + y A_{(2)} = A_{(3)}.$$

Il corrispondente $SL(5, 2, \mathbf{R})$

$$\begin{cases} x - y = 2 \\ -2y = 1 \\ -2y = 1 \\ -x - 5y = 1 \\ 4y = -2 \end{cases}$$

ammette, come subito si verifica, soluzione $(x, y) = (\frac{3}{2}, -\frac{1}{2})$ e pertanto

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

Segue che $\mathbf{c}_A = rg(A) = 2$.

* * *

3.3.2 Calcolare il rango della seguente matrice A , interpretandolo sia come massimo numero di righe linearmente indipendenti, sia come ordine massimo dei minori non nulli:

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

Soluzione. Le prime due righe di A sono linearmente indipendenti [infatti si osserva subito che non sono proporzionali]. Poniamo $A^{(3)} = a A^{(1)} + b A^{(2)}$. Otteniamo il $SL(4, 2, \mathbf{R})$

$$\begin{cases} a + 2b = 4 \\ b = 1 \\ -a = -2 \\ 2a - b = 3. \end{cases}$$

Tale SL è compatibile, con (unica) soluzione $(a, b) = (2, 1)$. Pertanto $A^{(3)} = 2A^{(1)} + A^{(2)}$ e quindi $\langle A^{(1)}, A^{(2)}, A^{(3)} \rangle = \langle A^{(1)}, A^{(2)} \rangle$.

Ora poniamo $A^{(4)} = a A^{(1)} + b A^{(2)}$. Otteniamo il $SL(4, 2, \mathbf{R})$

$$\begin{cases} a + 2b = 5 \\ b = 2 \\ -a = -1 \\ 2a - b = 0. \end{cases}$$

Anche tale SL è compatibile, con (unica) soluzione $(a, b) = (1, 2)$. Concludiamo che

$$\langle A^{(1)}, A^{(2)}, A^{(3)}, A^{(4)} \rangle = \langle A^{(1)}, A^{(2)} \rangle.$$

Quindi $rg(A) = \mathbf{r}_A = 2$.

Fissiamo in A un minore di ordine 2 non nullo. Tra le molte possibili scelte, consideriamo il minore $\det(A(1, 2 | 1, 2)) = 1$. Tale minore possiede quattro minori orlati. Se verifichiamo che sono tutti nulli, allora $rg(A) = \rho(A) = 2$. Infatti:

$$\det(A(1, 2, 3 | 1, 2, 3)) = \begin{vmatrix} 1 & 0 & -1 \\ 2 & 1 & 0 \\ 4 & 1 & -2 \end{vmatrix} = 0, \quad \det(A(1, 2, 3 | 1, 2, 4)) = \begin{vmatrix} 1 & 0 & 2 \\ 2 & 1 & -1 \\ 4 & 1 & 3 \end{vmatrix} = 0,$$

$$\det(A(1, 2, 4 | 1, 2, 3)) = \begin{vmatrix} 1 & 0 & -1 \\ 2 & 1 & 0 \\ 5 & 2 & -1 \end{vmatrix} = 0, \quad \det(A(1, 2, 4 | 1, 2, 4)) = \begin{vmatrix} 1 & 0 & 2 \\ 2 & 1 & -1 \\ 5 & 2 & 0 \end{vmatrix} = 0.$$

* * *

3.3.3 Al variare di $a, b \in \mathbf{R}$, descrivere il rango della matrice

$$A = \begin{pmatrix} a & 1 & b \\ 1 & a & b \\ b & 0 & 1 \end{pmatrix}.$$

Soluzione. La sottomatrice $A(2, 3 | 2, 3)$ ha determinante a . Distinguiamo due casi:

$$(i) \ a \neq 0, \quad (ii) \ a = 0.$$

(i) Se $a \neq 0$, $2 \leq rg(A) \leq 3$. Risulta in particolare

$$rg(A) = 2 \iff \det(A) = 0 \iff a^2 + b^2 - ab^2 - 1 = 0.$$

(ii) Se $a = 0$, $A = \begin{pmatrix} 0 & 1 & b \\ 1 & 0 & b \\ b & 0 & 1 \end{pmatrix}$. Tale matrice ammette la sottomatrice invertibile $A(1, 2 | 1, 2)$.

Dunque $2 \leq rg(A) \leq 3$. Risulta in particolare

$$rg(A) = 2 \iff \det(A) = 0 \iff b^2 - 1 = 0 \iff b = \pm 1.$$

Concludendo,

$$rg(A) = 3 \iff \begin{cases} a^2 + b^2 - ab^2 - 1 \neq 0 \text{ e } a \neq 0 \\ \text{oppure} \\ b \neq \pm 1 \text{ e } a = 0. \end{cases}$$

Altrimenti $rg(A) = 2$.

* * *

3.3.4. Sia $A = \begin{pmatrix} 0 & a & b \\ -a & 0 & c \\ -b & -c & 0 \end{pmatrix} \in \mathfrak{M}_3(\mathbf{R})$, con $a \neq 0$.

Verificare che $rg(A) = 2$ ed esprimere la terza riga come combinazione lineare delle prime due.

Soluzione. Con un calcolo diretto [ovvero tenendo conto che A è antisimmetrica di ordine dispari] risulta subito che $\det(A) = 0$. Poiché A possiede la sottomatrice quadrata invertibile $A(1, 2 | 1, 2)$, necessariamente $rg(A) = 2$. Inoltre le prime due righe sono linearmente indipendenti e quindi la terza è combinazione lineare delle prime due. Pertanto $\exists x, y \in \mathbf{R}$ tali che

$$(0 \ a \ b)x + (-a \ 0 \ c)y = (-b \ -c \ 0),$$

ovvero il $SL(3, 2, \mathbf{R})$

$$\begin{cases} -ay = -b \\ ax = -c \\ bx + cy = 0 \end{cases}$$

è compatibile. Risulta subito: $x = -\frac{c}{a}$, $y = \frac{b}{a}$. Pertanto

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

* * *

3.3.5. Sia $A \in \mathfrak{M}_2(\mathbf{R})$. Verificare che

$$rg(A^2) < rg(A) \iff A = \begin{pmatrix} a & b \\ c & -a \end{pmatrix}, \text{ con } a^2 + bc = 0 \text{ e } a, b, c \text{ non tutti nulli.}$$

Soluzione. (\Leftarrow). Poiché $\det(A) = -(a^2 + bc) = 0$, $rg(A) < 2$. D'altra parte $A \neq \mathbf{0}$ e quindi $rg(A) \geq 1$; dunque necessariamente $rg(A) = 1$. Si verifica poi subito che $A^2 = \mathbf{0}$. Pertanto $rg(A^2) = 0$ e dunque $rg(A^2) < rg(A)$.

(\Rightarrow). Se $rg(A^2) < rg(A)$, necessariamente $\det(A) = 0$ [altrimenti A^2 sarebbe invertibile come A e dunque $rg(A^2) = rg(A) = 2$]. Segue allora che

$$0 \leq rg(A^2) < rg(A) < 2$$

e dunque necessariamente $rg(A) = 1$ e $A^2 = \mathbf{0}$. Posto $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, la condizione $A^2 = \mathbf{0}$ equivale al sistema

$$\begin{cases} a^2 + bc = 0 \\ (a+d)b = 0 \\ (a+d)c = 0 \\ bc + d^2 = 0. \end{cases}$$

Se fosse $a+d \neq 0$, allora $b = c = 0$ e quindi $a^2 = b^2 = 0$, cioè $a = d = 0$: una contraddizione. Pertanto $a+d = 0$. La matrice A verifica quindi le condizioni:

$$A = \begin{pmatrix} a & b \\ c & -a \end{pmatrix}, \text{ con } \det(A) = -(a^2 + bc) = 0 \text{ e } a, b, c \text{ non tutti nulli,}$$

come richiesto.

* * *

3.3.6. Sono assegnate le tre matrici (a valori reali) dipendenti da un parametro $a \in \mathbf{R}$:

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

Sia $M = ABC$. Determinare $rg(M)$, al variare di $a \in \mathbf{R}$.

Soluzione. La matrice M è quadrata di ordine 4. Poiché $rg(ABC) \leq rg(AB) \leq rg(B) \leq 2$, allora

$$0 \leq rg(M) \leq 2.$$

Calcoliamo la matrice M . Risulta:

$$M = ABC = \begin{pmatrix} a+a^2 & a & a^2 & 2a \\ a^2-a+1 & a & a-a^2 & 1 \\ 3a & 1 & 2a^2 & 2a+1 \\ 2+3a^2 & 3a & 2a & 2+3a \end{pmatrix}.$$

Se riusciamo ad individuare due colonne di M linearmente indipendenti ($\forall a \in \mathbf{R}$) potremo concludere che $rg(M) = 2$, $\forall a \in \mathbf{R}$.

Sceglieremo la seconda e quarta colonna di M (hanno elementi più semplici) e sia N la matrice da esse formata:

$$N = \begin{pmatrix} a & 2a \\ a & 1 \\ 1 & 2a+1 \\ 3a & 2+3a \end{pmatrix}$$

Fissiamo in N la sottomatrice $N(3|1)$ ed orliamola. Otteniamo i tre determinanti:

$$\left| \begin{array}{cc} a & 2a \\ 1 & 2a+1 \end{array} \right|, \quad \left| \begin{array}{cc} a & 1 \\ 1 & 2a+1 \end{array} \right|, \quad \left| \begin{array}{cc} 1 & 2a+1 \\ 3a & 2+3a \end{array} \right|,$$

cioè (rispettivamente)

$$a(2a-1), \quad 2a^2+a-1, \quad 2(1-3a^2).$$

Il terzo determinante si annulla $\iff a = \pm \frac{1}{\sqrt{3}}$. Ma, per ciascuno di tali valori, il primo determinante non è nullo. Si conclude che $rg(N) = 2$, $\forall a \in \mathbf{R}$, e quindi $rg(M) = 2$, $\forall a \in \mathbf{R}$.

* * *

3.4.1 Risolvere il seguente $SLO(4, 3, \mathbf{R})$:

$$\begin{cases} x + y + z = 0 \\ -y + z = 0 \\ -x - 2y = 0 \\ 2x + 3y + z = 0. \end{cases}$$

Soluzione. Il SLO ha matrice dei coefficienti $A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & -1 & 1 \\ -1 & -2 & 0 \\ 2 & 3 & 1 \end{pmatrix}$.

Si verifica subito che $rg(A) = 2$ [infatti le prime due colonne di A sono linearmente indipendenti e $A^{(3)} = 2A^{(1)} - A^{(2)}$].

Poiché ad esempio $A(1, 2 | 1, 2)$ è invertibile, il SLO dato è equivalente al SLO formato dalle prime due equazioni. Si tratta di un sistema a scala. Posto $z = t$, si ha $y = t$, $x = -2t$. Pertanto le soluzioni del SLO assegnato formano il sottospazio vettoriale di \mathbf{R}^3 :

$$\Sigma_0 = \{(-2t, t, t), \forall t \in \mathbf{R}\} = \langle(-2, 1, 1)\rangle.$$

* * *

3.4.2 Risolvere, al variare di $a, b \in \mathbf{R}$, il seguente $SL(3, 3, \mathbf{R})$:

$$\begin{cases} ax + by = 0 \\ y + bz = a \\ x - az = b. \end{cases}$$

Soluzione. La matrice A dei coefficienti e la matrice completa M del SL assegnato sono rispettivamente

$$A = \begin{pmatrix} a & b & 0 \\ 0 & 1 & b \\ 1 & 0 & -a \end{pmatrix}, \quad M = \begin{pmatrix} a & b & 0 & 0 \\ 0 & 1 & b & a \\ 1 & 0 & -a & b \end{pmatrix}.$$

La sottomatrice $A(2, 3 | 1, 2)$ ha determinante non nullo. Quindi $2 \leq rg(A) \leq 3$. Si ha:

$$rg(A) = 2 \iff \det(A) = 0 \iff b^2 - a^2 = 0 \iff b = \pm a.$$

Altrimenti $rg(A) = 3$.

Se $rg(A) = 3$ il SL è compatibile ed ha $\infty^{3-3} = 1$ soluzione (x, y, z) , data dalla formula di Cramer:

$$\begin{aligned} x &= \frac{1}{b^2 - a^2} \begin{vmatrix} 0 & b & 0 \\ a & 1 & b \\ b & 0 & -a \end{vmatrix} = \frac{b^3 + a^2b}{b^2 - a^2} = b \frac{b^2 + a^2}{b^2 - a^2}, \\ y &= \frac{1}{b^2 - a^2} \begin{vmatrix} a & 0 & 0 \\ 0 & a & b \\ 1 & b & -a \end{vmatrix} = \frac{-a^3 - ab^2}{b^2 - a^2} = a \frac{b^2 + a^2}{b^2 - a^2}, \\ z &= \frac{1}{b^2 - a^2} \begin{vmatrix} a & b & 0 \\ 0 & 1 & a \\ 1 & 0 & b \end{vmatrix} = \frac{2ab}{b^2 - a^2}. \end{aligned}$$

Assumiamo ora $rg(A) = 2$ [cioè $b = \pm a$] e calcoliamo $rg(M)$. Ovviamente $2 \leq rg(M) \leq 3$. Orlando $M(2, 3 | 1, 2)$ si ottiene:

$$rg(M) = 2 \iff \det(A) = \begin{vmatrix} a & b & 0 \\ 0 & 1 & a \\ 1 & 0 & b \end{vmatrix} = 0 \iff \begin{cases} b^2 - a^2 = 0 \\ ab = 0 \end{cases}$$

e tale sistema (non lineare) ammette un'unica soluzione: $(a, b) = (0, 0)$. Pertanto:

- se $a = b = 0$, il SL è compatibile ed ha ∞^1 soluzioni. In tal caso il SL dato si riduce al $SL(2, 3, \mathbf{R})$ $\begin{cases} y = 0 \\ x = 0 \end{cases}$ e le sue soluzioni sono quindi date da $\{(0, 0, t), \forall t \in \mathbf{R}\}$.

- se invece $b = \pm a$ e $a \neq 0$ [ovvero $b \neq 0$], il SL dato è incompatibile.

* * *

3.4.3. Utilizzando il teorema di Rouché-Capelli discutere la risoluzione, al variare di tre parametri $a, b, c \in \mathbf{R}$, del $SL(1, 2, \mathbf{R})$: $\{ax + by = c\}$.

Soluzione. Risulta, in base al teorema di Rouché-Capelli:

$$\text{il } SL \text{ è compatibile} \iff rg((a \ 1 \ c)) = rg((a \ b)) \iff \begin{cases} (a, b) \neq (0, 0) \\ \text{oppure} \\ a = b = c = 0. \end{cases}$$

Nel primo caso $rg((a \ b)) = rg((a \ 1 \ c)) = 1$ ed il SL ha ∞^1 soluzioni. Distinguiamo due casi:

- (i) $a \neq 0$. Il SL si risolve ponendo $y = t$ e si ha:

$$\Sigma = \left\{ \left(\frac{c}{a} - \frac{b}{a}t, t \right), \forall t \in \mathbf{R} \right\}.$$

- (ii) $a = 0$. Il SL si risolve ponendo $x = t$ e si ha:

$$\Sigma = \left\{ (t, \frac{c}{b}), \forall t \in \mathbf{R} \right\}.$$

Nel secondo caso il sistema diventa $\{0 = 0$ ed ovviamente ha ∞^2 soluzioni. L'insieme delle sue soluzioni è $\Sigma = \mathbf{R}^2$.

* * *

3.4.4. Rifare l'**Esercizio 3.1.3** (senza usare l'algoritmo di Gauss).

Soluzione. Il SLO assegnato ha matrice dei coefficienti

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

Ovviamente $1 \leq rg(A) \leq 2$. Si ha (in base al principio degli orlati applicato ad $A(1 | 1)$):

$$rg(A) = 1 \iff \begin{vmatrix} 1 & a \\ 2 & 2 \end{vmatrix} = \begin{vmatrix} 1 & a \\ a & 0 \end{vmatrix} = 0 \iff a = 0 = 1 \text{ (incomp.)}.$$

Dunque $rg(A) = 2, \forall a \in \mathbf{R}$.

Se $a \neq 0$, possiamo scegliere in A la sottomatrice invertibile $A(1, 3 | 1, 2) = \begin{pmatrix} 1 & a \\ a & 0 \end{pmatrix}$.

Se $a = 0$, scegliamo invece in A la sottomatrice invertibile $A(1, 2 | 1, 2) = \begin{pmatrix} 1 & 0 \\ 2 & 2 \end{pmatrix}$.

In entrambi i casi il SLO dato si riduce ad un $SLO(2, 2, \mathbf{R})$ avente matrice dei coefficienti di rango 2. Tale SLO è quindi privo di autosoluzioni, $\forall a \in \mathbf{R}$.

* * *

3.4.5. Rifare l'**Esercizio 3.1.4** (senza usare l'algoritmo di Gauss).

Soluzione. Il SLO assegnato ha matrice dei coefficienti

$$A = \begin{pmatrix} 0 & a & b \\ -a & 0 & 1 \\ -b & -1 & 0 \end{pmatrix}.$$

Tale matrice ha rango 2 [infatti è antisimmetrica di ordine dispari e quindi $\det(A) = 0$; inoltre possiede la sottomatrice invertibile $A(2, 3 | 2, 3)$].

Il SLO assegnato è equivalente al $SLO(2, 3, \mathbf{R})$

$$\begin{cases} -ax + z = 0 \\ -bx - y = 0, \end{cases} \text{ cioè } \begin{cases} z = ax \\ y = -bx. \end{cases}$$

Lo spazio vettoriale Σ_0 delle soluzioni di tale SLO è

$$\Sigma_0 = \{(t, -bt, at), \forall t \in \mathbf{R}\} = \langle(1, -b, a)\rangle.$$

N.B. Si osservi che l'ipotesi che a, b siano parametri non nulli non è stata utilizzata. Era stata introdotta nell'**Esercizio 3.1.4** soltanto per agevolare l'algoritmo di Gauss.

* * *

3.4.6 È assegnato il $SL(3, 3, \mathbf{R})$:

$$\begin{cases} ay + bz = c \\ -ax + az = -c \\ -bx - ay = 0, \end{cases}$$

dipendente da tre parametri $a, b, c \in \mathbf{R}$, con $a, b \neq 0$. Determinare per quali valori dei parametri il SL è compatibile e scriverne l'insieme Σ delle soluzioni.

Soluzione. Il SL ha matrice dei coefficienti

$$A = \begin{pmatrix} 0 & a & b \\ -a & 0 & a \\ -b & -a & 0 \end{pmatrix}.$$

Tale matrice ha determinante nullo [come del resto tutte le matrici antisimmetriche di ordine dispari].

Dunque $rg(A) \leq 2$. Poiché $a \neq 0$, A possiede ad esempio il minore non nullo $\begin{vmatrix} a & b \\ 0 & a \end{vmatrix} = a^2$. Dunque $rg(A) = 2$, $\forall a, b \in \mathbf{R}$.

Dal teorema di Rouché-Capelli, il SL è compatibile per tutti e soli i valori $c \in \mathbf{R}$ per cui la matrice completa

$$M = \begin{pmatrix} 0 & a & b & c \\ -a & 0 & a & -c \\ -b & -a & 0 & 0 \end{pmatrix}$$

del SL ha rango 2. Risulta:

$$rg(M) = 2 \iff \begin{vmatrix} a & b & c \\ 0 & a & -c \\ -a & 0 & 0 \end{vmatrix} = 0 \iff -a(-bc - ac) = 0 \iff ac(a + b) = 0 \iff \begin{cases} c = 0 \\ b = -a. \end{cases}$$

Pertanto il SL è compatibile $\iff c = 0$ oppure $b = -a$ ($\neq 0$). In entrambi i casi ha $\infty^{3-2} = \infty^1$ soluzioni.

Sia $c = 0$. In tal caso il SL è omogeneo ed è ad esempio equivalente al $SLO(2, 3, \mathbf{R})$ formato dalle prime due equazioni del SL dato. Le sue ∞^1 soluzioni sono proporzionali a

$$\left(\begin{vmatrix} a & b \\ 0 & a \end{vmatrix}, -\begin{vmatrix} 0 & b \\ -a & a \end{vmatrix}, \begin{vmatrix} 0 & a \\ -a & 0 \end{vmatrix} \right) = (a^2, -ab, a^2).$$

Dunque

$$\Sigma = \langle (a^2, -ab, a^2) \rangle.$$

Sia ora $b = -a$ ($\neq 0$). Anche in tal caso il SL è equivalente al $SL(2, 3, \mathbf{R})$ formato dalle prime due equazioni del SL dato, cioè

$$\begin{cases} ay - az = c \\ -ax + az = -c, \end{cases} \quad \text{ovvero} \quad \begin{cases} y - z = \frac{c}{a} \\ x + z = -\frac{c}{a}. \end{cases}$$

Posto $x = t$, si ottiene $z = -\frac{c}{a} + t$ e $y = t$. Dunque

$$\Sigma = \{(t, t, -\frac{c}{a} + t), \forall t \in \mathbf{R}\}.$$

* * *

3.4.7 È assegnato il $SLO(2, 3, \mathbf{R})$:

$$\begin{cases} bx + ay = 0 \\ y + az = 0, \end{cases}$$

dipendente da due parametri $a, b \in \mathbf{R}$. Risolvere tale SLO , al variare dei parametri.

Soluzione. La matrice dei coefficienti del SLO è

$$A = \begin{pmatrix} b & a & 0 \\ 0 & 1 & a \end{pmatrix}.$$

Ovviamente $1 \leq rg(A) \leq 2$ [infatti $|A(2|2)| = 1 \neq 0$]. Risulta:

$$rg(A) = 1 \iff \begin{vmatrix} b & a \\ 0 & 1 \end{vmatrix} = \begin{vmatrix} a & 0 \\ 1 & a \end{vmatrix} = 0 \iff b = a^2 = 0 \iff a = b = 0.$$

Se $(a, b) \neq (0, 0)$, $rg(A) = 2$ ed il *SLO* ha ∞^1 soluzioni, proporzionali all'autosoluzione

$$\left(\begin{vmatrix} a & 0 \\ 1 & a \end{vmatrix}, -\begin{vmatrix} b & 0 \\ 0 & a \end{vmatrix}, \begin{vmatrix} b & a \\ 0 & 1 \end{vmatrix} \right) = (a^2, -ab, b).$$

Pertanto l'insieme delle soluzioni è il sottospazio vettoriale di \mathbf{R}^3 :

$$\Sigma_0 = \langle (a^2, -ab, b) \rangle.$$

Se $(a, b) = (0, 0)$, $rg(A) = 1$ ed il *SLO* dato si riduce al *SLO* $(1, 3, \mathbf{R})$

$$\{ y = 0 \}.$$

In tal caso il *SLO* ha ∞^2 soluzioni, date dal sottospazio vettoriale di \mathbf{R}^3 :

$$\Sigma_0 = \langle (1, 0, 0), (0, 0, 1) \rangle.$$

* * *

3.4.8. Risolvere il seguente $SL(3, 5, \mathbf{Z}_3)$: $\begin{cases} x_1 + x_3 + x_5 = \bar{2} \\ x_2 + x_3 = \bar{1} \\ x_1 + x_4 + x_5 = \bar{0}. \end{cases}$

Quante sono le sue soluzioni? Determinare una base dello spazio delle soluzioni del *SLO* associato.

Soluzione. La matrice del *SL* è

$$A = \begin{pmatrix} \bar{1} & \bar{0} & \bar{1} & \bar{0} & \bar{1} \\ \bar{0} & \bar{1} & \bar{1} & \bar{0} & \bar{0} \\ \bar{1} & \bar{0} & \bar{0} & \bar{1} & \bar{1} \end{pmatrix} \in \mathfrak{M}_{3,5}(\mathbf{Z}_3).$$

Risulta $rg(A) = 3$ [in quanto ad esempio le prime tre colonne sono linearmente indipendenti]. Ne segue che il *SL* è compatibile.

Per calcolarne le soluzioni poniamo $x_4 = t$, $x_5 = s$, con $t, s \in \mathbf{Z}_3$. Allora

$$\begin{cases} x_1 + x_3 = \bar{2} - s = \bar{2} + \bar{2}s \\ x_2 + x_3 = \bar{1} \\ x_1 = -t - s = \bar{2}t + \bar{2}s. \end{cases}$$

Dalla terza equazione, $x_1 = \bar{2}(t+s)$. Sostituendo nella prima, $x_3 = \bar{2} + \bar{2}s - \bar{2}(t+s) = \bar{2} + t$. Infine, sostituendo nella seconda, $x_2 = \bar{1} - (\bar{2} + t) = \bar{2} + \bar{2}t$.

L'insieme delle soluzioni è quindi

$$\Sigma = \{(\bar{2}(t+s), \bar{2} + \bar{2}t, \bar{2} + t, t, s), \forall t, s \in \mathbf{Z}_3\}$$

Il *SL* dato ha esattamente 9 soluzioni [ottenute facendo variare t, s in \mathbf{Z}_3].

Una soluzione di tale *SL* è $(\bar{0}, \bar{2}, \bar{2}, \bar{0}, \bar{0})$. Sottraendo tale soluzione a Σ si ottengono le soluzioni del *SLO* associato:

$$\Sigma_0 = \{(\bar{2}(t+s), \bar{2}t, t, t, s), \forall t, s \in \mathbf{Z}_3\} = \langle (\bar{2}, \bar{2}, \bar{1}, \bar{1}, \bar{0}), (\bar{2}, \bar{0}, \bar{0}, \bar{0}, \bar{1}) \rangle.$$

* * *

3.4.9. È assegnato il seguente $SL(3, 3, \mathbf{Z}_3)$, dipendente da un parametro $a \in \mathbf{Z}_3$:

$$\begin{cases} x + \bar{2}z = a \\ \bar{2}x + \bar{2}y + z = a \\ \bar{2}x + z = \bar{2} + a. \end{cases}$$

Determinare per quali eventuali $a \in \mathbf{Z}_3$ il *SL* è compatibile e calcolarne le soluzioni.

Soluzione. Il *SL* ha matrice dei coefficienti

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

Risulta: $det(A) = \bar{2} - \bar{8} = -\bar{6} = \bar{0}$. Pertanto $rg(A) \leq 2$. Ad esempio A contiene la sottomatrice invertibile

$$B = A(1, 2 \mid 1, 2) = \begin{pmatrix} \bar{1} & \bar{0} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

Ne segue che $rg(A) = 2$, $\forall a \in \mathbf{Z}_3$. Risulta:

$$\begin{aligned} \text{il } SL \text{ è compatibile} &\iff \text{la matrice completa del } SL \text{ ha rango 2} \iff \\ &\iff \text{è nullo l'orlato di } B \text{ con la colonna dei termini noti} \iff \\ &\iff \begin{vmatrix} \bar{1} & \bar{0} & a \\ \frac{1}{2} & \frac{1}{2} & a \\ \frac{1}{2} & \bar{0} & a + \bar{2} \end{vmatrix} = 0 \iff \bar{2}(a + \bar{2}) - \bar{4}a = \bar{0} \iff a = \bar{2}. \end{aligned}$$

Per $a \neq \bar{2}$ il SL è incompatibile.

Sia $a = \bar{2}$. In tal caso il SL ha $\infty^1 [= 3]$ soluzioni. Per ottenere tali soluzioni eliminiamo la terza equazione del SL e poniamo $z = t \in \mathbf{Z}_3$. Si ottiene il SL

$$\begin{cases} x = \bar{2} - \bar{2}t \\ \bar{2}x + \bar{2}y = \bar{2} - t, \end{cases}$$

che, risolto, fornisce $x = \bar{2} + t$, $y = \bar{2}$ e pertanto

$$\Sigma = \{(\bar{2} + t, \bar{2}, t), \forall t \in \mathbf{Z}_3\} = \{(\bar{2}, \bar{2}, \bar{0}), (\bar{0}, \bar{2}, \bar{1}), (\bar{1}, \bar{2}, \bar{2})\}.$$

* * *

3.4.10 È assegnato il $SL(3, 3, \mathbf{R})$:

$$\begin{cases} ax - y + bz = -a \\ x - bz = 0 \\ ax + 2y = b, \end{cases}$$

dipendente da due parametri $a, b \in \mathbf{R}$. Risolvere tale SL , al variare dei parametri.

Soluzione. La matrice dei coefficienti del SL è

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

Risulta: $2 \leq rg(A) \leq 3$ [infatti $|A(2, 3 \mid 1, 2)| = 2 \neq 0$]. Si ha:

$$rg(A) = 2 \iff \det(A) = 0 \iff b(2 + 3a) = 0 \iff b = 0 \text{ oppure } a = -\frac{2}{3}.$$

Se $a \neq -\frac{2}{3}$ e $b \neq 0$, il SL è compatibile, con un'unica soluzione, data dalla formula di Cramer:

$$\frac{1}{\det(A)} \left(\begin{vmatrix} -a & -1 & b \\ 0 & 0 & -b \\ b & 2 & 0 \end{vmatrix}, \begin{vmatrix} a & -a & b \\ 1 & 0 & -b \\ a & b & 0 \end{vmatrix}, \begin{vmatrix} a & -1 & -a \\ 1 & 0 & 0 \\ a & 2 & b \end{vmatrix} \right).$$

Sia $b = 0$. In tal caso $rg(A) = 2$ e la matrice completa del SL è

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

Risulta:

$$rg(M) = 2 \iff \begin{vmatrix} a & -1 & -a \\ 1 & 0 & 0 \\ a & 2 & 0 \end{vmatrix} = 0 \iff -2a = 0 \iff a = 0.$$

Dunque, se $b = 0$ e $a \neq 0$ il SL è incompatibile. Se invece $a = b = 0$, il SL ha ∞^1 soluzioni. Per ottenerle eliminiamo la prima equazione del SL . Il SL si riduce al $SLO(2, 3, \mathbf{R})$

$$\begin{cases} x = 0 \\ 2y = 0, \end{cases}$$

avente come spazio delle soluzioni $\langle e_3 \rangle \subset \mathbf{R}^3$.

Sia infine $b \neq 0$ e $a = -\frac{2}{3}$. Anche in tal caso $rg(A) = 2$ e la matrice completa del SL è

$$M = \begin{pmatrix} -\frac{2}{3} & -1 & b & \frac{2}{3} \\ 1 & 0 & -b & 0 \\ -\frac{2}{3} & 2 & 0 & b \end{pmatrix}.$$

Risulta:

$$rg(M) = 2 \iff \begin{vmatrix} -\frac{2}{3} & -1 & \frac{2}{3} \\ 1 & 0 & 0 \\ -\frac{2}{3} & 2 & b \end{vmatrix} = 0 \iff b = -\frac{4}{3}.$$

Dunque, se $a = -\frac{2}{3}$, $b \neq -\frac{4}{3}$ il SL è incompatible. Se invece $a = -\frac{2}{3}$, $b = 2a$, il SL ha ∞^1 soluzioni. Per ottenerle eliminiamo la prima equazione del SL . Il SL si riduce al $SL(2, 3, \mathbf{R})$

$$\begin{cases} x + \frac{4}{3}z = 0 \\ -\frac{2}{3}x + 2y = -\frac{4}{3}. \end{cases}$$

Posto $z = t$, si ottiene che le soluzioni del SL sono date dall'insieme

$$\Sigma = \left\{ \left(-\frac{4}{3}t, -\frac{4}{3} - \frac{8}{9}t, t \right), \forall t \in \mathbf{R} \right\}.$$

* * *

3.4.11. Al variare dei parametri $a, b \in \mathbf{R}$, risolvere il seguente $SL(2, 2, \mathbf{R})$

$$\begin{cases} ax + by = a - b \\ bx + ay = b - a. \end{cases}$$

Soluzione. La matrice A dei coefficienti e la matrice completa M del SL sono rispettivamente

$$A = \begin{pmatrix} a & b \\ b & a \end{pmatrix}, \quad M = \begin{pmatrix} a & b & a - b \\ b & a & b - a \end{pmatrix}.$$

Calcoliamo $rg(A)$. Si ha:

$$\begin{aligned} rg(A) = 0 &\iff a = b = 0; \\ rg(A) = 1 &\iff \det(A) = 0 \text{ e } (a, b) \neq (0, 0) \iff a^2 = b^2 \text{ e } (a, b) \neq (0, 0) \iff \\ &\iff b = \pm a \text{ e } a \neq 0; \\ rg(A) = 2 &\iff \det(A) \neq 0 \iff b \neq \pm a. \end{aligned}$$

Sia $rg(A) = 0$. In tal caso $M = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ e quindi anche $rg(M) = 0$. Il SL è compatibile e ha $\infty^{2-0} = \infty^2$ soluzioni. Pertanto l'insieme delle soluzioni è lo spazio vettoriale $\Sigma_0 = \mathbf{R}^2$.

Sia $rg(A) = 1$. Vanno discussi due casi: $b = a$, $a \neq 0$; $b = -a$, $a \neq 0$.

Nel primo caso, $M = \begin{pmatrix} a & a & 0 \\ a & a & 0 \end{pmatrix}$ e quindi anche $rg(M) = 1$. Il SL è compatibile e si riduce al $SLO(1, 2, \mathbf{R})$ $\{ax + ay = 0\}$, che ha soluzioni $\Sigma_0 = \{(t, -t), \forall t \in \mathbf{R}\} = \langle(1, -1)\rangle$.

Nel secondo caso, $M = \begin{pmatrix} a & -a & 2a \\ -a & a & -2a \end{pmatrix}$ e quindi anche $rg(M) = 1$. Il SL è compatibile e si riduce al $SL(1, 2, \mathbf{R})$ $\{ax - ay = 2a\}$, che ha soluzioni $\Sigma = \{(2 + t, t), \forall t \in \mathbf{R}\}$.

Sia $rg(A) = 2$. In tal caso il SL ha un'unica soluzione, che si ottiene con la formula di Cramer:

$$x = \frac{\begin{vmatrix} a-b & b \\ b-a & a \end{vmatrix}}{\begin{vmatrix} a & b \\ b & a \end{vmatrix}} = 1, \quad y = \frac{\begin{vmatrix} a & a-b \\ b & b-a \end{vmatrix}}{\begin{vmatrix} a & b \\ b & a \end{vmatrix}} = -1.$$

Dunque $\Sigma = \{(1, -1)\}$.

* * *

ESERCIZI PROPOSTI

Capitolo 4

4.1.1. Sia $V = V_{\mathbf{R}}^3$, con base $\mathbf{E} = (\underline{e}_1 \ \underline{e}_2 \ \underline{e}_3)$. Sono assegnati in V i tre vettori

$$\underline{u}_1 = \underline{e}_1 - \underline{e}_2, \quad \underline{u}_2 = \underline{e}_2 - \underline{e}_3, \quad \underline{u}_3 = \underline{e}_3 + \underline{e}_1.$$

- (i) Verificare che $\mathbf{F} = (\underline{u}_1 \ \underline{u}_2 \ \underline{u}_3)$ è una base di V .
- (ii) Esprimere in base \mathbf{F} il vettore $\underline{v} = \underline{e}_1 + 2\underline{e}_2 + 3\underline{e}_3$.
- (iii) Scrivere la formula di cambiamento di coordinate di vettore dalla base \mathbf{E} alla base \mathbf{F} .
- (iv) Esiste in V un vettore non nullo avente le stesse coordinate nelle due basi?

Soluzione. (i) Si ha:

$$\mathbf{F} = \mathbf{E} B, \quad \text{con } B = \begin{pmatrix} 1 & 0 & 1 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix}.$$

Poiché $\det(B) \neq 0$, allora $B \in \mathbf{GL}_3(\mathbf{R})$ e quindi \mathbf{F} è una base di V .

(ii) Si ha:

$$\underline{v} = \mathbf{E} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \mathbf{F} \mathbf{y} = \mathbf{E} B \mathbf{y} \quad \text{e dunque } B \mathbf{y} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \quad \text{cioè} \quad \begin{cases} y_1 + y_3 = 1 \\ -y_1 + y_2 = 2 \\ -y_2 + y_3 = 3. \end{cases}$$

Per ottenere le coordinate di \underline{v} in base \mathbf{F} basta risolvere tale SL. Utilizzando la formula di Cramer, si ottiene la soluzione $(-2, 0, 3)$. Dunque $\underline{v} = -2\underline{u}_1 + 3\underline{u}_3$.

(iii) Se $\underline{v} = \mathbf{E} \mathbf{x} = \mathbf{F} \mathbf{y}$, la formula richiesta è $\mathbf{y} = B^{-1} \mathbf{x}$. Risulta:

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

Quindi

$$\begin{cases} y_1 = \frac{1}{2} (x_1 - x_2 - x_3) \\ y_2 = \frac{1}{2} (x_1 + x_2 - x_3) \\ y_3 = \frac{1}{2} (x_1 + x_2 + x_3). \end{cases}$$

(iv) Sia $\underline{v} = \mathbf{E} \mathbf{x}$ un vettore non nullo. Tale vettore ha le stesse coordinate rispetto alle due basi $\iff \mathbf{E} \mathbf{x} = \mathbf{F} \mathbf{x} \iff \mathbf{E} \mathbf{x} = \mathbf{E} B \mathbf{x} \iff \mathbf{x} = B \mathbf{x} \iff (B - I_3) \mathbf{x} = \mathbf{0}$. Dunque un tale vettore esiste \iff il $SLO(3, 3, \mathbf{R})$ $(B - I_3)X = \mathbf{0}$ ammette autosoluzioni $\iff rg((B - I_3)) < 3$. Ma risulta:

$$B - I_3 = \begin{pmatrix} 0 & 0 & 1 \\ -1 & 0 & 0 \\ 0 & -1 & 0 \end{pmatrix} \quad \text{e} \quad \det(B - I_3) = 1.$$

Dunque nessun vettore non nullo ha le stesse coordinate nelle due basi.

* * *

4.1.2. Sia $V = V_{\mathbf{R}}^3$, con base \mathbf{E} . Sia $\mathbf{F} = (\underline{f}_1 \ \underline{f}_2 \ \underline{f}_3)$ un'altra base. Indichiamo con \mathbf{x} la colonna delle coordinate di un generico vettore $\underline{v} \in V$ rispetto alla base \mathbf{E} e con \mathbf{y} la colonna delle coordinate dello stesso vettore rispetto alla base \mathbf{F} . Determinare la base \mathbf{F} in funzione di \mathbf{E} , sapendo che

$$\begin{cases} y_1 = x_1 - x_2 \\ y_2 = x_2 - x_3 \\ y_3 = x_1 + x_3. \end{cases}$$

Soluzione. Scriviamo i dati dell'esercizio in forma matriciale:

$$\mathbf{y} = C\mathbf{x}, \text{ con } C = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 1 & 0 & 1 \end{pmatrix}.$$

[Si noti che $C \in \mathbf{GL}_3(\mathbf{R})$, essendo $\det(C) = 2$]. Dobbiamo determinare una matrice $B \in \mathbf{GL}_3(\mathbf{R})$ tale che $\mathbf{F} = \mathbf{E}B$. Si ha:

$$\underline{v} = \mathbf{E}\mathbf{x} = \mathbf{F}\mathbf{y} = \mathbf{E}B\mathbf{y} \text{ e dunque } \mathbf{x} = B\mathbf{y}.$$

Da $\mathbf{x} = B\mathbf{y}$ e $\mathbf{y} = C\mathbf{x}$ segue:

$$\mathbf{x} = BC\mathbf{x}, \text{ ovvero } (I_3 - BC)\mathbf{x} = \mathbf{0}, \forall \mathbf{x} \in \mathfrak{M}_{3,1}(\mathbf{R}).$$

Ciò significa che il $SLO(3, 3, \mathbf{R})$ $(I_3 - BC)X = \mathbf{0}$ ammette ∞^3 soluzioni, da cui $rg(I_3 - BC) = 0$. Pertanto $BC = I_3$, cioè $B = C^{-1}$. La matrice B cercata è quindi $B = C^{-1}$ e risulta:

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

Pertanto

$$\underline{f}_1 = \frac{1}{2}(\underline{e}_1 - \underline{e}_2 - \underline{e}_3), \quad \underline{f}_2 = \frac{1}{2}(\underline{e}_1 + \underline{e}_2 - \underline{e}_3), \quad \underline{f}_3 = \frac{1}{2}(\underline{e}_1 + \underline{e}_2 + \underline{e}_3).$$

N.B. Potevamo ugualmente ottenere $\underline{f}_1, \underline{f}_2, \underline{f}_3$ (in base \mathbf{E}), senza usare il linguaggio matriciale, risolvendo ordinatamente i tre $SL(3, 3, \mathbf{R})$:

$$\begin{cases} x_1 - x_2 = 1 \\ x_2 - x_3 = 0 \\ x_1 + x_3 = 0, \end{cases} \quad \begin{cases} x_1 - x_2 = 0 \\ x_2 - x_3 = 1 \\ x_1 + x_3 = 0, \end{cases} \quad \begin{cases} x_1 - x_2 = 0 \\ x_2 - x_3 = 0 \\ x_1 + x_3 = 1. \end{cases}$$

* * *

4.1.3. Siano $\{\underline{e}_1, \underline{e}_2, \underline{e}_3\}$ e $\{\underline{f}_1, \underline{f}_2, \underline{f}_3\}$ due basi di uno spazio vettoriale $V = V_{\mathbf{R}}^3$. Sia:

$$\underline{f}_1 = \underline{e}_1 - \underline{e}_2, \quad \underline{f}_2 = -\underline{e}_1 + \underline{e}_3, \quad \underline{f}_3 = 2\underline{e}_1 + \underline{e}_2.$$

Assegnato il sottospazio vettoriale $W = \langle \underline{e}_1 - \underline{e}_3, \underline{e}_2 + \underline{e}_3 \rangle$, determinarne un sistema di generatori espresso rispetto alla base $\{\underline{f}_1, \underline{f}_2, \underline{f}_3\}$.

Soluzione. Sia $\mathbf{E} = (\underline{e}_1 \ \underline{e}_2 \ \underline{e}_3)$ ed $\mathbf{F} = (\underline{f}_1 \ \underline{f}_2 \ \underline{f}_3)$. Risulta: $\mathbf{F} = \mathbf{E}C$, con

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

Risulta inoltre: $(\underline{e}_1 - \underline{e}_3 \ \underline{e}_2 + \underline{e}_3) = \mathbf{E}B$, con $B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -1 & 1 \end{pmatrix}$. Pertanto

$$(\underline{e}_1 - \underline{e}_3 \ \underline{e}_2 + \underline{e}_3) = \mathbf{F}C^{-1}B.$$

I due generatori di W in base \mathbf{F} hanno coordinate date dalla matrice $C^{-1}B$. Si ha:

$$C^{-1} = -\frac{1}{3} \begin{pmatrix} -1 & 2 & -1 \\ 0 & 0 & -3 \\ -1 & -1 & -1 \end{pmatrix}$$

e

$$C^{-1}B = -\frac{1}{3} \begin{pmatrix} -1 & 2 & -1 \\ 0 & 0 & -3 \\ -1 & -1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 0 & -\frac{1}{3} \\ -1 & 1 \\ 0 & \frac{2}{3} \end{pmatrix}.$$

Dunque $W = \langle -\underline{f}_2, -\frac{1}{3}\underline{f}_1 + \underline{f}_2 + \frac{2}{3}\underline{f}_3 \rangle$.

* * *

4.1.4. Sono assegnate in $V = \mathfrak{M}_2(\mathbf{R})$ le quattro matrici

$$J_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad J_2 = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \quad J_3 = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}, \quad J_4 = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

(i) Verificare che $\{J_1, J_2, J_3, J_4\}$ è una base di V .

(ii) Esprimere in tale base la matrice $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$.

Soluzione. (i) Si consideri la base canonica $\{E^{11}, E^{12}, E^{21}, E^{22}\}$ di $V = \mathfrak{M}_2(\mathbf{R})$. Posto:

$$\mathbf{E} = (E^{11} \quad E^{12} \quad E^{21} \quad E^{22}), \quad \mathbf{F} = (J_1 \quad J_2 \quad J_3 \quad J_4),$$

risulta:

$$\mathbf{F} = \mathbf{E} C, \text{ con } C = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Poiché $C \in \mathbf{GL}_4(\mathbf{R})$, \mathbf{F} è una base di $V = \mathfrak{M}_2(\mathbf{R})$.

(ii) Ovviamente

$$A = 1E^{11} + 2E^{12} + 3E^{21} + 4E^{22} = \mathbf{E} \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}.$$

Si ponga:

$$A = \mathbf{F} \mathbf{y}, \text{ con } \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} \in \mathfrak{M}_{4,1}(\mathbf{R}).$$

Allora:

$$A = \mathbf{F} \mathbf{y} = \mathbf{E} C \mathbf{y} = \mathbf{E} \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix} \text{ e dunque } C \mathbf{y} = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix},$$

ovvero:

$$\begin{cases} y_1 + y_2 + y_3 + y_4 = 1 \\ y_2 + y_3 + y_4 = 2 \\ y_3 + y_4 = 3 \\ y_4 = 4. \end{cases}$$

Risolvendo tale SL (a scala) si ottiene $y_4 = 4$, $y_3 = y_2 = y_1 = -1$ e pertanto

$$A = -J_1 - J_2 - J_3 + 4J_4.$$

* * *

4.1.5. Sia $V = V_{\mathbf{R}}^3$, con base \mathbf{E} . Siano $\underline{f}_1, \underline{f}_2 \in V$ tali che

$$\underline{f}_1 = \underline{e}_1 - \underline{e}_2 + \underline{e}_3, \quad \underline{f}_2 = \underline{e}_2 + \underline{e}_3.$$

Verificare se esiste una base $\mathbf{F} = (\underline{f}_1 \quad \underline{f}_2 \quad \underline{f}_3)$ tale che un vettore $\underline{v} \in V$ abbia in base \mathbf{E} coordinate $(1, 1, 1)$ ed in base \mathbf{F} coordinate $(0, 1, 1)$.

Soluzione. Indichiamo con \underline{f}_3 un vettore non noto, avente coordinate $(a, b, c) \in \mathbf{R}^3$, in base \mathbf{E} . Se tale vettore esiste ed è linearmente indipendente da $\underline{f}_1, \underline{f}_2$, otterremo una base $\mathbf{F} = (\underline{f}_1 \quad \underline{f}_2 \quad \underline{f}_3)$, con $\mathbf{F} = \mathbf{E} C$, dove

$$C = \begin{pmatrix} 1 & 0 & a \\ -1 & 1 & b \\ 1 & 1 & c \end{pmatrix} \in \mathbf{GL}_3(\mathbf{R}).$$

Deve risultare

$$\underline{v} = \mathbf{E} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \mathbf{F} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = \mathbf{E} C \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}.$$

Ne segue

$$\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = C \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \text{ cioè } \begin{cases} a = 1 \\ 1 + b = 1 \\ 1 + c = 1. \end{cases}$$

Tale sistema ammette la soluzione $(a, b, c) = (1, 0, 0)$. Per concludere basta verificare che la matrice C ottenuta con tali valori dei parametri è invertibile [si verifica infatti che $\det(C) = -2$]. Pertanto il problema ammette l'unica soluzione $f_3 = \underline{e}_1$.

* * *

4.1.6. Rifare l'Esercizio 2.5.5.

Soluzione. Sia $\mathbf{E} = (\underline{e}_1 \ \underline{e}_2)$ ed $\mathbf{F} = (\underline{f}_1 \ \underline{f}_2)$. Dai dati assegnati:

$$\mathbf{F} = \mathbf{E} C, \text{ con } C = \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix}.$$

Poiché $\det(C) \neq 0$, \mathbf{F} è una base di V . Risulta:

$$\mathbf{E} = \mathbf{F} C^{-1} \text{ e } C^{-1} = \frac{1}{5} \begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix}.$$

Pertanto

$$\underline{e}_1 = \frac{1}{5} (\underline{f}_1 + 2 \underline{f}_2), \quad \underline{e}_2 = \frac{1}{5} (-2 \underline{f}_1 + \underline{f}_2).$$

Si ha:

$$\underline{v} = \mathbf{E} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \mathbf{F} C^{-1} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \mathbf{F} \begin{pmatrix} \frac{3}{5} \\ \frac{1}{5} \end{pmatrix}.$$

Pertanto \underline{v} ha coordinate $(\frac{3}{5}, \frac{1}{5})$ in base \mathbf{F} .

* * *

4.1.7. In $V = V_{\mathbf{R}}^5$, con base \mathbf{E} , è assegnato il sottospazio vettoriale $U = \langle \underline{u}_1, \underline{u}_2, \underline{u}_3 \rangle$, con

$$\underline{u}_1 = \underline{e}_1 - \underline{e}_2 + \underline{e}_5, \quad \underline{u}_2 = 2\underline{e}_1 - \underline{e}_3 + \underline{e}_4, \quad \underline{u}_3 = -2\underline{e}_2 + \underline{e}_3 - \underline{e}_4 + 2\underline{e}_5.$$

Determinare equazioni cartesiane di U , cioè un *SLO* in cinque incognite ed a valori in \mathbf{R} , le cui soluzioni siano tutte e sole le coordinate dei vettori di U , in base \mathbf{E} .

Soluzione. Risulta:

$$(\underline{u}_1 \ \underline{u}_2 \ \underline{u}_3) = \mathbf{E} B, \text{ con } B = \begin{pmatrix} 1 & 2 & 0 \\ -1 & 0 & -2 \\ 0 & -1 & 1 \\ 0 & 1 & -1 \\ 1 & 0 & 2 \end{pmatrix}.$$

Poiché, come si verifica, $B_{(1)}, B_{(2)}$ sono linearmente indipendenti e $B_{(3)} = 2B_{(1)} - B_{(2)}$, allora $\text{rg}(B) = 2$. Pertanto $\dim(U) = 2$ ed una base di U è ad esempio $\{\underline{u}_1, \underline{u}_2\}$.

Sia ora $\underline{v} = \mathbf{E} \mathbf{x} \in V$. Si ha:

$$\begin{aligned} \underline{v} \in U &\iff \underline{u}_1, \underline{u}_2, \underline{v} \text{ sono linearmente dipendenti} \iff \\ &\iff \text{rg}((B_{(1)} \ B_{(2)} \ \mathbf{x})) < 3 \iff \text{rg}((B_{(1)} \ B_{(2)} \ \mathbf{x})) = 2. \end{aligned}$$

Poniamo $A := (B_{(1)} \ B_{(2)} \ \mathbf{x})$. Fissiamo ad esempio la sottomatrice $A(4, 5 | 1, 2)$ ed annulliamone gli orlati. Otteniamo un *SLO*(3, 5, \mathbf{R}) cercato:

$$\left\{ \begin{vmatrix} 1 & 2 & x_1 \\ 0 & 1 & x_4 \\ 1 & 0 & x_5 \end{vmatrix} = \begin{vmatrix} -1 & 0 & x_2 \\ 0 & 1 & x_4 \\ 1 & 0 & x_5 \end{vmatrix} = \begin{vmatrix} 0 & -1 & x_3 \\ 0 & 1 & x_4 \\ 1 & 0 & x_5 \end{vmatrix} = 0, \right.$$

cioè

$$\left\{ \begin{array}{l} x_1 - 2x_4 - x_5 = 0 \\ x_2 + x_5 = 0 \\ x_3 + x_4 = 0. \end{array} \right.$$

* * *

4.2.1. Siano V, W due K -spazi vettoriali di dimensione n , con basi rispettivamente \mathbf{E}, \mathbf{F} . Sia $T : V \rightarrow W$ un'applicazione lineare avente matrice A , rispetto alle basi \mathbf{E}, \mathbf{F} .

Verificare che T è un isomorfismo $\iff A \in \mathbf{GL}_n(K)$.

Soluzione. Poiché $\dim(V) = \dim(W) [= n]$, sappiamo che

$$T \text{ è un isomorfismo} \iff \text{Ker}(T) = \{\underline{0}\}.$$

I vettori di $\text{Ker}(T)$ sono rappresentati dalle soluzioni del $SLO(n, n, K)$ $AX = \mathbf{0}$. Pertanto

$$\text{Ker}(T) = \{\underline{0}\} \iff AX = \mathbf{0} \text{ non ha autosoluzioni} \iff n - rg(A) = 0 \iff A \in \mathbf{GL}_n(K).$$

* * *

4.2.2. Sia \mathbf{E} la base canonica di \mathbf{R}^3 e sia $\mathbf{F} = (\underline{f}_1 \ \underline{f}_2 \ \underline{f}_3) \in \mathfrak{M}_{1,3}(\mathbf{R}^3)$, con

$$\underline{f}_1 = (1, 0, 1), \quad \underline{f}_2 = (0, 1, -2), \quad \underline{f}_3 = (1, 1, 0).$$

Sia $T : \mathbf{R}^3 \rightarrow \mathbf{R}^3$ un operatore lineare tale che

$$T(\underline{f}_1) = (1, 0, 0), \quad T(\underline{f}_2) = (1, 1, 0), \quad T(\underline{f}_3) = (1, 1, 1).$$

Determinare la matrice A di T rispetto alla base canonica \mathbf{E} .

Soluzione. I dati assegnati nell'esercizio sono i seguenti:

$$(1) \quad \mathbf{F} = \mathbf{E} C, \quad \text{con} \quad C = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & -2 & 0 \end{pmatrix} \quad [\text{si noti che } C \in \mathbf{GL}_3(\mathbf{R}) \text{ e quindi } \mathbf{F} \text{ è una base di } \mathbf{R}^3];$$

$$(2) \quad (T(\underline{f}_1) \ T(\underline{f}_2) \ T(\underline{f}_3)) = \mathbf{E} B, \quad \text{con} \quad B = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

Bisogna determinare $A \in \mathfrak{M}_3(\mathbf{R})$ tale che

$$(T(\underline{e}_1) \ T(\underline{e}_2) \ T(\underline{e}_3)) = \mathbf{E} A.$$

Da $\mathbf{F} = \mathbf{E} C$ segue $\mathbf{E} = \mathbf{F} C^{-1}$. Allora

$$T(\underline{e}_i) = T((\mathbf{F} C^{-1})_{(i)}) = (T(\underline{f}_1) \ T(\underline{f}_2) \ T(\underline{f}_3)) C^{-1}_{(i)}.$$

Ne segue:

$$(T(\underline{e}_1) \ T(\underline{e}_2) \ T(\underline{e}_3)) = (T(\underline{f}_1) \ T(\underline{f}_2) \ T(\underline{f}_3)) C^{-1} = (\mathbf{E} B) C^{-1} = \mathbf{E} (B C^{-1}).$$

La matrice richiesta è quindi $A = B C^{-1}$. Risulta:

$$C^{-1} = \begin{pmatrix} 2 & -2 & -1 \\ 1 & -1 & -1 \\ -1 & 2 & 1 \end{pmatrix} \quad \text{e quindi} \quad A = B C^{-1} = \begin{pmatrix} 2 & -1 & -1 \\ 0 & 1 & 0 \\ -1 & 2 & 1 \end{pmatrix}.$$

N.B. Usando la formula di definizione $T(\mathbf{F}) = \mathbf{E} B$ e la formula di cambiamento di base $\mathbf{F} = \mathbf{E} C$, si ha subito:

$$T(\mathbf{E}) = T(\mathbf{F} C^{-1}) = T(\mathbf{F}) C^{-1} = (\mathbf{E} B) C^{-1} = \mathbf{E} (B C^{-1})$$

e dunque $A = B C^{-1}$.

* * *

4.2.3. Siano $S : \mathbf{R}^3 \rightarrow \mathbf{R}^5$ e $T : \mathbf{R}^5 \rightarrow \mathbf{R}^3$ applicazioni lineari così definite:

$$S((a_1, a_2, a_3)) = (0, a_1, a_2, a_3, 0), \quad T((b_1, b_2, b_3, b_4, b_5)) = (b_1, b_3, b_5),$$

$$\forall (a_1, a_2, a_3) \in \mathbf{R}^3, \quad \forall (b_1, b_2, b_3, b_4, b_5) \in \mathbf{R}^5.$$

(i) Scrivere le matrici di $T \circ S$ e di $S \circ T$ rispetto alle basi canoniche \mathbf{E} di \mathbf{R}^3 ed \mathbf{F} di \mathbf{R}^5 .

(ii) Verificare se $\text{Ker}(T \circ S)$ e $\text{Im}(T \circ S)$ sono supplementari in \mathbf{R}^3 e se $\text{Ker}(S \circ T)$ e $\text{Im}(S \circ T)$ lo sono in \mathbf{R}^5 .

Soluzione. (i) Rispetto alle basi canoniche, S e T hanno rispettivamente matrici:

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

Ne segue che l'operatore lineare $T \circ S : \mathbf{R}^3 \rightarrow \mathbf{R}^3$ ha matrice (in base \mathbf{E})

$$BA = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

mentre l'operatore lineare $S \circ T : \mathbf{R}^5 \rightarrow \mathbf{R}^5$ ha matrice (in base \mathbf{F})

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

(ii) I vettori di $\text{Ker}(T \circ S)$ sono le soluzioni del $SLO(3, 3, \mathbf{R})$ $BAX = \mathbf{0}$, che si riduce a $\{x_2 = 0\}$. Una base di $\text{Ker}(T \circ S)$ è data quindi da $\{\underline{e}_1, \underline{e}_3\}$.

$\text{Im}(T \circ S)$ è generato dai vettori le cui coordinate sono le colonne di BA . Dunque $\text{Im}(T \circ S)$ ha base $\{\underline{e}_2\}$. I due sottospazi sono supplementari.

$\text{Ker}(S \circ T)$ è invece descritto dal $SLO(5, 5, \mathbf{R})$ $ABX = \mathbf{0}$, cioè

$$\begin{cases} y_1 = 0 \\ y_3 = 0 \\ y_5 = 0. \end{cases}$$

Una base di $\text{Ker}(S \circ T)$ è data da $\{\underline{f}_2, \underline{f}_4\}$. Invece $\text{Im}(S \circ T)$ ha come base $\{\underline{f}_2, \underline{f}_3, \underline{f}_4\}$. I due sottospazi $\text{Ker}(S \circ T), \text{Im}(S \circ T)$ non sono supplementari [l'intersezione è $\langle \underline{f}_2, \underline{f}_4 \rangle$].

* * *

4.2.4. È assegnata la matrice $A = \begin{pmatrix} 0 & -1 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 2 \end{pmatrix} \in \mathfrak{M}_{3,4}(\mathbf{R})$. Sia $T : \mathbf{R}^4 \rightarrow \mathbf{R}^3$ l'applicazione

lineare definita dalla matrice A , rispetto alle basi canoniche \mathbf{E} di \mathbf{R}^4 ed \mathbf{F} di \mathbf{R}^3 .

Sia poi $S : \mathbf{R}^3 \rightarrow \mathbf{R}^4$ l'applicazione lineare definita (sempre rispetto ad \mathbf{F} ed \mathbf{E}) dalla matrice ${}^t A$.

Determinare la matrice dell'operatore lineare $S \circ T$ di \mathbf{R}^4 (rispetto ad \mathbf{E}) e calcolare dimensioni e basi di $\text{Ker}(S \circ T)$ e $\text{Im}(S \circ T)$.

Soluzione. L'operatore $S \circ T$ ha matrice

$$B = {}^t A A = \begin{pmatrix} 0 & 1 & 1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 2 \end{pmatrix} \begin{pmatrix} 0 & -1 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 2 \end{pmatrix} = \begin{pmatrix} 2 & 1 & 1 & 3 \\ 1 & 2 & -1 & 0 \\ 1 & -1 & 2 & 3 \\ 3 & 0 & 3 & 6 \end{pmatrix}.$$

Calcoliamo $\text{rg}(B)$. Osserviamo che $\text{rg}(A) = 2$ [infatti $A^{(3)} = A^{(1)} + A^{(2)}$]. Dunque

$$\text{rg}(B) = \text{rg}({}^t A A) \leq \min\{\text{rg}({}^t A), \text{rg}(A)\} = \min\{2, 2\} = 2, \text{ cioè } \text{rg}(B) \leq 2.$$

Poiché B contiene ad esempio la sottomatrice quadrata invertibile $B(1, 2 | 1, 2) = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$, allora $\text{rg}(B) = 2$. I vettori di $\text{Ker}(S \circ T)$ sono le soluzioni del $SLO(4, 4, \mathbf{R})$ $BX = \mathbf{0}$. Risulta quindi:

$$\dim(\text{Ker}(S \circ T)) = 4 - \text{rg}(B) = 2.$$

Il SLO $BX = \mathbf{0}$ è equivalente al $SLO(2, 4, \mathbf{R})$ formato dalle sole prime due equazioni, cioè

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

Posto $x_3 = t, x_4 = s$, si ottiene il $SL(2, 2, \mathbf{R})$

$$\begin{cases} 2x_1 + x_2 = -t - 3s \\ x_1 + 2x_2 = t. \end{cases}$$

Con la formula di Cramer, si ottiene

$$x_1 = -t - 2s, \quad x_2 = t + s.$$

Pertanto la generica soluzione del *SLO* è $(-t - 2s, t + s, t, s)$, $\forall t, s \in \mathbf{R}$.

Una base di $\text{Ker}(S \circ T)$ è ad esempio $\{(-1, 1, 1, 0), (-2, 1, 0, 1)\}$.

Dal teorema della nullità più rango, $\dim(\text{Im}(S \circ T)) = 4 - 2 = 2$. Una base di $\text{Im}(S \circ T)$ è ottenuta ad esempio dalle prime due colonne di B (linearmente indipendenti).

Dunque una base di $\text{Im}(S \circ T)$ è $\{(2, 1, 1, 3), (1, 2, -1, 0)\}$.

* * *

4.2.5. Siano $V = V_{\mathbf{R}}^3$ e $W = W_{\mathbf{R}}^4$, con basi rispettivamente \mathbf{E} ed \mathbf{F} . Sia $T : V \rightarrow W$ definito rispetto a tali basi dalla matrice

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \in \mathfrak{M}_{4,3}(\mathbf{R}).$$

Sia W_1 il sottospazio vettoriale di W generato dai vettori $\underline{w}_1, \underline{w}_2$, aventi in base \mathbf{F} rispettivamente coordinate $(-1, 0, 1, 2), (1, 1, 0, 1)$.

Determinare una base e la dimensione di $T^{-1}(W_1)$.

Soluzione. I due vettori $\underline{w}_1, \underline{w}_2$ sono linearmente indipendenti. Infatti la matrice

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

delle loro coordinate (in base \mathbf{F}) ha rango 2.

Si ha:

$\underline{v} \in T^{-1}(W_1) \iff T(\underline{v}) \in W_1 \iff T(\underline{v}) \in \langle \underline{w}_1, \underline{w}_2 \rangle \iff \underline{w}_1, \underline{w}_2, T(\underline{v})$ sono linearmente dipendenti.

Se $\underline{v} = \mathbf{E}\underline{x}$, con $\underline{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$, allora $T(\underline{v})$ in base \mathbf{F} ha coordinate

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ z \\ y \\ y+z \end{pmatrix}.$$

Pertanto

$$\underline{v} \in T^{-1}(W_1) \iff \text{la matrice } \begin{pmatrix} -1 & 1 & x \\ 0 & 1 & z \\ 1 & 0 & y \\ 2 & 1 & y+z \end{pmatrix} \text{ ha rango 2.}$$

Orliamo il minore formato dalla seconda e terza riga di B . Si ottiene che $T^{-1}(W_1)$ è descritto dal *SLO* $(2, 3, \mathbf{R})$

$$\left\{ \begin{vmatrix} -1 & 1 & x \\ 0 & 1 & z \\ 1 & 0 & y \end{vmatrix} = 0, \quad \begin{vmatrix} 0 & 1 & z \\ 1 & 0 & y \\ 2 & 1 & y+z \end{vmatrix} = 0, \right.$$

ovvero

$$\left\{ \begin{array}{l} x+y-z=0 \\ y=0. \end{array} \right.$$

Tale *SLO* ha ∞^1 soluzioni proporzionali a $(1, 0, 1)$. Pertanto $\dim(T^{-1}(W_1)) = 1$ e

$$T^{-1}(W_1) = \langle \underline{e}_1 + \underline{e}_3 \rangle.$$

* * *

4.2.6 Sia $V = V_{\mathbf{R}}^4$ con base \mathbf{E} . Sia $T : V \rightarrow V$ un operatore lineare avente in base \mathbf{E} matrice

$$A = \begin{pmatrix} 1 & 0 & 1 & -1 \\ 1 & 2 & 0 & -1 \\ 0 & -2 & 1 & 0 \\ 2 & 0 & 2 & -2 \end{pmatrix} \in \mathfrak{M}_4(\mathbf{R}).$$

Sia U il sottospazio vettoriale di V generato dai tre vettori $\underline{e}_1 + \underline{e}_4$, $\underline{e}_2 - \underline{e}_3$, $\underline{e}_2 - \underline{e}_4$.

(i) Determinare la dimensione ed una base di $T(U)$.

(ii) Determinare equazioni cartesiane di $T(U)$, cioè un *SLO* in quattro incognite le cui soluzioni sono le coordinate dei vettori di $T(U)$.

Soluzione. (i) $T(U)$ è generato dai tre vettori $T(\underline{e}_1 + \underline{e}_4)$, $T(\underline{e}_2 - \underline{e}_3)$, $T(\underline{e}_2 - \underline{e}_4)$, le cui coordinate in base \mathbf{E} sono rispettivamente

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

La matrice formata dalle coordinate di questi tre vettori è quindi

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

Le ultime due colonne di B sono linearmente indipendenti e quindi $rg(B) = 2$. Ne segue che $dim(T(U)) = 2$ e $T(U)$ ha base

$$-\underline{e}_1 + 2\underline{e}_2 - 3\underline{e}_3 - 2\underline{e}_4, \quad \underline{e}_1 + 3\underline{e}_2 - 2\underline{e}_3 + 2\underline{e}_4.$$

(ii) Per ottenere un *SLO* le cui soluzioni sono le coordinate dei vettori $\underline{v} \in T(U)$, basta imporre la condizione che \underline{v} , $T(\underline{e}_2 - \underline{e}_3)$, $T(\underline{e}_2 - \underline{e}_4)$ siano linearmente dipendenti. Se $\underline{v} = \mathbf{E}\underline{x}$, bisogna imporre che

$$rg((\mathbf{x} \quad B_{(2)} \quad B_{(3)})) = 2, \text{ cioè che } rg \begin{pmatrix} x_1 & -1 & 1 \\ x_2 & 2 & 3 \\ x_3 & -3 & -2 \\ x_4 & -2 & 2 \end{pmatrix} = 2.$$

Orlando la sottomatrice $B(1, 2 | 2, 3)$ si ottiene il $SL(2, 4, \mathbf{R})$

$$\left\{ \begin{vmatrix} x_1 & -1 & 1 \\ x_2 & 2 & 3 \\ x_3 & -3 & -2 \end{vmatrix} = \begin{vmatrix} x_1 & -1 & 1 \\ x_2 & 2 & 3 \\ x_4 & -2 & 2 \end{vmatrix} = 0, \right.$$

cioè

$$\left\{ \begin{array}{l} 5x_1 - 5x_2 - 5x_3 = 0 \\ 10x_1 - 5x_4 = 0, \end{array} \right. \text{ ovvero } \left\{ \begin{array}{l} x_1 - x_2 - x_3 = 0 \\ 2x_1 - x_4 = 0. \end{array} \right.$$

N.B. Si noti che un'altra base di $T(U)$ (con vettori più semplici) è

$$\{\underline{e}_2 - \underline{e}_3, \underline{e}_1 + \underline{e}_2 + 2\underline{e}_4\},$$

ottenuta risolvendo tale *SLO*.

* * *

4.2.7 Sia T un'applicazione lineare da $V = V_{\mathbf{R}}^3$ a $W = W_{\mathbf{R}}^4$ che, espressa rispetto alle basi $\mathbf{E} = (\underline{e}_1 \ \underline{e}_2 \ \underline{e}_3)$ di V ed $\mathbf{F} = (\underline{f}_1 \ \underline{f}_2 \ \underline{f}_3 \ \underline{f}_4)$ di W , è definita da

$$T(\underline{e}_1) = \underline{f}_1 - \underline{f}_2, \quad T(\underline{e}_2) = \underline{f}_2 - \underline{f}_3, \quad T(\underline{e}_3) = \underline{f}_3 - \underline{f}_4.$$

(i) Scrivere la matrice A e le equazioni di T , rispetto alle basi \mathbf{E}, \mathbf{F} . Calcolare poi equazioni cartesiane e basi di $Ker(T)$ e di $Im(T)$.

(ii) Sia $S : W \rightarrow V$ l'applicazione lineare definita dalla matrice $B = {}^t A$, rispetto alle basi \mathbf{F}, \mathbf{E} . Determinare la formula di definizione e le equazioni di S , rispetto alle basi \mathbf{F}, \mathbf{E} . Calcolare poi equazioni cartesiane e basi di $Ker(S)$ e di $Im(S)$.

(iii) Determinare matrice ed equazioni di $T \circ S$. Calcolare poi equazioni cartesiane e basi di $Ker(T \circ S)$ e di $Im(T \circ S)$.

Soluzione. (i) Si ha:

$$T(\mathbf{E}) = \mathbf{F} A, \text{ con } A = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{pmatrix}.$$

T ha equazioni $\mathbf{y} = A\mathbf{x}$, cioè

$$\begin{cases} y_1 = x_1 \\ y_2 = -x_1 + x_2 \\ y_3 = -x_2 + x_3 \\ y_4 = -x_3. \end{cases}$$

Il sottospazio $Ker(T)$ ha equazioni cartesiane date dal $SLO(4, 3, \mathbf{R})$ $A X = \mathbf{0}$, cioè

$$\begin{cases} x_1 = 0 \\ -x_1 + x_2 = 0 \\ -x_2 + x_3 = 0 \\ -x_3 = 0. \end{cases}$$

Poiché, come subito si verifica, $rg(A) = 3$, allora $dim(Ker(T)) = 3 - rg(A) = 3 - 3 = 0$. Dunque $Ker(T) = \langle \emptyset \rangle$, cioè T è iniettiva.

Il sottospazio $Im(T)$ è generato dai tre vettori $T(\underline{e}_1)$, $T(\underline{e}_2)$, $T(\underline{e}_3)$ ed ha dimensione $rg(A) = 3$. Pertanto tali vettori sono una base di $Im(T)$. Per ottenere equazioni cartesiane di $Im(T)$ basta imporre la condizione

$$rg((A \ \mathbf{y})) = 3,$$

che si traduce in

$$\det((A \ \mathbf{y})) = 0.$$

Risulta:

$$\det((A \ \mathbf{y})) = \begin{vmatrix} 1 & 0 & 0 & y_1 \\ -1 & 1 & 0 & y_2 \\ 0 & -1 & 1 & y_3 \\ 0 & 0 & -1 & y_4 \end{vmatrix} = y_1 + y_2 + y_3 + y_4.$$

Dunque $Im(T)$ ha equazione cartesiana $y_1 + y_2 + y_3 + y_4 = 0$.

(ii) Poiché

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

S ha formula di definizione

$$S(\underline{f}_1) = \underline{e}_1, \quad S(\underline{f}_2) = -\underline{e}_1 + \underline{e}_2, \quad S(\underline{f}_3) = -\underline{e}_2 + \underline{e}_3, \quad S(\underline{f}_4) = -\underline{e}_3.$$

Le equazioni di S sono date da $\mathbf{x} = B\mathbf{y}$, cioè

$$\begin{cases} x_1 = y_1 - y_2 \\ x_2 = y_2 - y_3 \\ x_3 = y_3 - y_4. \end{cases}$$

Il sottospazio $Ker(S)$ ha equazioni cartesiane date dal $SLO(3, 4, \mathbf{R})$ $B Y = \mathbf{0}$, cioè

$$\begin{cases} y_1 - y_2 = 0 \\ y_2 - y_3 = 0 \\ y_3 - y_4 = 0. \end{cases}$$

Risulta: $dim(Ker(S)) = 4 - rg(B) = 4 - 3 = 1$. Risolvendo il precedente SLO , si ottiene l'autosoluzione $(1, 1, 1, 1)$. Dunque $Ker(S) = \langle \underline{f}_1 + \underline{f}_2 + \underline{f}_3 + \underline{f}_4 \rangle$.

Il sottospazio $Im(S)$ è generato dai quattro vettori $S(\underline{f}_1)$, $S(\underline{f}_2)$, $S(\underline{f}_3)$, $S(\underline{f}_4)$. Poiché $rg(B) = 3$ e ad esempio le prime tre colonne di B sono linearmente indipendenti, una base di $Im(S)$ è data dai tre vettori $S(\underline{f}_1)$, $S(\underline{f}_2)$, $S(\underline{f}_3)$. Ma una base più semplice di $Im(S)$ si ottiene osservando che $Im(S) = V$ e dunque che una base di $Im(S)$ è data da \underline{e}_1 , \underline{e}_2 , \underline{e}_3 . Si noti infine che $Im(S)$

non ha equazioni cartesiane [ovvero ha equazione cartesiana $0 = 0$, corrispondente alla condizione $\text{rg}((B \ x)) = 3$, identicamente verificata].

(iii) L'applicazione lineare $T \circ S : W \rightarrow W$ ha, rispetto alla base \mathbf{F} di W , matrice AB [infatti $(T \circ S)(\mathbf{F}) = T(S(\mathbf{F})) = T(\mathbf{E}B) = T(\mathbf{E})B = (\mathbf{F}A)B = \mathbf{F}AB$]. Risulta:

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

Le equazioni di $T \circ S$ traducono in base \mathbf{F} l'uguaglianza vettoriale $\underline{w}' = (T \circ S)(\underline{w})$. Posto $\underline{w} = \mathbf{F}\mathbf{y}$ e $\underline{w}' = \mathbf{F}\mathbf{y}'$, le equazioni cercate sono date da $\mathbf{y}' = AB\mathbf{y}$, cioè

$$\begin{cases} y'_1 = y_1 - y_2 \\ y'_2 = -y_1 + 2y_2 - y_3 \\ y'_3 = -y_2 + 2y_3 - y_4 \\ y'_4 = -y_3 + y_4. \end{cases}$$

Il sottospazio $\text{Ker}(T \circ S)$ ha equazioni cartesiane date dal $SLO(4, 4, \mathbf{R})$ $ABY = \mathbf{0}$, cioè

$$\begin{cases} y_1 - y_2 = 0 \\ -y_1 + 2y_2 - y_3 = 0 \\ -y_2 + 2y_3 - y_4 = 0 \\ -y_3 + y_4 = 0. \end{cases}$$

Si verifica subito che $\text{rg}(AB) = 3$. Ne segue che $\dim(\text{Ker}(T \circ S)) = 4 - 3 = 1$. Per ottenere una base di $\text{Ker}(T \circ S)$ è sufficiente determinare un'autosoluzione del precedente SLO . Basta quindi estrarre dalle prime tre righe di AB (linearmente indipendenti) i quattro minori di ordine 3, a segni alterni. Si ottiene l'autosoluzione $(-1, -1, -1, 1)$. Dunque $\text{Ker}(T \circ S) = \langle \underline{f}_1 + \underline{f}_2 + \underline{f}_3 - \underline{f}_4 \rangle$.

Il sottospazio $\text{Im}(T \circ S)$ ha dimensione data da $\text{rg}(AB) = 3$. Ad esempio le prime tre colonne di AB sono linearmente indipendenti. Quindi una base di $\text{Im}(T \circ S)$ è data da

$$\{(T \circ S)(\underline{f}_1), (T \circ S)(\underline{f}_2), (T \circ S)(\underline{f}_3)\}.$$

Un sistema di equazioni cartesiane di $\text{Im}(T \circ S)$ è ottenuta dalla condizione

$$\text{rg}((AB \ \mathbf{y})) = \text{rg}(AB) [= 3].$$

Eliminando la quarta colonna di AB , tale condizione equivale all'annullamento del seguente determinante:

$$\begin{vmatrix} 1 & -1 & 0 & y_1 \\ -1 & 2 & -1 & y_2 \\ 0 & -1 & 2 & y_3 \\ 0 & 0 & -1 & y_4 \end{vmatrix} = y_1 + y_2 + y_3 + y_4.$$

Dunque $\text{Im}(T \circ S)$ ha equazione cartesiana $y_1 + y_2 + y_3 + y_4 = 0$.

N.B. Si può rilevare che $\text{Im}(T)$ e $\text{Im}(T \circ S)$ hanno la stessa equazione e dunque coincidono. Tale fatto non è casuale. Infatti, essendo S un'applicazione suriettiva, si ha: $\text{Im}(T \circ S) = (T \circ S)(W) = T(S(W)) = T(V) = \text{Im}(T)$.

* * *

4.2.8. In \mathbf{R}^3 , con base canonica \mathbf{E} , sono assegnati i due sottospazi vettoriali:

U , rappresentato dal $SLO(1, 3, \mathbf{R})$ $\{x_1 - x_2 = 0\}$;

W , rappresentato dal $SLO(1, 3, \mathbf{R})$ $\{x_2 - x_3 = 0\}$.

Determinare un isomorfismo $T : \mathbf{R}^3 \rightarrow \mathbf{R}^3$ tale che $T(U) = W$. Esprimere T in base \mathbf{E} .

Soluzione. Il problema ha molte soluzioni. Ne otterremo una procedendo come segue. Per prima cosa determineremo una base $\{\underline{u}_1, \underline{u}_2\}$ di U ed una base $\{\underline{w}_1, \underline{w}_2\}$ di W . Poi completeremo tali basi a basi di \mathbf{R}^3 , ottenendo due basi

$$\mathbf{F} = (\underline{u}_1 \ \underline{u}_2 \ \underline{u}_3), \quad \mathbf{G} = (\underline{w}_1 \ \underline{w}_2 \ \underline{w}_3).$$

Definiremo poi "per linearità", $T : \mathbf{R}^3 \rightarrow \mathbf{R}^3$ tale che

$$T(\underline{u}_1) = \underline{w}_1, \quad T(\underline{u}_2) = \underline{w}_2, \quad T(\underline{u}_3) = \underline{w}_3.$$

T è un isomorfismo [in quanto trasforma una base in un'altra base] e risulta, per costruzione, $T(U) = W$. Infine rappresenteremo T rispetto ad \mathbf{E} .

Risolvendo il $SLO(1, 3, \mathbf{R})$ { $x_1 - x_2 = 0$ si ottiene ad esempio la base di U

$$\{\underline{u}_1 = (1, 1, 0), \underline{u}_2 = (0, 0, 1)\}.$$

Un ulteriore vettore $\underline{u} \notin \langle \underline{u}_1, \underline{u}_2 \rangle$ è ad esempio $\underline{u} = (1, 0, 0)$. Dunque

$$\mathbf{F} = (\underline{u}_1 \ \underline{u}_2 \ \underline{u}) = \mathbf{E} C, \text{ con } C = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \in \mathbf{GL}_3(\mathbf{R}).$$

Risolviamo ora il $SLO(1, 3, \mathbf{R})$ { $x_2 - x_3 = 0$. Si ottiene ad esempio la base di W

$$\{\underline{w}_1 = (1, 0, 0), \underline{w}_2 = (0, 1, 1)\}.$$

Un ulteriore vettore $\underline{w} \notin \langle \underline{w}_1, \underline{w}_2 \rangle$ è ad esempio $\underline{w} = (0, 0, 1)$. Dunque

$$\mathbf{G} = (\underline{w}_1 \ \underline{w}_2 \ \underline{w}) = \mathbf{E} D, \text{ con } D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \in \mathbf{GL}_3(\mathbf{R}).$$

Si definisce allora T tale che

$$(T(\underline{u}_1) \ T(\underline{u}_2) \ T(\underline{u})) = \mathbf{G} = \mathbf{E} D.$$

Invertendo la formula $\mathbf{F} = \mathbf{E} C$ si ottiene $\mathbf{E} = \mathbf{F} C^{-1}$, cioè:

$$\underline{e}_1 = \underline{u}, \quad \underline{e}_2 = \underline{u}_1 - \underline{u}, \quad \underline{e}_3 = \underline{u}_2.$$

Quindi

$$T(\underline{e}_1) = T(\underline{u}) = \mathbf{E} D_{(3)}, \quad T(\underline{e}_2) = T(\underline{u}_1 - \underline{u}) = \mathbf{E} (D_{(1)} - D_{(3)}), \quad T(\underline{e}_3) = T(\underline{u}_2) = \mathbf{E} D_{(2)}.$$

Pertanto T in base \mathbf{E} ha matrice:

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

* * *

4.2.9. Sia $V = (\mathbf{Z}_2)^2 [= \mathbf{Z}_2 \times \mathbf{Z}_2]$.

(i) Determinare la cardinalità dell'anello $\mathcal{E}nd(V) [= \mathcal{E}nd_{\mathbf{Z}_2}(V)]$ degli operatori lineari di V .

(ii) Determinare il gruppo $\mathcal{U}(\mathcal{E}nd(V))$ degli elementi invertibili di $\mathcal{E}nd(V)$.

Soluzione. (i) Lo spazio vettoriale V è formato dai quattro vettori

$$(\bar{0}, \bar{0}), (\bar{0}, \bar{1}), (\bar{1}, \bar{0}), (\bar{1}, \bar{1}),$$

con $\bar{0}, \bar{1} \in \mathbf{Z}_2$. Dunque $V \times V$ ha cardinalità 16.

Si fissi ora in V la base canonica $\{\underline{e}_1 = (\bar{1}, \bar{0}), \underline{e}_2 = (\bar{0}, \bar{1})\}$. È noto che $\mathcal{E}nd(V)$ è in corrispondenza biunivoca con $V \times V$, tramite l'applicazione

$$\Phi : \mathcal{E}nd(V) \rightarrow V \times V \text{ tale che } \Phi(T) = (T(\underline{e}_1), T(\underline{e}_2)), \quad \forall T \in \mathcal{E}nd(V).$$

Segue che $|\mathcal{E}nd(V)| = 16$.

(ii) Gli elementi di $\mathcal{U}(\mathcal{E}nd(V))$ sono tutti e soli gli operatori lineari invertibili, cioè gli automorfismi di V . Formano un gruppo, usualmente denotato $\mathcal{A}ut(V)$.

Fissata la base canonica di V [sia nello spazio vettoriale di partenza che in quello di arrivo], gli automorfismi di V sono tutti e soli gli operatori lineari individuati dalle matrici invertibili di $\mathfrak{M}_2(\mathbf{Z}_2)$, cioè dalle matrici di $\mathbf{GL}_2(\mathbf{Z}_2)$. Si osserva facilmente che

$$\mathbf{GL}_2(\mathbf{Z}_2) = \left\{ \begin{pmatrix} \bar{1} & \bar{0} \\ \bar{0} & \bar{1} \end{pmatrix}, \begin{pmatrix} \bar{0} & \bar{1} \\ \bar{1} & \bar{0} \end{pmatrix}, \begin{pmatrix} \bar{0} & \bar{1} \\ \bar{1} & \bar{1} \end{pmatrix}, \begin{pmatrix} \bar{1} & \bar{0} \\ \bar{1} & \bar{1} \end{pmatrix}, \begin{pmatrix} \bar{1} & \bar{1} \\ \bar{0} & \bar{1} \end{pmatrix}, \begin{pmatrix} \bar{1} & \bar{1} \\ \bar{1} & \bar{0} \end{pmatrix} \right\}$$

[mentre le altre dieci matrici di $\mathfrak{M}_2(\mathbf{Z}_2)$ hanno determinante nullo].

Le due colonne di tali sei matrici forniscono le immagini rispettivamente dei vettori $\underline{e}_1, \underline{e}_2$, nella stessa base. Ad esempio, la seconda matrice definisce l'automorfismo

$$\varphi_2 : \begin{cases} \underline{e}_1 \rightarrow \underline{e}_2 \\ \underline{e}_2 \rightarrow \underline{e}_1, \end{cases}$$

mentre l'ultima matrice definisce l'automorfismo

$$\varphi_6 : \begin{cases} \underline{e}_1 \rightarrow \underline{e}_1 + \underline{e}_2 \\ \underline{e}_2 \rightarrow \underline{e}_1. \end{cases}$$

N.B. Si noti che $(\varphi_3)^3 = (\varphi_6)^3 = \mathbf{1}_V$, mentre $(\varphi_2)^2 = (\varphi_4)^2 = (\varphi_5)^2 = \mathbf{1}_V$. Si potrebbe verificare che $\mathcal{U}(\text{End}(V)) \cong S_3$.

* * *

4.2.10. Siano \mathbf{E} ed \mathbf{F} due basi di $V = V_K^n$, tali che $\mathbf{F} = \mathbf{E}C$, con $C \in \mathbf{GL}_n(K)$. Sia T un operatore lineare di V definito da $T(\mathbf{E}) = \mathbf{F}A$. Determinare la matrice di T rispetto:

- alla base \mathbf{E} sia nello spazio vettoriale di partenza che in quello di arrivo,
- alla base \mathbf{F} sia nello spazio vettoriale di partenza che in quello di arrivo,
- alla base \mathbf{F} nello spazio vettoriale di partenza ed alla base \mathbf{E} in quello di arrivo.

Soluzione. Le matrici richieste sono rispettivamente le matrici $B_1, B_2, B_3 \in \mathfrak{M}_3(K)$ tali che:

$$T(\mathbf{E}) = \mathbf{E}B_1, \quad T(\mathbf{F}) = \mathbf{F}B_2, \quad T(\mathbf{F}) = \mathbf{E}B_3.$$

Si ha:

$$T(\mathbf{E}) = \mathbf{F}A = (\mathbf{E}C)A = \mathbf{E}(CA) \text{ e quindi } B_1 = CA;$$

$$T(\mathbf{F}) = T(\mathbf{E}C) = T(\mathbf{E})C = (\mathbf{F}A)C = \mathbf{F}(AC) \text{ e quindi } B_2 = AC;$$

$$T(\mathbf{F}) = T(\mathbf{E}C) = T(\mathbf{E})C = (\mathbf{F}A)C = \mathbf{F}(AC) = (\mathbf{E}C)(AC) = \mathbf{E}(CAC) \text{ e quindi } B_3 = CAC.$$

* * *

4.2.11. Sono assegnati i due spazi vettoriali

$$V = V_K^n, \text{ con basi } \mathbf{E} \text{ ed } \mathbf{E}' \text{ tali che } \mathbf{E}' = \mathbf{E}C;$$

$$W = W_K^m, \text{ con basi } \mathbf{F} \text{ ed } \mathbf{F}' \text{ tali che } \mathbf{F}' = \mathbf{F}D.$$

Sia $T : V \rightarrow W$ un'applicazione lineare avente matrice $A \in \mathfrak{M}_{m,n}(K)$, rispetto alle basi \mathbf{E} ed \mathbf{F} . Qual'è la matrice di T rispetto alle basi \mathbf{E}' ed \mathbf{F}' ?

Soluzione. La formula di definizione assegnata è $T(\mathbf{E}) = \mathbf{F}A$ e quella cercata è $T(\mathbf{E}') = \mathbf{F}'A'$, con $A' \in \mathfrak{M}_{m,n}(K)$ da determinare. Si ha (tenuto conto che $\mathbf{F} = \mathbf{F}'D^{-1}$):

$$T(\mathbf{E}') = T(\mathbf{E}C) = T(\mathbf{E})C = \mathbf{F}AC = (\mathbf{F}'D^{-1})AC = \mathbf{F}'(D^{-1}AC).$$

Pertanto la matrice richiesta è $A' = D^{-1}AC$.

* * *

4.3.1. Sia $T : V_K^n \rightarrow V_K^n$ un operatore lineare.

(i) Verificare che, se λ è un autovalore di T , λ^2 è un autovalore di T^2 .

(ii) È vero che, se T è diagonalizzabile, anche T^2 lo è?

(iii) È vero che, se T^2 è diagonalizzabile, anche T lo è?

Soluzione. (i) Sia \underline{v} un autovettore di T associato a λ [$T(\underline{v}) = \lambda\underline{v}$]. Si ha:

$$T^2(\underline{v}) = T(T(\underline{v})) = T(\lambda\underline{v}) = \lambda T(\underline{v}) = \lambda^2\underline{v}.$$

Ne segue che \underline{v} è un autovettore di T^2 associato a λ^2 . Pertanto λ^2 è un autovalore di T^2 .

(ii) Sia \mathbf{F} una base di autovettori di T . In base \mathbf{F} T ha matrice diagonale D . Nella stessa base \mathbf{F} , T^2 ha matrice D^2 , anch'essa diagonale. Pertanto T^2 è diagonalizzabile.

(iii) La risposta è negativa. Lo proviamo con il seguente esempio.

Sia $T : \mathbf{R}^2 \rightarrow \mathbf{R}^2$ l'operatore lineare definito (rispetto alla base canonica \mathbf{E} di \mathbf{R}^2) dalla matrice

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

L'operatore T non è diagonalizzabile: infatti ha il solo autovalore 0 con molteplicità geometrica 1 [in quanto $\mathbf{E}_0(T) = \langle \underline{e}_1 \rangle$].

Si ha invece che T^2 è diagonalizzabile, in quanto (in base \mathbf{E}) ha matrice

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

[e dunque $T^2 = \mathbf{0}$, operatore nullo (che è diagonalizzabile)].

* * *

4.3.2. Sia $V = V_K^n$ e T un operatore lineare di V , tale che

$$T \neq \mathbf{1}_V, T \neq \mathbf{0}_V \text{ e } T^2 = T.$$

(i) Verificare che 0, 1 sono autovalori di T .

(ii) Verificare che gli autospazi $\mathbf{E}_0(T)$, $\mathbf{E}_1(T)$ sono sottospazi vettoriali supplementari di V .

Soluzione. (i) Poiché $T \neq \mathbf{1}_V$, esiste $\underline{v} \in V$ tale che $T(\underline{v}) \neq \underline{v}$. Allora $\underline{v} - T(\underline{v}) \neq \underline{0}$ e si ha:

$$T(\underline{v} - T(\underline{v})) = T(\underline{v}) - T^2(\underline{v}) = T(\underline{v}) - T(\underline{v}) = \underline{0}.$$

Pertanto $\underline{v} - T(\underline{v}) \in \text{Ker}(T)$ e quindi $0 \in \Lambda(T)$ e $\underline{v} - T(\underline{v}) \in \mathbf{E}_0(T)$.

Poiché $T \neq \mathbf{0}_V$, esiste $\underline{v} \in V$ tale che $T(\underline{v}) \neq \underline{0}$. Allora

$$T(T(\underline{v})) = T^2(\underline{v}) = T(\underline{v}) = 1 \cdot T(\underline{v}).$$

Pertanto $1 \in \Lambda(T)$ e $T(\underline{v}) \in \mathbf{E}_1(T)$.

(ii) Ovviamente $\mathbf{E}_0(T) \cap \mathbf{E}_1(T) = \{\underline{0}\}$. Basta quindi verificare che $\mathbf{E}_0(T) + \mathbf{E}_1(T) = V$. Si ha infatti, $\forall \underline{v} \in V$:

$$\underline{v} = \underline{v} - T(\underline{v}) + T(\underline{v}) = (\underline{v} - T(\underline{v})) + T(\underline{v})$$

e $\underline{v} - T(\underline{v}) \in \mathbf{E}_0(T)$, mentre $T(\underline{v}) \in \mathbf{E}_1(T)$.

N.B. Si noti che per ottenere un operatore lineare verificante le ipotesi dell'esercizio basta determinare una matrice quadrata $A \neq \mathbf{0}$, I_n tale che $A^2 = A$. Ad esempio la matrice

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \in \mathfrak{M}_3(\mathbf{R}).$$

* * *

4.3.3. Al variare di $a, b \in \mathbf{R}$, sia $T : \mathbf{R}^3 \rightarrow \mathbf{R}^3$ l'operatore lineare avente matrice

$$A = \begin{pmatrix} 0 & a & b \\ -a & 0 & -b \\ -b & b & 0 \end{pmatrix}$$

[rispetto alla base canonica \mathbf{E} di \mathbf{R}^3]. Per quali $a, b \in \mathbf{R}$ l'operatore T è diagonalizzabile?

Soluzione. T ha polinomio caratteristico

$$P_T = |A - x I_3| = \begin{vmatrix} -x & a & b \\ -a & -x & -b \\ -b & b & -x \end{vmatrix} = -x(x^2 + (2b^2 + a^2)).$$

Se $a = b = 0$, la matrice A è nulla ed il polinomio caratteristico è $-x^3$. In tal caso $T = \mathbf{0}$ è diagonalizzabile.

Se invece $(a, b) \neq (0, 0)$, il fattore $x^2 + (2b^2 + a^2)$ di P_T è irriducibile. Quindi $\lambda = 0$ è l'unico autovalore di T ed ha molteplicità geometrica 1. Pertanto T non è diagonalizzabile.

* * *

4.3.4. Sia $T : \mathbf{R}^3 \rightarrow \mathbf{R}^3$ l'operatore lineare definito, rispetto alla base canonica \mathbf{E} , dalla matrice

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

dipendente da un parametro reale a . Determinare per quali valori di a T è diagonalizzabile.

Soluzione. T ha polinomio caratteristico

$$P_T = \begin{vmatrix} 1-\lambda & 0 & -\frac{1}{2} \\ -2 & 2-\lambda & 1 \\ a & 0 & 3-\lambda \end{vmatrix} = (2-\lambda)(\lambda^2 - 4\lambda + 3 + \frac{a}{2}).$$

Un autovalore di T è 2. Gli altri autovalori sono gli eventuali zeri reali del polinomio $\lambda^2 - 4\lambda + 3 + \frac{a}{2}$. Tale polinomio ammette zeri reali $\iff 1 - \frac{a}{2} \geq 0 \iff a \leq 2$. Distinguiamo tre casi:

$$a > 2, \quad a < 2, \quad a = 2.$$

Se $a > 2$, T ha soltanto l'autovalore 2. L'autospazio $\mathbf{E}_2(T)$ ha equazioni cartesiane date dal *SLO* $(A - 2I_3)X = \mathbf{0}$. Si verifica subito che la matrice $A - 2I_3$ ha rango 2. Pertanto la molteplicità geometrica dell'autovalore 2 è $d_2 = 3 - 2 = 1$. Ne segue che T non è diagonalizzabile.

N.B. Si osservi che 2 ha molteplicità algebrica $h_2 = 1$. Dunque necessariamente $d_2 = 1$.

Se $a < 2$, T ha tre autovalori distinti: $2 - \sqrt{1 - \frac{a}{2}}, 2, 2 + \sqrt{1 - \frac{a}{2}}$. In tal caso T è diagonalizzabile.

Se $a = 2$, T ha polinomio caratteristico $P_T = (2-\lambda)(\lambda^2 - 4\lambda + 4) = -(\lambda-2)^3$ e quindi T ha soltanto l'autovalore 2. L'autospazio $\mathbf{E}_2(T)$ ha equazioni cartesiane date dal *SLO* $(A - 2I_3)X = \mathbf{0}$. Poiché la matrice

$$A - 2I_3 = \begin{pmatrix} -1 & 0 & -\frac{1}{2} \\ -2 & 0 & 1 \\ 2 & 0 & 1 \end{pmatrix}$$

ha rango 2, allora $d_2 = 3 - 2 = 1$. Segue che T non è diagonalizzabile.

* * *

4.3.5. Sia $T : \mathbf{R}^4 \rightarrow \mathbf{R}^4$ l'operatore lineare definito, rispetto alla base canonica \mathbf{E} , dalla matrice

$$A = \begin{pmatrix} -5 & 0 & 4 & 4 \\ -10 & 5 & 6 & -4 \\ 0 & 0 & 3 & 8 \\ 0 & 0 & 2 & -3 \end{pmatrix}.$$

Determinare gli autovalori di T e, se esiste, una base di autovettori di T .

Soluzione. Il polinomio caratteristico è

$$\begin{aligned} P = |A - xI_4| &= \begin{vmatrix} -5-x & 0 & 4 & 4 \\ -10 & 5-x & 6 & -4 \\ 0 & 0 & 3-x & 8 \\ 0 & 0 & 2 & -3-x \end{vmatrix} = \begin{vmatrix} -5-x & 0 & 0 \\ -10 & 5-x & 0 \\ 0 & 0 & -3-x \end{vmatrix} \cdot \begin{vmatrix} 3-x & 8 \\ 2 & -3-x \end{vmatrix} = \\ &= -(25-x^2)(-(9-x^2)-16) = (25-x^2)^2 = (x-5)^2(x+5)^2. \end{aligned}$$

Pertanto

$$\Lambda(T) = \{5, -5\}.$$

Inoltre ognuno dei due autovalori ha molteplicità algebrica 2. Pertanto l'operatore T è diagonalizzabile \iff entrambi gli autovalori hanno molteplicità geometrica 2. Calcoliamo quindi i due autospazi.

Per ottenere una base di $\mathbf{E}_5(T)$ basta risolvere il *SLO* $(4, 4, \mathbf{R})$ avente matrice

$$B := A - 5I_4 = \begin{pmatrix} -10 & 0 & 4 & 4 \\ -10 & 0 & 6 & -4 \\ 0 & 0 & -2 & 8 \\ 0 & 0 & 2 & -8 \end{pmatrix}.$$

Si verifica subito che $rg(B) = 2$ [infatti $B^{(3)} = B^{(1)} - B^{(2)}$ e $B^{(4)} = -B^{(3)}$]. Ne segue che il *SLO* ha ∞^2 soluzioni. Per ottenerle riduciamo il *SLO* dato al *SLO* $(2, 4, \mathbf{R})$ (equivalente) formato dalle sole prime due righe di B , cioè

$$\begin{cases} -10x_1 + 4x_3 + 4x_4 = 0 \\ -10x_1 + 6x_3 - 4x_4 = 0. \end{cases}$$

Posto $x_1 = t$, $x_2 = s$, si ottiene: $x_3 = 2t$, $x_4 = \frac{1}{2}t$. Una coppia di autovettori (indipendenti) associati a 5 è ad esempio

$$\underline{v}_1 = (0, 1, 0, 0), \quad \underline{v}_2 = (2, 0, 4, 1).$$

Per ottenere invece una base di $\mathbf{E}_{-5}(T)$ basta risolvere il *SLO* $(4, 4, \mathbf{R})$ avente matrice

$$C := A + 5 I_4 = \begin{pmatrix} 0 & 0 & 4 & 4 \\ -10 & 10 & 6 & -4 \\ 0 & 0 & 8 & 8 \\ 0 & 0 & 2 & 2 \end{pmatrix}.$$

Anche C ha rango 2 [infatti $C^{(4)} = C^{(3)}$, $C^{(1)} = \frac{1}{2}C^{(3)}$]. Due righe di C linearmente indipendenti sono $C^{(2)}, C^{(3)}$ ed il SLO dato si riduce al $SLO(2, 4, \mathbf{R})$ da esse formato, cioè

$$\begin{cases} -10x_1 + 10x_2 + 6x_3 - 4x_4 = 0 \\ 8x_3 + 8x_4 = 0. \end{cases}$$

Posto $x_1 = t$, $x_2 = s$, si ottiene: $x_3 = t - s$, $x_4 = -t + s$. Una coppia di autovettori (indipendenti) associati a -5 è ad esempio

$$\underline{v}_3 = (1, 0, 1, -1), \quad \underline{v}_4 = (0, 1, -1, 1).$$

In base $\mathbf{F} = (\underline{v}_1 \ \underline{v}_2 \ \underline{v}_3 \ \underline{v}_4)$ l'operatore T ha matrice diagonale

$$D = \begin{pmatrix} 5 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 0 & 0 & -5 & 0 \\ 0 & 0 & 0 & -5 \end{pmatrix}.$$

* * *

4.3.6. È assegnata l'applicazione lineare $T : \mathbf{R}^3 \rightarrow \mathbf{R}^2$ definita, rispetto alle basi canoniche dei due spazi vettoriali, dalla matrice

$$A = \begin{pmatrix} 0 & a & 1 \\ 0 & 1 & a \end{pmatrix}, \text{ dipendente da un parametro } a \in \mathbf{R}.$$

Sia $S : \mathbf{R}^2 \rightarrow \mathbf{R}^3$ l'applicazione lineare definita dalla matrice ${}^t A$ (sempre rispetto alle basi canoniche). Determinare per quali eventuali $a \in \mathbf{R}$ l'operatore lineare $S \circ T$ non è diagonalizzabile.

Soluzione. L'operatore lineare $S \circ T : \mathbf{R}^3 \rightarrow \mathbf{R}^3$ ha matrice (rispetto alla base canonica \mathbf{E} di \mathbf{R}^3)

$${}^t A A = \begin{pmatrix} 0 & 0 \\ a & 1 \\ 1 & a \end{pmatrix} \begin{pmatrix} 0 & a & 1 \\ 0 & 1 & a \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & a^2 + 1 & 2a \\ 0 & 2a & a^2 + 1 \end{pmatrix}.$$

Il polinomio caratteristico di $S \circ T$ è

$$P = \begin{vmatrix} -x & 0 & 0 \\ 0 & a^2 + 1 - x & 2a \\ 0 & 2a & a^2 + 1 - x \end{vmatrix} = -x [(a^2 + 1 - x)^2 - 4a^2] = \\ = -x [(a^2 + 1 - x - 2a)(a^2 + 1 - x + 2a)] = -x ((a-1)^2 - x)((a+1)^2 - x).$$

Pertanto

$$\Lambda(S \circ T) = \{0, (a-1)^2, (a+1)^2\}.$$

Se i tre autovalori $0, (a-1)^2, (a+1)^2$ sono a due a due distinti, l'operatore è ovviamente diagonalizzabile. Si ha:

i tre autovalori non sono a due a due distinti \iff vale una delle tre condizioni $\begin{cases} (a-1)^2 = (a+1)^2, \\ (a-1)^2 = 0, \\ (a+1)^2 = 0. \end{cases}$

Tali condizioni equivalgono nell'ordine a:

$$a = 0, a = 1, a = -1.$$

Esaminiamo allora la diagonalizzabilità di $S \circ T$, per $a = 0, 1, -1$.

(i) Sia $a = 0$. In tal caso $S \circ T$ ha matrice

$$A_0 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

La matrice è diagonale e dunque $S \circ T$ è diagonalizzabile [anzi è già diagonalizzato, rispetto ad \mathbf{E}].

(ii) Sia $a = 1$. In tal caso $S \circ T$ ha matrice

$$A_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 2 & 2 \\ 0 & 2 & 2 \end{pmatrix}$$

e $\Lambda(S \circ T) = \{0, 0, 4\}$. Per decidere se $S \circ T$ è diagonalizzabile occorre verificare se la dimensione dell'autospazio $\mathbf{E}_0(S \circ T)$ è 2. Basta risolvere il $SLO(3, 3, \mathbf{R})$ $A_1 X = \mathbf{0}$. Tale SLO è equivalente al $SLO(1, 3, \mathbf{R})$ $\{y + z = 0\}$, che ha ∞^2 soluzioni, generate ad esempio da $(1, 0, 0), (0, 1, -1)$. Inoltre $\mathbf{E}_4(S \circ T) = \langle (0, 1, 1) \rangle$. Dunque una base di autovettori di $S \circ T$ è

$$(0, 1, 1), (1, 0, 0), (0, 1, -1).$$

(iii) Sia $a = -1$. In tal caso $S \circ T$ ha matrice

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

e $\Lambda(S \circ T) = \{0, 4, 0\}$.

Calcoliamo ancora l'autospazio $\mathbf{E}_0(S \circ T)$, risolvendo il $SLO(3, 3, \mathbf{R})$ $A_{-1} X = \mathbf{0}$. Tale SLO è equivalente al $SLO(1, 3, \mathbf{R})$ $\{y - z = 0\}$, che ha ∞^2 soluzioni, generate da $(1, 0, 0), (0, 1, 1)$. Inoltre $\mathbf{E}_4(S \circ T) = \langle (0, -1, 1) \rangle$. Dunque una base di autovettori di $S \circ T$ è

$$(0, -1, 1), (1, 0, 0), (0, 1, 1).$$

Abbiamo provato che l'operatore lineare $S \circ T$ è diagonalizzabile, $\forall a \in \mathbf{R}$.

* * *

4.3.7. Sia $T : \mathbf{R}^2 \rightarrow \mathbf{R}^2$ l'operatore lineare definito dai seguenti dati:

$$\underline{v} = (1, 2), \quad T(\underline{v}) = (2, 1), \quad T^2(\underline{v}) = (-2, -1).$$

(i) Determinare la matrice A di T rispetto alla base canonica \mathbf{E} di \mathbf{R}^2 .

(ii) Determinare gli autovalori di T e, se esiste, una base di autovettori di T .

Soluzione. (i) Dai dati dell'esercizio si osserva che è assegnata (in base \mathbf{E}) l'immagine di T dei due vettori $\underline{v}, T(\underline{v})$. Se quindi $\mathbf{F} = (\underline{v} \ T(\underline{v}))$ è una base di \mathbf{R}^2 , T è completamente individuata, rispetto alle basi \mathbf{F}, \mathbf{E} . Si ha:

$$\mathbf{F} = (\underline{v} \ T(\underline{v})) = \mathbf{E} C, \text{ con } C = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}.$$

Essendo $C \in \mathbf{GL}_2(\mathbf{R})$, \mathbf{F} è una base. Si ha:

$$T(\mathbf{F}) = (T(\underline{f}_1) \ T(\underline{f}_2)) = (T(\underline{v}) \ T(T(\underline{v}))) = \mathbf{E} B, \text{ con } B = \begin{pmatrix} 2 & -2 \\ 1 & -1 \end{pmatrix}.$$

Ora determiniamo la matrice A tale che

$$T(\mathbf{E}) = \mathbf{E} A.$$

Risulta:

$$T(\mathbf{E}) = T(\mathbf{F} C^{-1}) = T(\mathbf{F}) C^{-1} = (\mathbf{E} B) C^{-1} = \mathbf{E}(B C^{-1})$$

e dunque la matrice cercata è $A = B C^{-1}$. Poiché

$$C^{-1} = \begin{pmatrix} -\frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{1}{3} \end{pmatrix},$$

risulta:

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

(ii) Il polinomio caratteristico di T è:

$$P_T = \begin{vmatrix} -2 - x & 2 \\ -1 & 1 - x \end{vmatrix} = x^2 + x = x(x + 1).$$

Dunque $\Lambda(T) = \{0, -1\}$. T ha due autospazi 1-dimensionali ed è quindi diagonalizzabile. Calcoliamo i due autospazi. $\mathbf{E}_0(T) = Ker(T)$ è ottenuto risolvendo il SLO

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

Si ottiene l'autosoluzione $(1, 1)$. Dunque $\mathbf{E}_0(T) = \langle (1, 1) \rangle$. $\mathbf{E}_{-1}(T)$ è ottenuto risolvendo il SLO

$$\begin{cases} -x + 2y = 0 \\ -x + 2y = 0. \end{cases}$$

Si ottiene l'autosoluzione $(2, 1)$. Dunque $\mathbf{E}_{-1}(T) = \langle (2, 1) \rangle = \langle T(\underline{v}) \rangle$.

N.B. Che T possedesse l'autovalore -1 con autovettore associato $T(\underline{v})$ era evidente dai dati.

* * *

4.3.8. Sia $T : \mathbf{R}^3 \rightarrow \mathbf{R}^3$ l'operatore lineare individuato dai seguenti dati:

$$\begin{aligned} T \text{ è diagonalizzabile, } \Lambda(T) &= \{0, 1\}, \quad \mathbf{E}_1(T) = \langle (2, 0, 1) \rangle, \\ T(W) &\subseteq W, \text{ con } W = \langle (1, 1, 1), (0, 1, -1) \rangle. \end{aligned}$$

Determinare la matrice di T rispetto alla base canonica \mathbf{E} di \mathbf{R}^3 .

Soluzione. I tre vettori assegnati (cioè i generatori di $\mathbf{E}_1(T)$ e di W) formano un base \mathbf{F} di \mathbf{R}^3 . Infatti

$$\mathbf{F} = \mathbf{E}C, \text{ con } C = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & -1 \end{pmatrix},$$

e $C \in \mathbf{GL}_3(\mathbf{R})$.

Osserviamo ora che, in base alle due ipotesi $\underline{f}_1 \in \mathbf{E}_1(T)$, $T(W) \subseteq W$, T ha, in base \mathbf{F} , matrice del tipo

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & a & c \\ 0 & b & d \end{pmatrix}, \text{ con } a, b, c, d \in \mathbf{R}.$$

Poiché T è diagonalizzabile e $\Lambda(T) = \{0, 1\}$, allora $d_0 + d_1 = 3$; ma $d_1 = 1$ e quindi $d_0 = 2$.

L'autospazio $\mathbf{E}_0(T) [= \text{Ker}(T)]$ ha perciò dimensione 2 ed i vettori di tale autospazio sono ottenuti risolvendo il SLO $AX = \mathbf{0}$. Dunque $3 - rg(A) = 2$, cioè $rg(A) = 1$. In A sono quindi nulli tutti i minori di ordine 2. Quindi in particolare

$$\begin{vmatrix} 1 & 0 \\ 0 & a \end{vmatrix} = \begin{vmatrix} 1 & 0 \\ 0 & c \end{vmatrix} = \begin{vmatrix} 1 & 0 \\ 0 & b \end{vmatrix} = \begin{vmatrix} 1 & 0 \\ 0 & d \end{vmatrix} = 0, \text{ cioè } a = b = c = d = 0.$$

Si conclude che la matrice A di T in base \mathbf{F} è

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

Ora possiamo calcolare la matrice B di T in base canonica \mathbf{E} . Risulta:

$$T(\mathbf{E}) = T(\mathbf{F}C^{-1}) = T(\mathbf{F})C^{-1} = (\mathbf{F}A)C^{-1} = \mathbf{F}(AC^{-1}) = \mathbf{E}C(AC^{-1}) = \mathbf{E}(CAC^{-1}).$$

Dunque $B = CAC^{-1}$. Risulta:

$$C^{-1} = \frac{1}{3} \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & 2 \\ 1 & 1 & -2 \end{pmatrix}.$$

Pertanto

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

* * *

4.3.9. In $V = V_{\mathbf{R}}^3$ sono assegnate due basi $\mathbf{E} = (\underline{e}_1 \ \underline{e}_2 \ \underline{e}_3)$, $\mathbf{F} = (\underline{f}_1 \ \underline{f}_2 \ \underline{f}_3)$ tali che:

$$\underline{e}_1 = \underline{f}_2, \quad \underline{e}_2 = -\underline{f}_1 + \underline{f}_3, \quad \underline{e}_3 = \underline{f}_1 + \underline{f}_2.$$

Sia $T : V \rightarrow V$ l'operatore lineare tale che

$$T(\underline{f}_1) = \underline{e}_1 - \underline{e}_2, \quad T(\underline{f}_2) = \underline{e}_2 - \underline{e}_3, \quad T(\underline{f}_3) = \underline{e}_3 - \underline{e}_1.$$

(i) Esprimere T in base \mathbf{E} ed in base \mathbf{F} .

(ii) Verificare se T è diagonalizzabile e calcolare una base di ogni autospazio.

(iii) Sia U il sottospazio vettoriale di V definito in base \mathbf{F} dal $SLO(1, 3, \mathbf{R})$ $\{y_1 - y_2 = 0\}$. Determinare una base di $T(U)$ (in base \mathbf{F}).

Soluzione. (i) Si ha:

$$\mathbf{E} = \mathbf{F}C, \text{ con } C = \begin{pmatrix} 0 & -1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

[C è ovviamente invertibile (infatti $\det(C) = 1$)]. Inoltre

$$(T(\underline{f}_1) \ T(\underline{f}_2) \ T(\underline{f}_3)) = \mathbf{E}D, \text{ con } D = \begin{pmatrix} 1 & 0 & -1 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix}.$$

Allora

$$(T(\underline{f}_1) \ T(\underline{f}_2) \ T(\underline{f}_3)) = \mathbf{F}CD, \text{ con } CD = \begin{pmatrix} 1 & -2 & 1 \\ 1 & -1 & 0 \\ -1 & 1 & 0 \end{pmatrix}.$$

In base \mathbf{F} , T ha quindi matrice $B = CD$.

Per rappresentare T in base \mathbf{E} , osserviamo che

$$\begin{cases} T(\underline{e}_1) = T(\underline{f}_2) = \underline{e}_2 - \underline{e}_3 \\ T(\underline{e}_2) = T(-\underline{f}_1 + \underline{f}_3) = -T(\underline{f}_1) + T(\underline{f}_3) = -2\underline{e}_1 + \underline{e}_2 + \underline{e}_3 \\ T(\underline{e}_3) = T(\underline{f}_1 + \underline{f}_2) = T(\underline{f}_1) + T(\underline{f}_2) = \underline{e}_1 - \underline{e}_3. \end{cases}$$

Pertanto T in base \mathbf{E} ha matrice

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

(ii) Per calcolare il polinomio caratteristico di T possiamo usare indifferentemente la matrice B o la matrice A . Useremo la matrice B (operando quindi in base \mathbf{F}). Si ha:

$$P_T = \begin{vmatrix} 1-x & -2 & 1 \\ 1 & -1-x & 0 \\ -1 & 1 & -x \end{vmatrix} = x(1-x^2) + 1 - (1+x) - 2x = -x^3 - 2x = -x(x^2 + 2).$$

Poiché $x^2 + 2$ è un polinomio irriducibile in \mathbf{R} , T non è diagonalizzabile. Inoltre $\Lambda(T) = \{0\}$. L'unico autospazio $\mathbf{E}_0(T) = Ker(T)$ è ottenuto risolvendo il SLO $BY = 0$, cioè

$$\begin{cases} y_1 - 2y_2 + y_3 = 0 \\ y_1 - y_2 = 0 \\ -y_1 + y_2 = 0, \end{cases} \text{ cioè } \begin{cases} y_1 - y_2 = 0 \\ y_1 - y_3 = 0. \end{cases}$$

Tale SLO ha autosoluzione $(1, 1, 1)$. Dunque

$$\mathbf{E}_0(T) = \langle \underline{f}_1 + \underline{f}_2 + \underline{f}_3 \rangle.$$

(iii) Il $SLO(1, 3, \mathbf{R})$ $\{y_1 - y_2 = 0\}$ ha ∞^2 soluzioni. Una base di tali soluzioni è $\{(1, 1, 0), (0, 0, 1)\}$. Dunque

$$U = \langle \underline{f}_1 + \underline{f}_2, \underline{f}_3 \rangle.$$

Risulta:

$$T(\underline{f}_1 + \underline{f}_2) = \mathbf{F}B \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \mathbf{F} \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix} = -\underline{f}_1, \quad T(\underline{f}_3) = \underline{f}_1.$$

Allora

$$T(U) = \langle T(\underline{f}_1 + \underline{f}_2), T(\underline{f}_3) \rangle = \langle -\underline{f}_1, \underline{f}_1 \rangle = \langle \underline{f}_1 \rangle.$$

Una base di $T(U)$ è dunque $\{\underline{f}_1\}$.

* * *

4.3.10. Sia $T : \mathbf{R}^3 \rightarrow \mathbf{R}^3$ l'operatore lineare definito dai seguenti dati: posto $\underline{x} = (1, 0, -1) \in \mathbf{R}^3$,

$$T(\underline{x}) = (0, 1, 2), \quad T^2(\underline{x}) = (1, 1, 0), \quad T^3(\underline{x}) = (0, 1, 2).$$

- (i) Determinare la matrice A di T rispetto alla base canonica \mathbf{E} di \mathbf{R}^3 .
- (ii) Verificare che T è diagonalizzabile e indicarne gli autovalori.
- (iii) Assegnato il sottospazio $W = \langle T(\underline{x}), T^2(\underline{x}) \rangle$, determinare una base di $T^{-1}(W)$.

Soluzione. (i) È assegnato il comportamento di T sui tre vettori $\underline{x}, T(\underline{x}), T^2(\underline{x})$. Verifichiamo che tali vettori formano una base. Sia $\mathbf{F} = (\underline{x} \ T(\underline{x}) \ T^2(\underline{x}))$. Risulta:

$$\mathbf{F} = \mathbf{E} C, \text{ con } C = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ -1 & 2 & 0 \end{pmatrix}.$$

Poiché $\det(C) = -1$, \mathbf{F} è una base di \mathbf{R}^3 .

Risulta, in base ai dati assegnati:

$$T(\mathbf{F}) = (T(\underline{x}) \ T^2(\underline{x}) \ T^3(\underline{x})) = \mathbf{E} D, \text{ con } D = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 2 & 0 & 2 \end{pmatrix}.$$

Per esprimere T nella base canonica \mathbf{E} , calcoliamo \mathbf{E} in funzione di \mathbf{F} . Risulta: $\mathbf{E} = \mathbf{F} C^{-1}$, con

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

Si ha:

$$T(\mathbf{E}) = T(\mathbf{F} C^{-1}) = T(\mathbf{F}) C^{-1} = (\mathbf{E} D) C^{-1} = \mathbf{E} (D C^{-1}).$$

La matrice A di T in base \mathbf{E} è quindi

$$A = D C^{-1} = \begin{pmatrix} 1 & -1 & 1 \\ 2 & -1 & 1 \\ 2 & 0 & 0 \end{pmatrix}.$$

(ii) Risulta:

$$P_T = \begin{vmatrix} 1-x & -1 & 1 \\ 2 & -1-x & 1 \\ 2 & 0 & -x \end{vmatrix} = x(1-x^2) - 2 + 2(1+x) - 2x = x(1-x^2).$$

Pertanto $\Lambda(T) = \{0, 1, -1\}$. Poiché T ha tre autovalori distinti (tutti di molteplicità algebrica 1) T è diagonalizzabile.

(iii) Poniamo

$$(T(\underline{x}) \ T^2(\underline{x})) = \mathbf{E} H, \text{ con } H = \begin{pmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 0 \end{pmatrix}.$$

Ovviamente $\text{rg}(H) = 2$. Sia $T(\underline{v}) = \mathbf{E} A \underline{x}$, con $\underline{v} = \mathbf{E} \underline{x}$, generico vettore di \mathbf{R}^3 . Si ha:

$$\underline{v} \in T^{-1}(W) \iff T(\underline{v}) \in W \iff \text{rg}((A \underline{x} \ H)) = \text{rg}(H) \iff \text{rg}((A \underline{x} \ H)) = 2.$$

Si ha:

$$\text{rg}((A \underline{x} \ H)) = 2 \iff \begin{vmatrix} x_1 - x_2 + x_3 & 0 & 1 \\ 2x_1 - x_2 + x_3 & 1 & 1 \\ 2x_1 & 2 & 0 \end{vmatrix} = 0.$$

Tale determinante è identicamente nullo. Ciò significa che la condizione $\text{rg}((A \underline{x} \ H)) = 2$ è sempre verificata, $\forall \underline{v} \in \mathbf{R}^3$, ovvero che $T^{-1}(W) = \mathbf{R}^3$. Una base di $T^{-1}(W)$ è quindi \mathbf{E} .

N.B. I calcoli fatti in (iii) sono inutili. In effetti si sarebbe potuto osservare che $W = \text{Im}(T)$ [infatti $\text{Im}(T) = \langle T(\underline{x}), T^2(\underline{x}), T^3(\underline{x}) \rangle$ e $T^3(\underline{x}) = T(\underline{x})$. Dunque $\text{Im}(T) = \langle T(\underline{x}), T^2(\underline{x}) \rangle$].

Ogni volta che un sottospazio W contiene $\text{Im}(T)$, la sua controimmagine è tutto lo spazio vettoriale di partenza (in questo caso \mathbf{R}^3).

* * *

4.3.11. Sia $V = V_K^n$ e sia $T : V \rightarrow V$ l'operatore lineare definito, rispetto ad una base \mathbf{E} di V , dalla matrice $A \in \mathfrak{M}_n(K)$. Sia $T' : V \rightarrow V$ l'operatore lineare definito, sempre rispetto ad \mathbf{E} , dalla matrice ${}^t A$ (trasposta di A).

(i) Verificare che $P_T = P_{T'}$.

(ii) Verificare che T è diagonalizzabile $\iff T'$ è diagonalizzabile.

Soluzione. (i) Si osservi che

$${}^t A - \lambda I_n = {}^t A - \lambda {}^t I_n = {}^t (A - \lambda I_n).$$

Poiché il determinante di una matrice coincide con quello della sua trasposta, allora

$$P_{T'} = |{}^t A - \lambda I_n| = |{}^t (A - \lambda I_n)| = |A - \lambda I_n| = P_T.$$

(ii)) Da (i)) segue che $\Lambda(T) = \Lambda(T')$. Basta allora verificare che, $\forall \lambda \in \Lambda(T)$,

$$d_\lambda(T) = d_\lambda(T')$$

[cioè che le molteplicità geometriche di ogni autovalore coincidono]. Si ha:

$$d_\lambda(T) = \dim(\mathbf{E}_\lambda(T)) = n - rg(A - \lambda I_n),$$

$$d_\lambda(T') = \dim(\mathbf{E}_\lambda(T')) = n - rg({}^t A - \lambda I_n).$$

Poiché ${}^t A - \lambda I_n = {}^t (A - \lambda I_n)$ ed il rango di una matrice coincide con quello della sua trasposta, allora

$$rg({}^t A - \lambda I_n) = rg(A - \lambda I_n),$$

da cui $d_\lambda(T) = d_\lambda(T')$.

* * *

4.3.12. Determinare, in una base opportuna, gli operatori lineari $T : \mathbf{R}^3 \rightarrow \mathbf{R}^3$ definiti dalle seguenti tre condizioni:

$$\mathbf{E}_1(T) = \langle (1, 0, -2) \rangle, \quad \mathbf{E}_0(T) = \langle (1, 1, 1) \rangle, \quad h_0 = 2$$

[dove h_0 denota la molteplicità algebrica di 0 come autovalore di T].

Soluzione. Poiché $h_0 + h_1 \leq 3$ e $h_0 = 2$, $h_1 \geq 1$, allora necessariamente $h_0 + h_1 = 3$ e $h_1 = 1$. Dunque T non ha altri autovalori (oltre a 0, 1) e non è diagonalizzabile [in quanto $d_0 = 1 < h_0 = 2$].

Consideriamo i due autovettori

$$\underline{v}_1 = (1, 0, -2), \quad \underline{v}_2 = (1, 1, 1)$$

e completiamoli ad una base di \mathbf{R}^3 con un terzo vettore. Indicata con \mathbf{E} la base canonica di \mathbf{R}^3 , possiamo ad esempio scegliere come terzo vettore \underline{e}_3 . Infatti

$$\mathbf{F} = (\underline{v}_1 \quad \underline{v}_2 \quad \underline{e}_3) = \mathbf{E} C \quad \text{e} \quad C = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ -2 & 1 & 1 \end{pmatrix} \in \mathbf{GL}_3(\mathbf{R}).$$

In base \mathbf{F} , T ha matrice del tipo

$$A = \begin{pmatrix} 1 & 0 & a \\ 0 & 0 & b \\ 0 & 0 & c \end{pmatrix}, \quad \text{con } a, b, c \in \mathbf{R},$$

[dove ovviamente $T(\underline{e}_3) = a\underline{v}_1 + b\underline{v}_2 + c\underline{e}_3$].

Ora verifichiamo che $c = 0$. Abbiamo due strade differenti da poter seguire.

(1) Poiché $h_0 = 2$, $h_1 = 1$, T ha polinomio caratteristico

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

T ha quindi traccia 1 [in quanto è il coefficiente di x^2]. Poiché la traccia è un invariante di T e $Tr(A) = 1 + c$, segue che $c = 0$.

(2) Poiché A è triangolare superiore, gli autovalori di T (con la relativa molteplicità algebrica) si leggono sulla diagonale di A . Pertanto sulla diagonale devono comparire 1, 0, 0. Dunque $c = 0$.

Pertanto in base \mathbf{F} T ha matrice

$$A = \begin{pmatrix} 1 & 0 & a \\ 0 & 0 & b \\ 0 & 0 & 0 \end{pmatrix}, \quad \text{con } a, b \in \mathbf{R}.$$

Poiché infine $\dim(Ker(T)) = \dim(\mathbf{E}_0(T)) = 1$ e $\dim(Ker(T)) = 3 - rg(A)$, allora $rg(A) = 2$. Osserviamo che ciò avviene $\iff b \neq 0$.

Concludiamo quindi che, in base \mathbf{F} , T ha una delle seguenti matrici

$$A = \begin{pmatrix} 1 & 0 & a \\ 0 & 0 & b \\ 0 & 0 & 0 \end{pmatrix}, \quad \forall a, b \in \mathbf{R}, \quad b \neq 0.$$

* * *

4.3.13. Sia T un operatore lineare di $V = V_K$, tale che $T^3 \neq \mathbf{0}$ e $T^4 = \mathbf{0}$ (operatore nullo).

- (i) Verificare che $\Lambda(T) = \{0\}$.
- (ii) Verificare che T non è diagonalizzabile.
- (iii) Determinare un esempio di un siffatto operatore T di \mathbf{R}^4 .

Soluzione. (i) Proveremo prima che, se $\lambda \in \Lambda(T)$, allora $\lambda = 0$. Poi proveremo che effettivamente $0 \in \Lambda(T)$.

Sia $\lambda \in \Lambda(T)$ e sia \underline{u} un autovettore associato. Si ha:

$$\underline{0} = T^4(\underline{u}) = T^3(T(\underline{u})) = T^3(\lambda\underline{u}) = \lambda(T^3(\underline{u})) = \dots = \lambda^3 T(\underline{u}) = \lambda^4 \underline{u}.$$

Da $\lambda^4 \underline{u} = \underline{0}$ segue che $\lambda^4 = 0$ e quindi $\lambda = 0$.

Si consideri ora un vettore $\underline{v} \in V$ tale che $T^3(\underline{v}) \neq \underline{0}$. Ovviamente $T(T^3(\underline{v})) = T^4(\underline{v}) = \underline{0}$. Dunque $T^3(\underline{v}) \in \text{Ker}(T) = \mathbf{E}_0(T)$ e pertanto $T^3(\underline{v})$ è un autovettore associato all'autovalore 0. Quindi $\Lambda(T) = \{0\}$.

(ii) Poiché $T^3 \neq \mathbf{0}$, necessariamente $T \neq \mathbf{0}$. Allora $\text{Ker}(T) \neq V$. Poiché $\text{Ker}(T)$ è l'unico autospazio di T , si conclude che non esiste una base di autovettori di T .

(iii) Si tratta di determinare una matrice $A \in \mathfrak{M}_4(\mathbf{R})$ tale che $A^4 = \mathbf{0}$ ed $A^3 \neq \mathbf{0}$ (matrice nulla). Consideriamo ad esempio la seguente matrice triangolare superiore A (con diagonale tutta nulla):

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

Si ha:

$$A^2 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad A^3 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad A^4 = \mathbf{0}.$$

Dunque l'operatore T di \mathbf{R}^4 avente (in base canonica) matrice A verifica le condizioni richieste ($T^3 \neq \mathbf{0}$, $T^4 = \mathbf{0}$).

* * *

4.3.14. È assegnata la matrice

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

Verificare che esiste un polinomio non nullo $M = M(x) = a_0 + a_1x + a_2x^2 + a_3x^3$, a coefficienti in \mathbf{R} , tale che

$$\sum_{i=0}^3 a_i A^i = \mathbf{0}$$

[dove A^i denota la i -sima potenza di A e $\mathbf{0}$ denota la matrice quadrata nulla di ordine 3]. Confrontare il polinomio M con il polinomio caratteristico P_A .

Soluzione. Calcoliamo A^2 e A^3 . Si ha:

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

Pertanto:

$$a_0 I_3 + a_1 A + a_2 A^2 + a_3 A^3 = \begin{pmatrix} a_0 + a_1 - a_2 - a_3 & -2a_1 + 2a_3 & 0 \\ a_1 - a_3 & a_0 - a_1 - a_2 + a_3 & 0 \\ 0 & 0 & a_0 \end{pmatrix}.$$

Uguagliando tale matrice con la matrice $\mathbf{0} \in \mathfrak{M}_3(\mathbf{R})$ si ottiene il $SLO(9, 4, \mathbf{R})$

$$\begin{cases} a_0 + a_1 - a_2 - a_3 = 0, & -2a_1 + 2a_3 = 0, & 0 = 0 \\ a_1 - a_3 = 0, & a_0 - a_1 - a_2 + a_3 = 0, & 0 = 0 \\ 0 = 0, & 0 = 0, & a_0 = 0, \end{cases}$$

che è equivalente al $SLO(4, 4, \mathbf{R})$

$$\begin{cases} a_0 = 0 \\ a_1 - a_3 = 0 \\ a_0 + a_1 - a_2 - a_3 = 0 \\ a_0 - a_1 - a_2 + a_3 = 0, \end{cases} \text{ cioè } \begin{cases} a_0 = 0 \\ a_1 = a_3 \\ a_1 - a_2 - a_3 = 0 \\ -a_1 - a_2 + a_3 = 0, \end{cases} \text{ ovvero } \begin{cases} a_0 = 0 \\ a_2 = 0 \\ a_1 = a_3. \end{cases}$$

Tale SLO ha ∞^1 soluzioni ed un'autosoluzione è $(0, 1, 0, 1)$. Dunque il polinomio M cercato è [a meno di un fattore di proporzionalità non nullo]:

$$M = M(x) = x + x^3.$$

Si ha quindi $M(A) = \mathbf{0}$ [infatti $A + A^3 = \mathbf{0}$, come si controlla subito da dati sopra ottenuti].

Il polinomio caratteristico P_A è

$$P_A = \begin{vmatrix} 1-x & -2 & 0 \\ 1 & -1-x & 0 \\ 0 & 0 & -x \end{vmatrix} = -x(x^2 + 1) = -x^3 - x$$

e quindi coincide con M [a meno di un fattore di proporzionalità non nullo]. Dunque $P_A(A) = \mathbf{0}$.

N.B. Tale fatto non è certamente casuale: per ogni matrice $A \in \mathfrak{M}_n(K)$, risulta: $P_A(A) = \mathbf{0}$. Si tratta di un importante risultato: il *teorema di Hamilton-Cayley*.

* * *

ESERCIZI PROPOSTI

Capitolo 5

5.1.1. Determinare il periodo dell'elemento x^{320} del gruppo ciclico $C_{15} = \langle x \mid x^{15} = 1 \rangle$. Indicare tutti i generatori del sottogruppo $\langle x^{320} \rangle$.

Soluzione. Dividiamo 320 per 15. Si ha: $320 = 15 \cdot 21 + 5$ e quindi:

$$x^{320} = (x^{15})^{21} x^5 = 1^{21} x^5 = x^5.$$

Allora

$$\circ(x^{320}) = \circ(x^5) = \frac{15}{MCD(15, 5)} = \frac{15}{5} = 3.$$

Il sottogruppo ciclico $\langle x^{320} \rangle$ è quindi:

$$\langle x^{320} \rangle = \langle x^5 \rangle = \{1, x^5, x^{10}\}.$$

Tale gruppo ha ordine 3 ed ha quindi due generatori: x^5, x^{10} .

* * *

5.1.2. Determinare il reticolo dei sottogruppi di C_{20} .

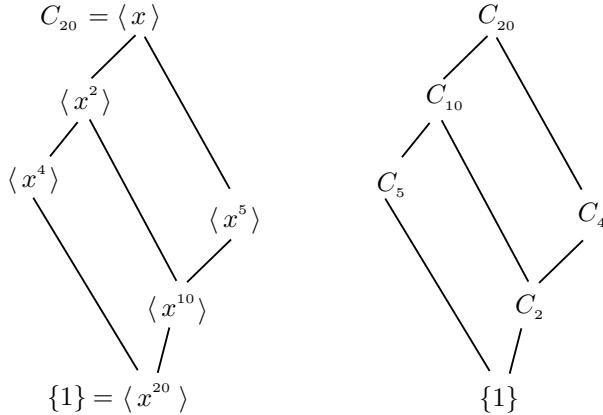
Soluzione. Sia $C_{20} = \langle x \mid x^{20} = 1 \rangle$. I divisori positivi di 20 sono i seguenti naturali $d = 1, 2, 4, 5, 10, 20$. Esiste in C_{20} un unico sottogruppo di ordine d , cioè

$$\langle x^{20/d} \rangle \cong C_d.$$

Si ha quindi:

$$\begin{aligned} \langle x^{20/1} \rangle &= \langle x^{20} \rangle = \{1\}, & \langle x^{20/2} \rangle &= \langle x^{10} \rangle \cong C_2, & \langle x^{20/4} \rangle &= \langle x^5 \rangle \cong C_4, & \langle x^{20/5} \rangle &= \langle x^4 \rangle \cong C_5, \\ \langle x^{20/10} \rangle &= \langle x^2 \rangle \cong C_{10}, & \langle x^{20/20} \rangle &= \langle x \rangle \cong C_{20}. \end{aligned}$$

Il reticolo dei sottogruppi di C_{20} è quindi il seguente:



* * *

5.1.3. Determinare i sottogruppi ciclici di S_4 .

Soluzione. S_4 possiede sei 4-cicli. Ogni 4-ciclo genera un gruppo ciclico di ordine 4, al cui interno ci sono due 4-cicli (inversi tra loro). Dunque S_4 possiede tre sottogruppi C_4 :

$$\begin{aligned} \langle (1 2 3 4) \rangle &= \{(1), (1 2 3 4), (1 3)(2 4), (1 4 3 2)\} = \langle (1 4 3 2) \rangle, \\ \langle (1 2 4 3) \rangle &= \{(1), (1 2 4 3), (1 4)(2 3), (1 3 4 2)\} = \langle (1 3 4 2) \rangle, \\ \langle (1 3 2 4) \rangle &= \{(1), (1 3 2 4), (1 2)(3 4), (1 4 2 3)\} = \langle (1 4 2 3) \rangle. \end{aligned}$$

Inoltre S_4 possiede otto 3-cicli. Ogni 3-ciclo genera un gruppo ciclico di ordine 3, al cui interno ci sono due 3-cicli (inversi tra loro). Pertanto S_4 possiede quattro sottogruppi C_3 :

$$\begin{aligned}\langle (123) \rangle &= \{(1), (123), (132)\} = \langle (132) \rangle, \\ \langle (124) \rangle &= \{(1), (124), (142)\} = \langle (142) \rangle, \\ \langle (134) \rangle &= \{(1), (134), (143)\} = \langle (143) \rangle, \\ \langle (234) \rangle &= \{(1), (234), (243)\} = \langle (243) \rangle.\end{aligned}$$

Infine S_4 possiede sei 2-cicli e tre coppie di 2-cicli disgiunti. Ogni gruppo ciclico di ordine 2 possiede un solo elemento di periodo 2. Dunque S_4 possiede nove sottogruppi C_2 :

$$\begin{aligned}\langle (12) \rangle, \quad \langle (13) \rangle, \quad \langle (14) \rangle, \quad \langle (23) \rangle, \quad \langle (24) \rangle, \quad \langle (34) \rangle, \\ \langle (12)(34) \rangle, \quad \langle (13)(24) \rangle, \quad \langle (14)(23) \rangle.\end{aligned}$$

* * *

5.1.4. Verificare che $U(\mathbf{Z}_9)$ è un gruppo ciclico e determinarne tutti i generatori.

Soluzione. Risulta:

$$U(\mathbf{Z}_9) = \{\bar{1}, \bar{2}, \bar{4}, \bar{5}, \bar{7}, \bar{8}\}.$$

Calcoliamo il periodo di $\bar{2}$. Si ha:

$$\bar{2}^2 = \bar{4}, \quad \bar{2}^3 = \bar{8}, \quad \bar{2}^4 = \bar{16} = \bar{7}, \quad \bar{2}^5 = \bar{14} = \bar{5}, \quad \bar{2}^6 = \bar{10} = \bar{1}.$$

Dunque $\circ(\bar{2}) = 6$ e pertanto $U(\mathbf{Z}_9)$ è ciclico.

Si noti che in ogni C_6 , se x è un generatore, l'altro generatore è x^5 . Dunque l'altro generatore di $U(\mathbf{Z}_9)$ è $\bar{2}^5 = \bar{5}$.

* * *

5.1.5. Verificare che un gruppo di ordine 4 non possiede elementi di periodo 3. Dedurne che un gruppo di ordine 4 o è ciclico o è un gruppo di Klein [cioè che, a meno di isomorfismi, esistono due soli gruppi di ordine 4: C_4 e V].

Soluzione. [Si noti che la prima affermazione da dimostrare è ovvia, se si assume acquisito il teorema di Lagrange. Ma a questo punto il teorema di lagrange non lo conosciamo ancora].

Sia (G, \cdot) un gruppo di ordine 4 ed assumiamo, per assurdo, che G ammetta un elemento x di periodo 3.

Allora $1, x, x^2 \in G$. Sia y l'ulteriore elemento di G . Si osserva subito che $\circ(y) \neq 4$ [altrimenti G sarebbe ciclico ed un gruppo C_4 non ha elementi di periodo 3]. Inoltre $\circ(y) \neq 3$ [altrimenti anche $y^2 \in G$ e quindi $|G| \geq 5$]. Necessariamente allora $\circ(y) = 2$. Poiché $xy \in G$, allora xy coincide con uno degli elementi $1, x, x^2, y$. Ma $xy \neq 1$ (essendo $y \neq x^2$); $xy \neq x$ (essendo $y \neq 1$); $xy \neq x^2$ (essendo $y \neq x$); $xy \neq x$ (essendo $y \neq 1$). Da ciò l'assurdo.

In base a quanto provato sopra, un gruppo non ciclico G di ordine 4 possiede necessariamente tre elementi di periodo 2. Sia quindi

$$G = \{1, a, b, c\}, \quad \text{con } \circ(a) = \circ(b) = \circ(c) = 2.$$

Si osserva subito che $ab \neq 1, a, b$; analogamente $ba \neq 1, a, b$. Pertanto necessariamente

$$ab = c = ba.$$

Si conclude che

$$G = \langle a, b \mid a^2 = b^2 = 1, ba = ab \rangle \cong V \quad (\text{gruppo di Klein}).$$

* * *

5.1.6. Sia G un gruppo di ordine 6.

(i) Verificare che G non può possedere cinque elementi di periodo 2.

(ii) Verificare che G non può possedere tre elementi di periodo 3.

Soluzione. (i) Supponiamo che esista un gruppo G di ordine 6, formato (oltreché dall'unità 1) da cinque elementi x_1, \dots, x_5 di periodo 2. Verifichiamo che un tale gruppo è abeliano. Infatti, $\forall i \neq j$, $x_i x_j \in \{x_1, \dots, x_5\}$ e quindi ha periodo 2. Pertanto $(x_i x_j)(x_i x_j) = 1$. D'altra parte

$$(x_i x_j)(x_j x_i) = x_i (x_j x_i) x_i = x_i 1 x_i = 1.$$

Dunque $x_i x_j = x_j x_i$.

Ora calcoliamo $x_1 x_2$. In base alla legge di cancellazione, $x_1 x_2 \neq 1$, x_1, x_2 . Allora (eventualmente rinumerando x_3, x_4, x_5) risulta:

$$x_1 x_2 = x_3.$$

Segue (moltiplicando a sinistra per x_1) $x_1 x_3 = x_2$ e quindi (moltiplicando a destra per x_3) $x_2 x_3 = x_1$.

Ora calcoliamo $x_1 x_4$. Ovviamente $x_1 x_4 \neq 1$, x_1, x_4 . Inoltre si ha:

$$x_1 x_4 \neq x_2 [= x_1 x_3] \text{ e } x_1 x_4 \neq x_3 [= x_1 x_2].$$

Quindi necessariamente $x_1 x_4 = x_5$. Ne segue subito che $x_1 x_5 = x_4$.

Infine calcoliamo $x_2 x_4$. Si ha: $x_2 x_4 \neq 1$, x_2, x_4 . Inoltre

$$x_2 x_4 \neq x_1 [= x_2 x_3], \quad x_2 x_4 \neq x_3 [= x_1 x_2 = x_2 x_1] \text{ e } x_2 x_4 \neq x_5 [= x_1 x_4].$$

Si è così ottenuto un assurdo.

(ii) Osserviamo che in ogni gruppo finito G il numero degli elementi di periodo 3 è sempre pari. Se infatti $x \in G$ e $\circ(x) = 3$, allora $x^2 \neq x$ e $\circ(x^2) = 3$. Assumiamo per assurdo che G possieda tre elementi di periodo 3. Allora ne possiede almeno quattro.

Se $|G| = 6$, allora G contiene $1, x, x^2, y, y^2$, con x, x^2, y, y^2 di periodo 3 e distinti tra loro. Poiché $xy \in G$ e $xy \neq 1, x, x^2, y, y^2$, allora xy è il sesto elemento di G . Ma anche $x^2y \in G$ e anche $x^2y \neq 1, x, x^2, y, y^2$. Allora $x^2y = xy$ e quindi $x = 1$: assurdo.

* * *

5.1.7. Sono assegnati due simboli x, y , legati soltanto dalle seguenti tre relazioni (moltiplicative):

$$x^4 = 1, \quad y^2 = 1, \quad yx = x^3y.$$

Verificare che gli elementi generati da tali simboli sono otto e scriverne la tavola moltiplicativa. Verificare che formano un gruppo [che è chiamato *gruppo diedrale di ordine 8*].

Soluzione. (i) G contiene gli elementi $1, x, x^2, x^3, y$. Inoltre contiene i prodotti xy, x^2y, x^3y (tutti diversi tra loro). Si noti che i prodotti $y x^k$ (per $k = 1, 2, 3$) si identificano rispettivamente con gli elementi $x^{4-k}y$ (in base alla relazione $yx = x^3y$). Ad esempio si ha:

$$yx^2 = (yx)x = x^3yx = x^3x^3y = x^6y = x^2y.$$

Per verificare che questi otto elementi formano un gruppo basta scriverne la tavola moltiplicativa e verificare che su ciascuna riga e colonna compaiono (una sola volta) tutti gli elementi.

	1	x	x^2	x^3	y	xy	x^2y	x^3y
1	1	x	x^2	x^3	y	xy	x^2y	x^3y
x	x	x^2	x^3	1	xy	x^2y	x^3y	y
x^2	x^2	x^3	1	x	x^2y	x^3y	y	xy
x^3	x^3	1	x	x^2	x^3y	y	xy	x^2y
y	y	x^3y	x^2y	xy	1	x^3	x^2	x
xy	xy	y	x^3y	x^2y	x	1	x^3	x^2
x^2y	x^2y	xy	y	x^3y	x^2	x	1	x_3
x^3y	x^3y	x^2y	xy	y	x^3	x^2	x	1

* * *

5.1.8. (i) Sia G un gruppo abeliano. Verificare che l'insieme H degli elementi di periodo finito di G è un sottogruppo di G .

(ii) Se invece G non è abeliano, H può non essere un sottogruppo di G . Per dimostrare tale affermazione si utilizzino i seguenti dati:

$$G = \mathbf{GL}_2(\mathbf{R}), \quad A = \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix} \in G.$$

Verificare che $A, {}^t A \in H$ mentre il prodotto $A {}^t A \notin H$.

Soluzione. (i) Sia (G, \circ) abeliano e siano $x, y \in H$, con $\circ(x) = t, \circ(y) = s$. Allora

$$(xy)^{st} = x^{st} y^{st} = (x^t)^s (y^s)^t = 1^s 1^t = 1.$$

Dunque $\circ(xy) \leq st < \infty$, cioè $xy \in H$. Ovviamente $1 \in H$ e se poi $x \in H$ (e $\circ(x) = t$), anche $x^{-1} \in H$ [infatti $(x^{-1})^t = (x^t)^{-1} = 1^{-1} = 1$].

(ii) Si ha: $A^2 = I_2 = ({}^t A)^2$. Dunque $\circ(A) = \circ({}^t A) = 2$. Sia ora $B = A {}^t A$. Si ha:

$$B = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}, \quad B^2 = \begin{pmatrix} 5 & -3 \\ -3 & 2 \end{pmatrix}, \quad B^3 = \begin{pmatrix} 13 & -8 \\ -8 & 5 \end{pmatrix}, \dots.$$

Si noti che $(B^k)_{1,1} < (B^{k+1})_{1,1}, \forall k \geq 1$. Dunque $B^k \neq I_2, \forall k \geq 1$, e quindi $\circ(B) = \infty$. Ne segue che H non è chiuso rispetto al prodotto e quindi non è un gruppo.

* * *

5.2.1. Determinare le relazioni di equivalenza su G associate alla partizione dei laterali destri ed a quella dei laterali sinistri modulo un sottogruppo H di G .

Soluzione. È noto dall'**Osserv. 1** di **Cap. 1.3** che ad ogni partizione resta associata una relazione di equivalenza, che mette in relazione due elementi se e solo se si trovano in uno stesso insieme della partizione. Nel caso della partizione $\mathfrak{L}_d(H)$, la relazione associata $\rho_{H,d}$ o, più semplicemente, ρ_d è quindi così ottenuta: $\forall a, b \in G$,

$$a\rho_d b \iff Ha = Hb \iff ab^{-1} \in H.$$

Pertanto ρ_d è così definita:

$$a\rho_d b \iff ab^{-1} \in H, \quad \forall a, b \in G.$$

La relazione di equivalenza associata alla partizione $\mathfrak{L}_s(H)$ è invece la seguente:

$$a\rho_s b \iff aH = bH \iff a^{-1}b \in H.$$

L'insieme quoziante G/ρ_d coincide quindi con $\mathfrak{L}_d(H)$, mentre G/ρ_s coincide con $\mathfrak{L}_s(H)$.

* * *

5.2.2. Determinare i sottogruppi ciclici di $\mathbf{U}(\mathbf{Z}_{15})$ e dire se tale gruppo è ciclico.

Soluzione. (i) Si tenga conto che $\varphi(15) = \varphi(3)\varphi(5) = 2 \cdot 4 = 8$. Risulta:

$$\mathbf{U}(\mathbf{Z}_{15}) = \{\bar{1}, \bar{2}, \bar{4}, \bar{7}, \bar{8}, \bar{11}, \bar{13}, \bar{14}\}.$$

Poiché $|\mathbf{U}(\mathbf{Z}_{15})| = 8$, gli elementi di $\mathbf{U}(\mathbf{Z}_{15})$ possono avere periodo 1, 2, 4, 8. Calcoliamo tali periodi. Si ha:

- $\circ(\bar{2}) = 4$ [infatti $\bar{2}^2 = \bar{4}, \bar{2}^4 = \bar{4}^2 = \bar{1}$];
- $\circ(\bar{4}) = 2$ [infatti $\bar{4}^2 = \bar{1}$];
- $\circ(\bar{7}) = 4$ [infatti $\bar{7}^2 = \bar{49} = \bar{4}, \bar{7}^4 = \bar{4}^2 = \bar{1}$];
- $\circ(\bar{8}) = 4$ [infatti $\bar{8}^2 = \bar{64} = \bar{4}, \bar{8}^4 = \bar{4}^2 = \bar{1}$];
- $\circ(\bar{11}) = 2$ [infatti $\bar{11}^2 = \bar{121} = \bar{1}$];
- $\circ(\bar{13}) = 4$ [infatti $\bar{13}^2 = \bar{169} = \bar{4}, \bar{13}^4 = \bar{4}^2 = \bar{1}$];
- $\circ(\bar{14}) = 2$ [infatti $\bar{14}^2 = \bar{2}^2 \cdot \bar{7}^2 = \bar{4} \cdot \bar{4} = \bar{1}$].

Si vede quindi che $\mathbf{U}(\mathbf{Z}_{15})$ non è ciclico. Tale gruppo contiene quattro elementi di periodo 4 e quindi due sottogruppi C_4 :

$$\langle \bar{2} \rangle = \{\bar{1}, \bar{2}, \bar{4}, \bar{8}\}, \quad \langle \bar{7} \rangle = \{\bar{1}, \bar{7}, \bar{4}, \bar{13}\}.$$

Contiene inoltre tre elementi di periodo 2 e quindi tre sottogruppi C_2 :

$$\langle \bar{4} \rangle, \quad \langle \bar{11} \rangle, \quad \langle \bar{14} \rangle.$$

* * *

5.2.3. Verificare che, a meno di isomorfismi, esistono due soli gruppi di ordine 6: C_6 e S_3 .

Soluzione. Un gruppo di ordine 6 se è ciclico ha elementi di periodo 2 e di periodo 3. Se non è ciclico, dal teorema di Lagrange ha elementi di periodo 1, 2, 3. In base ai risultati dell'**Eserc. 5.1.4**, il gruppo deve possedere almeno un elemento di periodo 2 ed almeno un elemento di periodo 3.

Assumiamo quindi che in (G, \cdot) esistano elementi x, y , con $\circ(x) = 2$, $\circ(y) = 3$. Allora G contiene 1, x, y, y^2 , distinti a due a due. Poiché $xy \neq 1, x, y, y^2$, xy è un altro elemento di G . Poiché poi $xy^2 \neq 1, x, y, y^2, xy$, allora xy^2 è il sesto elemento di G .

Consideriamo ora l'elemento $yx \in G$. Poiché $yx \neq 1, x, y, y^2$, allora si hanno due eventualità:

$$yx = xy \quad \text{oppure} \quad yx = xy^2.$$

Nel primo caso si ha:

$$(xy)^2 = (xy)(xy) = x(yx)y = x(xy)y = x^2y^2 = y^2, \quad (xy)^3 = xy y^2 = x.$$

Ma allora $\circ(xy) > 3$ e dunque necessariamente $\circ(xy) = 6$. Dunque $G \cong C_6$.

Nel secondo caso si ha: $yx = xy^2$. In tal caso $G \cong S_3$. Infatti risulta:

$$G = \langle x, y \mid x^2 = y^3 = 1, yx = xy^2 \rangle$$

e S_3 verifica tali relazioni, con $x = (1\ 2)$, $y = (1\ 2\ 3)$.

* * *

5.2.4. Verificare che $\mathcal{U}(\mathbf{Z}_{16})/\langle \bar{7} \rangle$ è un gruppo ciclico di ordine 4.

Soluzione. Poiché $\varphi(16) = 2^4 - 2^3 = 8$, $\mathcal{U}(\mathbf{Z}_{16})$ è un gruppo (abeliano) di ordine 8. Risulta:

$$\mathcal{U}(\mathbf{Z}_{16}) = \{\bar{1}, \bar{3}, \bar{5}, \bar{7}, \bar{9}, \bar{11}, \bar{13}, \bar{15}\}.$$

Poiché $\circ(\bar{7}) = 2$ [in quanto $\bar{7}^2 = \bar{49} = \bar{1}$], allora $\langle \bar{7} \rangle = \{\bar{1}, \bar{7}\}$. Ne segue:

$$|\mathcal{U}(\mathbf{Z}_{16})/\langle \bar{7} \rangle| = |\mathfrak{L}_s(\langle \bar{7} \rangle)| = \frac{8}{2} = 4.$$

Il gruppo quoziante $\mathcal{U}(\mathbf{Z}_{16})/\langle \bar{7} \rangle$ è formato da

$$\langle \bar{7} \rangle = \{\bar{1}, \bar{7}\}, \quad \bar{3}\langle \bar{7} \rangle = \{\bar{3}, \bar{5}\}, \quad \bar{9}\langle \bar{7} \rangle = \{\bar{9}, \bar{15}\}, \quad \bar{11}\langle \bar{7} \rangle = \{\bar{11}, \bar{13}\}.$$

Per verificare se tale gruppo è ciclico o di Klein esaminiamo i periodi dei suoi elementi. Si ha:

$$(\bar{3}\langle \bar{7} \rangle)^2 = \bar{9}\langle \bar{7} \rangle \neq \langle \bar{7} \rangle.$$

Dunque $\circ(\bar{3}\langle \bar{7} \rangle) > 2$ e quindi necessariamente $\circ(\bar{3}\langle \bar{7} \rangle) = 4$. Segue che

$$\mathcal{U}(\mathbf{Z}_{16})/\langle \bar{7} \rangle = \langle \bar{3}\langle \bar{7} \rangle \rangle \cong C_4.$$

[Si noti che $\circ(\bar{9}\langle \bar{7} \rangle) = 2$ mentre $\circ(\bar{11}\langle \bar{7} \rangle) = 4$].

* * *

5.2.5. Il sottogruppo $\langle (1\ 2\ 3\ 4) \rangle$ è normale in S_4 ?

Soluzione. Il sottogruppo $H = \langle (1\ 2\ 3\ 4) \rangle$ è ciclico di ordine 4. Si ha:

$$H = \langle (1\ 2\ 3\ 4) \rangle = \{(1), (1\ 2\ 3\ 4), (1\ 3)(2\ 4), (1\ 4\ 3\ 2)\}.$$

L'indice di H in S_4 è $i = \frac{24}{4} = 6$. Tra i sei laterali sinistri consideriamo ad esempio $(1\ 2)H$ e confrontiamolo con il corrispondente laterale destro $H(1\ 2)$. Si ha:

$$(1\ 2)H = \{(1\ 2)(1), (1\ 2)(1\ 2\ 3\ 4), (1\ 2)(1\ 3)(2\ 4), (1\ 2)(1\ 4\ 3\ 2)\} = \{(1\ 2), (1\ 3\ 4), (1\ 4\ 2\ 3), (2\ 4\ 3)\},$$

$$H(1\ 2) = \{(1)(1\ 2), (1\ 2\ 3\ 4)(1\ 2), (1\ 3)(2\ 4)(1\ 2), (1\ 4\ 3\ 2)(1\ 2)\} = \{(1\ 2), (2\ 3\ 4), (1\ 3\ 2\ 4), (1\ 4\ 3)\}.$$

Come si osserva, $(1\ 2)H \neq H(1\ 2)$. Pertanto H non è normale in S_4 .

* * *

5.2.6. Sia $\mathcal{H} = \{A \in \mathbf{GL}_n(K) \mid \det(A) = 1\}$. Verificare che è un sottogruppo normale di $\mathbf{GL}_n(K)$ e determinare il gruppo quoziante $\mathbf{GL}_n(K)/\mathcal{H}$.

Soluzione. Si consideri l'applicazione $\det : \mathbf{GL}_n(K) \rightarrow K^\times$ che associa ad ogni matrice invertibile il suo determinante: $\det(A) = |A| \forall A \in \mathbf{GL}_n(K)$.

Per il teorema di Binet,

$$\det(AB) = |AB| = |A||B| = \det(A)\det(B)$$

Dunque \det è un omomorfismo dal gruppo $(\mathbf{GL}_n(K), \cdot)$ al gruppo (K^\cdot, \cdot) . Inoltre \det è suriettivo. Infatti, $\forall a \in K^\cdot$, la matrice

$$\begin{pmatrix} a & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

ha determinante a . Infine

$$\text{Ker}(\det) = \{A \in \mathbf{GL}_n(K) \mid \det(A) = 1\} = \mathcal{H}.$$

Segue dal teorema fondamentale di omomorfismo che \mathcal{H} è normale in $\mathbf{GL}_n(K)$ e che

$$\mathbf{GL}_n(K)/_{\mathcal{H}} \cong K^\cdot.$$

* * *

5.2.7. Sia $\mathcal{H} = \{A \in \mathbf{GL}_2(\mathbf{Z}_5) \mid \det(A) = \bar{1} \text{ oppure } \det(A) = \bar{4}\}$. Verificare che \mathcal{H} è un sottogruppo normale di $\mathbf{GL}_2(\mathbf{Z}_5)$ e che il gruppo quoziante $\mathbf{GL}_2(\mathbf{Z}_5)/_{\mathcal{H}}$ è ciclico. Indicarne poi un generatore.

Soluzione. Poiché $\bar{4} = -\bar{1}$, allora

$$\mathcal{H} = \{A \in \mathbf{GL}_2(\mathbf{Z}_5) \mid \det(A) = \pm \bar{1}\} = \{A \in \mathbf{GL}_2(\mathbf{Z}_5) \mid \det(A)^2 = \bar{1}\}.$$

Si consideri l'applicazione $\varphi : \mathbf{GL}_2(\mathbf{Z}_5) \rightarrow \mathbf{Z}_5^\cdot$ tale che

$$\varphi(A) = (\det(A))^2, \quad \forall A \in \mathbf{GL}_2(\mathbf{Z}_5).$$

Per il teorema di Binet,

$$\varphi(AB) = |AB|^2 = |A|^2|B|^2 = \varphi(A)\varphi(B).$$

Dunque φ è un omomorfismo dal gruppo $(\mathbf{GL}_2(\mathbf{Z}_5), \cdot)$ al gruppo $(\mathbf{Z}_5^\cdot, \cdot)$. Inoltre $\text{Ker}(\varphi) = \mathcal{H}$ e pertanto \mathcal{H} è un sottogruppo normale di $\mathbf{GL}_2(\mathbf{Z}_5)$. Infine

$$\text{Im}(\varphi) = \{\det(A)^2, \quad \forall A \in \mathbf{GL}_2(\mathbf{Z}_5)\} = \{x^2, \quad \forall x \in \mathbf{Z}_5^\cdot\} = \{\bar{1}^2, \bar{2}^2, \bar{3}^2, \bar{4}^2\} = \{\bar{1}, \bar{4}\} = \{\pm \bar{1}\}.$$

Si tratta di un gruppo ciclico di ordine due. Segue dal teorema fondamentale di omomorfismo che

$$\mathbf{GL}_2(\mathbf{Z}_5)/_{\mathcal{H}} \equiv \text{Im}(\varphi)$$

e dunque tale gruppo è ciclico di ordine 2.

Per ottenerne un generatore basta scegliere una matrice $A \in \mathbf{GL}_2(\mathbf{Z}_5)$ avente determinante $\neq \pm \bar{1}$.

Ad esempio, posto $A = \begin{pmatrix} \bar{2} & \bar{0} \\ \bar{0} & \bar{1} \end{pmatrix}$, si ottiene

$$\mathbf{GL}_2(\mathbf{Z}_5)/_{\mathcal{H}} = \langle A\mathcal{H} \rangle.$$

* * *

5.2.8. Verificare che se G è un gruppo ciclico ed H è un suo sottogruppo, il gruppo quoziante G/H è ciclico.

Soluzione. Poiché ogni gruppo ciclico G è abeliano, ogni suo sottogruppo H è normale e quindi G/H è un gruppo.

Se G è infinito, $G \cong (\mathbf{Z}, +)$. Assumiamo quindi $G = \mathbf{Z}$ e $H = n\mathbf{Z}$. Se $n \neq 0$, è noto che

$$\mathbf{Z}/_H = \mathbf{Z}/_{n\mathbf{Z}} \cong \mathbf{Z}_n \text{ (gruppo ciclico di ordine } n\text{).}$$

Se invece $H = 0\mathbf{Z} = \{0\}$, applicando il teorema fondamentale di omomorfismo all'applicazione identica $\mathbf{1}_{\mathbf{Z}} : \mathbf{Z} \rightarrow \mathbf{Z}$, segue subito che

$$\mathbf{Z}/_{\{0\}} = \mathbf{Z}/_{\text{Ker}(\mathbf{1}_{\mathbf{Z}})} \cong \mathbf{Z} \text{ (gruppo ciclico).}$$

Dunque ogni gruppo $\mathbf{Z}/_H$ è ciclico.

Sia ora $G = C_n = \langle x \mid x^n = 1 \rangle$. È noto che i sottogruppi di C_n sono esattamente i seguenti sottogruppi ciclici $\langle x^{n/d} \rangle$, per ogni d divisore positivo di n . Poniamo quindi $H = \langle x^t \rangle$, con $t = \frac{n}{d}$.

Poiché $|H| = \circ(x^t) = d$, il gruppo quoziante $G/H = C_n/\langle x^t \rangle$ ha ordine $\frac{n}{d} = t$. Tale gruppo quoziante contiene i laterali

$$x^i \langle x^t \rangle, \quad \forall i = 0, 1, \dots, t-1,$$

e tali laterali sono a due a due distinti, come si può verificare [se ad esempio $x^i \langle x^t \rangle = x^j \langle x^t \rangle$, allora $x^{i-j} \in \langle x^t \rangle$ e quindi $i-j = tq$, da cui $i = j$, essendo $0 \leq i, j < t$]. Allora

$$G/H = \{\langle x^t \rangle, x \langle x^t \rangle, \dots, x^{t-1} \langle x^t \rangle\} = \langle x \langle x^t \rangle \rangle \cong C_t.$$

Dunque G/H è ciclico di ordine t .

* * *

5.2.9. Si consideri in $(\mathbf{Q}, +)$ il sottogruppo $(\mathbf{Z}, +)$ ed il gruppo quoziante $(\mathbf{Q}/\mathbf{Z}, +)$.

(i) Verificare che ogni elemento di \mathbf{Q}/\mathbf{Z} è rappresentabile con un razionale q tale che $0 \leq q < 1$.

(ii) Verificare che ogni elemento di \mathbf{Q}/\mathbf{Z} ha periodo finito e dedurne che \mathbf{Q}/\mathbf{Z} non è ciclico.

Soluzione. (i) Ogni elemento di \mathbf{Q}/\mathbf{Z} è del tipo

$$\frac{a}{b} + \mathbf{Z}, \quad \text{con } a, b \in \mathbf{Z}, b \neq 0.$$

Non è restrittivo assumere $b > 0$ [cambiando eventualmente segno ad a]. Sia ora

$$a = bh + r, \quad \text{con } h, r \in \mathbf{Z} \quad 0 \leq r < b.$$

Allora

$$\frac{a}{b} = h + \frac{r}{b}, \quad \text{con } 0 \leq \frac{r}{b} < 1.$$

Ne segue:

$$\frac{a}{b} + \mathbf{Z} = (h + \frac{r}{b}) + \mathbf{Z} = \frac{r}{b} + \mathbf{Z}.$$

Posto $q = \frac{r}{b} \in \mathbf{Q}$, si ha che $\frac{a}{b} + \mathbf{Z} = q + \mathbf{Z}$ e $0 \leq q < 1$.

(ii) Per ogni $\frac{a}{b} + \mathbf{Z} \in \mathbf{Q}/\mathbf{Z}$ si ha:

$$|b|(\frac{a}{b} + \mathbf{Z}) = a \frac{|b|}{b} + \mathbf{Z} = \pm a + \mathbf{Z} = \mathbf{Z}.$$

Dunque $\circ(\frac{a}{b} + \mathbf{Z}) \leq |b| < \infty$.

Poiché \mathbf{Q}/\mathbf{Z} è un gruppo infinito e non ha elementi di periodo ∞ , non è isomorfo a \mathbf{Z} e quindi non è ciclico.

* * *

5.2.10. Nel gruppo moltiplicativo $(\mathbf{C}^\cdot, \cdot)$ si considerino i due sottoinsiemi

$$\mathbf{R}^{>0} \quad (\text{numeri reali positivi}), \quad \mathcal{H} = \{z \in \mathbf{C}^\cdot \mid |z| = 1\}.$$

(i) Verificare che $\mathbf{R}^{>0}$ ed \mathcal{H} sono sottogruppi di $(\mathbf{C}^\cdot, \cdot)$.

(ii) Dimostrare che il gruppo quoziante $\mathbf{C}^\cdot /_{\mathbf{R}^{>0}}$ è isomorfo a \mathcal{H} .

Soluzione. (i) Se $r, s \in \mathbf{R}^{>0}$, si ha:

$$rs \in \mathbf{R}^{>0}, \quad \frac{1}{r} \in \mathbf{R}^{>0}.$$

Pertanto $\mathbf{R}^{>0}$ è un sottogruppo di $(\mathbf{C}^\cdot, \cdot)$.

Se $z_1, z_2 \in \mathcal{H}$, ovviamente $|z_1 z_2| = |z_1| |z_2| = 1 \cdot 1 = 1$ e quindi $z_1 z_2 \in \mathcal{H}$. Inoltre, da $z \frac{1}{z} = 1$ segue che $1 = |z| |\frac{1}{z}|$ e quindi $|\frac{1}{z}| = \frac{1}{|z|}$. Pertanto, se $z \in \mathcal{H}$, anche $\frac{1}{z} \in \mathcal{H}$. Dunque anche \mathcal{H} è un sottogruppo di $(\mathbf{C}^\cdot, \cdot)$.

(ii) Utilizziamo il Teorema fondamentale di omomorfismo, considerando l'applicazione

$$\varphi : \mathbf{C}^\cdot \rightarrow \mathcal{H} \quad \text{tale che } \varphi(z) = \frac{z}{|z|}, \quad \forall z \in \mathbf{C}^\cdot.$$

Si osservi che $\varphi(z) \in \mathcal{H}$ [in quanto $|\frac{z}{|z|}| = \frac{|z|}{|z|} = 1$]. Inoltre φ è un omomorfismo [infatti $\varphi(z_1 z_2) = \frac{z_1 z_2}{|z_1 z_2|} = \frac{z_1}{|z_1|} \frac{z_2}{|z_2|} = \varphi(z_1) \varphi(z_2)$]. L'omomorfismo φ è suriettivo [infatti, $\forall z \in \mathcal{H}, \varphi(z) = z$]. Calcoliamo $\text{Ker}(\varphi)$. Si ha:

$$\text{Ker}(\varphi) = \{z \in \mathbf{C}^\cdot \mid \varphi(z) = 1\} = \{z \in \mathbf{C}^\cdot \mid z = |z|\} = \{z \in \mathbf{C}^\cdot \mid z \in \mathbf{R}^{>0}\} = \mathbf{R}^{>0}.$$

Segue dal Teorema fondamentale di omomorfismo che

$$\mathbf{C}^\cdot /_{_{Ker(\varphi)}} \cong Im(\varphi) \text{ cioè } \mathbf{C}^\cdot /_{_{R^{>0}}} \cong \mathcal{H}.$$

* * *

ESERCIZI FINALI

[in preparazione al secondo esonero]

- 1.** Si verifichi che $\mathcal{U}(\mathbf{Z}_7)$ è un gruppo ciclico di ordine 6 e si tracci il reticolo dei suoi sottogruppi.

Soluzione. Risulta:

$$\mathcal{U}(\mathbf{Z}_7) = \{\bar{1}, \bar{2}, \bar{3}, \bar{4}, \bar{5}, \bar{6}\}.$$

Calcoliamo i periodi degli elementi di $\mathcal{U}(\mathbf{Z}_7)$. Si ha:

$$\circ(\bar{2}) = 3. \text{ Infatti } \bar{2}^2 = \bar{4}, \bar{2}^3 = \bar{4} \cdot \bar{2} = \bar{8} = \bar{1}.$$

$$\circ(\bar{3}) = 6. \text{ Infatti } \bar{3}^2 = \bar{9} = \bar{2}, \bar{3}^3 = \bar{6}, \bar{3}^4 = \bar{4}, \bar{3}^5 = \bar{5}, \bar{3}^6 = \bar{1}.$$

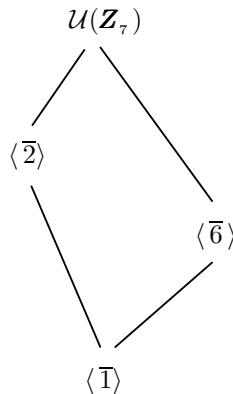
Poiché $\circ(\bar{3}) = 6$, allora $\mathcal{U}(\mathbf{Z}_7)$ è ciclico, generato da $\bar{3}$. I restanti elementi di $\mathcal{U}(\mathbf{Z}_7)$ hanno i seguenti periodi:

$$\circ(\bar{4}) = \circ(\bar{3}^4) = \frac{\circ(\bar{3})}{MCD(6,4)} = \frac{6}{2} = 3;$$

$$\circ(\bar{5}) = \circ(\bar{3}^5) = \frac{\circ(\bar{3})}{MCD(6,5)} = 6$$

$$\circ(\bar{6}) = \circ(\bar{3}^3) = \frac{\circ(\bar{3})}{MCD(6,3)} = \frac{6}{3} = 2.$$

Il gruppo $\mathcal{U}(\mathbf{Z}_7) [= \langle \bar{3} \rangle = \langle \bar{5} \rangle]$ ha due sottogruppi propri: $\langle \bar{2} \rangle = \langle \bar{4} \rangle$, ciclico di ordine 3 e $\langle \bar{6} \rangle$, ciclico di ordine 2. Il reticolo dei sottogruppi di $\mathcal{U}(\mathbf{Z}_7)$ è il seguente:



* * *

- 2.** Si considerino il gruppo moltiplicativo $\mathcal{U}(\mathbf{Z}_{16})$ ed il gruppo moltiplicativo $\mathcal{U}(\mathbf{Z}_{24})$. Dei due gruppi si determini il periodo di ogni elemento ed il reticolo dei sottogruppi. I due gruppi sono isomorfi?

Soluzione. I gruppi $\mathcal{U}(\mathbf{Z}_{16})$ e $\mathcal{U}(\mathbf{Z}_{24})$ hanno entrambi ordine 8. Infatti $\varphi(16) = 8 = \varphi(24)$.

Si ha:

$$\mathcal{U}(\mathbf{Z}_{16}) = \{\bar{1}, \bar{3}, \bar{5}, \bar{7}, \bar{9}, \bar{11}, \bar{13}, \bar{15}\}.$$

Calcoliamo i periodi degli elementi di $\mathcal{U}(\mathbf{Z}_{16})$.

Risulta, con semplici calcoli:

$$\circ(\bar{3}) = \circ(\bar{5}) = \circ(\bar{11}) = \circ(\bar{13}) = 4, \quad \circ(\bar{7}) = \circ(\bar{9}) = \circ(\bar{15}) = 2.$$

$\mathcal{U}(\mathbf{Z}_{16})$ possiede tre elementi di periodo 2 e quindi tre gruppi ciclici di ordine 3, cioè

$$\langle \bar{9} \rangle, \quad \langle \bar{7} \rangle, \quad \langle \bar{15} \rangle.$$

Poiché inoltre esistono quattro elementi di periodo 4, $\mathcal{U}(\mathbf{Z}_{16})$ possiede due gruppi ciclici di ordine 4. Si tratta di

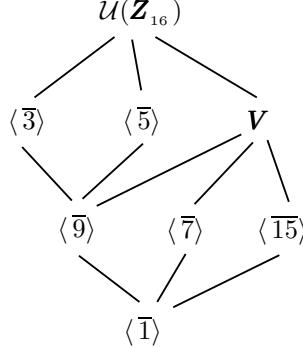
$$\langle \bar{3} \rangle = \langle \bar{11} \rangle, \quad \langle \bar{5} \rangle = \langle \bar{13} \rangle.$$

Si noti che entrambi contengono lo stesso sottogruppo ciclico $\langle \bar{9} \rangle$. Infine, i tre elementi di periodo 2 generano un gruppo di Klein

$$\mathbf{V} = \{\bar{1}, \bar{7}, \bar{9}, \bar{15}\}$$

[infatti tale insieme è chiuso rispetto al prodotto, essendo $\bar{7} \cdot \bar{9} = \bar{15}$]. Ovviamente \mathbf{V} contiene i tre sottogruppi ciclici di ordine 2.

Siamo in grado di tracciare il grafico dei sottogruppi di $\mathcal{U}(\mathbf{Z}_{16})$:



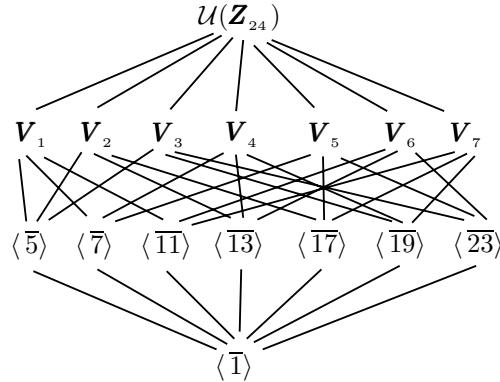
Ora esaminiamo $\mathcal{U}(\mathbf{Z}_{24})$. Si ha:

$$\mathcal{U}(\mathbf{Z}_{24}) = \{\bar{1}, \bar{5}, \bar{7}, \bar{11}, \bar{13}, \bar{17}, \bar{19}, \bar{23}\}.$$

Si verifica, con semplici calcoli, che $\mathcal{U}(\mathbf{Z}_{24})$ possiede sette elementi di periodo 2. Quindi tale gruppo possiede sette sottogruppi ciclici di ordine 2. Non può possedere sottogruppi ciclici di ordine 4, ma ha vari sottogruppi di Klein. Per ottenerli, basta considerare due elementi distinti di periodo 2 e moltiplicarli tra loro. I tre elemtni ottenuti generano un gruppo di Klein. Scegliendo i primi generatori con rappresentanti in ordine crescente (ed escludendo sottogruppi già ottenuti) si ottengono sette sottogruppi di Klein:

$$\begin{aligned} \mathbf{V}_1 &= \{\bar{1}, \bar{5}, \bar{7}, \bar{11}\}, & \mathbf{V}_2 &= \{\bar{1}, \bar{5}, \bar{13}, \bar{17}\}, & \mathbf{V}_3 &= \{\bar{1}, \bar{5}, \bar{19}, \bar{23}\}, & \mathbf{V}_4 &= \{\bar{1}, \bar{7}, \bar{13}, \bar{19}\}, \\ \mathbf{V}_5 &= \{\bar{1}, \bar{7}, \bar{17}, \bar{23}\}, & \mathbf{V}_6 &= \{\bar{1}, \bar{11}, \bar{13}, \bar{23}\}, & \mathbf{V}_7 &= \{\bar{1}, \bar{11}, \bar{17}, \bar{19}\}. \end{aligned}$$

Siamo ora in grado di tracciare il grafico dei sottogruppi di $\mathcal{U}(\mathbf{Z}_{24})$:



Si noti che i due gruppi $\mathcal{U}(\mathbf{Z}_{16})$ e $\mathcal{U}(\mathbf{Z}_{24})$ non sono isomorfi perché i periodi dei loro elementi non coincidono, mentre un isomorfismo fissa i periodi degli elementi.

* * *

3. Siano \mathbf{E} ed \mathbf{F} due basi di $V = V_{\mathbf{R}}^5$, legate dalla seguente formula di cambiamento di base:

$$\underline{f}_1 = \underline{e}_1 + \underline{e}_2 + \underline{e}_3, \quad \underline{f}_2 = \underline{e}_2 + \underline{e}_3 + \underline{e}_4, \quad \underline{f}_3 = \underline{e}_1, \quad \underline{f}_4 = \underline{e}_2, \quad \underline{f}_5 = \underline{e}_5.$$

Sia U il sottospazio vettoriale di V avente in base \mathbf{F} equazioni cartesiane

$$\begin{cases} y_1 - y_2 + y_4 - y_5 = 0 \\ y_2 - y_3 + y_4 = 0 \\ y_2 + y_3 + y_5 = 0. \end{cases}$$

Determinare equazioni cartesiane di U in base E ed una base di U (sempre in base E).

Soluzione. Dai dati assegnati, risulta: $\mathbf{F} = \mathbf{E} C$, con

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

Per ogni $\underline{v} \in V$, sia $\underline{v} = \mathbf{E} \mathbf{x} = \mathbf{F} \mathbf{y}$. Ne segue che $\mathbf{E} \mathbf{x} = \mathbf{E} C \mathbf{y}$, cioè $\mathbf{x} = C \mathbf{y}$. Invertendo,

$$\mathbf{y} = C^{-1} \mathbf{x}.$$

Dovendo determinare tale formula, osserviamo che è più semplice invertire la formula $\mathbf{x} = C \mathbf{y}$.

$$\text{Infatti da } \begin{cases} x_1 = y_1 + y_3 \\ x_2 = y_1 + y_2 + y_4 \\ x_3 = y_1 + y_2 \\ x_4 = y_2 \\ x_5 = y_5 \end{cases} \text{ segue subito } \begin{cases} y_1 = x_3 - x_4 \\ y_2 = x_4 \\ y_3 = x_1 - x_3 + x_4 \\ y_4 = x_2 - x_3 \\ y_5 = x_5. \end{cases}$$

Sostituendo tale formula nelle equazioni cartesiane di U in base F , si ottengono le seguenti equazioni cartesiane di U in base E :

$$\begin{cases} x_2 - 2x_4 - x_5 = 0 \\ x_1 - x_2 = 0 \\ x_1 - x_3 + 2x_4 + x_5 = 0. \end{cases}$$

Per ottenere una base di U (rispetto ad E), risolviamo tale *SLO*, che ha matrice dei coefficienti:

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

Scegliamo $B(1, 2, 3|1, 2, 3)$ come sottomatrice fondamentale. Ponendo $x_4 = 1$, $x_5 = 0$, si ottiene la soluzione $(2, 2, 4, 1, 0)$; ponendo $x_4 = 0$, $x_5 = 1$, si ottiene la soluzione $(1, 1, 2, 0, 1)$. Pertanto una base di U è data da

$$\underline{u}_1 = 2\underline{e}_1 + 2\underline{e}_2 + 4\underline{e}_3 + \underline{e}_4, \quad \underline{u}_2 = \underline{e}_1 + \underline{e}_2 + 2\underline{e}_3 + \underline{e}_5.$$

* * *

4. Sia $V = V_{\mathbf{R}}^3$ uno spazio vettoriale con base E . Sia W il sottospazio vettoriale di equazioni cartesiane (in base E) $\{x_1 - x_2 + 2x_3 = 0\}$. Determinare una base \mathbf{F} di V tale che W abbia, in tale base, equazioni cartesiane $\{y_2 = 0\}$.

Soluzione. Una base di W è ottenuta risolvendo il *SLO* $(1, 3, \mathbf{R})$ $\{x_1 - x_2 + 2x_3 = 0\}$. Si ottengono le due soluzioni linearmente indipendenti $(1, 1, 0)$, $(-2, 0, 1)$. Pertanto W ha base

$$\underline{w}_1 = \underline{e}_1 + \underline{e}_2, \quad \underline{w}_2 = -2\underline{e}_1 + \underline{e}_3.$$

Poiché W ha, in base \mathbf{F} , equazione $y_2 = 0$, risulta: $W = \langle \underline{f}_1, \underline{f}_3 \rangle$. Essendo $W = \langle \underline{w}_1, \underline{w}_2 \rangle = \langle \underline{f}_1, \underline{f}_3 \rangle$, dal momento che è richiesta una base \mathbf{F} , possiamo ad sempio porre

$$\underline{f}_1 = \underline{w}_1, \quad \underline{f}_3 = \underline{w}_2.$$

Per ottenere \underline{f}_2 basterà scegliere un vettore tale che $\underline{w}_1, \underline{w}_2, \underline{f}_2$ siano linearmente indipendenti. Tra le infinite possibili scelte per \underline{f}_2 , poniamo $\underline{f}_2 = \underline{e}_2$. Allora

$$\mathbf{F} = (\underline{w}_1 \quad \underline{f}_2 \quad \underline{w}_2) = \mathbf{E} C, \quad \text{con } C = \begin{pmatrix} 1 & 0 & -2 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

In base \mathbf{F} , W ha equazioni cartesiane $\{y_2 = 0\}$.

N.B. Possiamo eseguire una verifica dei calcoli. Da $\underline{v} = \mathbf{E} \mathbf{x} = \mathbf{F} \mathbf{y} = \mathbf{E} C \mathbf{y}$ segue $\mathbf{x} = C \mathbf{y}$, cioè

$$\begin{cases} x_1 = y_1 - 2y_3 \\ x_2 = y_1 + y_2 \\ x_3 = y_3. \end{cases}$$

Sostituendo tale formula nell'equazione $x_1 - x_2 + 2x_3 = 0$, si ottiene $-y_2 = 0$ dunque W ha in base \mathbf{F} equazione $y_2 = 0$, come previsto.

* * *

5. Sono assegnati in \mathbf{R}^4 due sottospazi vettoriali U_1, U_2 dipendenti da un parametro reale a e definiti (rispetto alla base canonica \mathbf{E} di \mathbf{R}^4) rispettivamente dalle seguenti equazioni cartesiane:

$$\begin{cases} ax_1 + x_2 - x_4 = 0 \\ x_1 + ax_3 = 0, \end{cases} \quad \begin{cases} -x_1 + ax_3 = 0 \\ ax_1 + x_2 + x_4 = 0. \end{cases}$$

- (i) Determinare per quali $a \in \mathbf{R}$ i due sottospazi U_1, U_2 sono supplementari.
- (ii) Posto $a = 1$, determinare (in una base opportuna) le matrici di tutti gli operatori lineari T di \mathbf{R}^4 tali che

$$T(U_1) \subseteq U_2 \text{ e } T(U_2) \subseteq U_1.$$

- (iii) Tra tali operatori lineari individuare quali sono automorfismi.

Soluzione. (i) Il sottospazio vettoriale $U_1 \cap U_2$ ha equazioni cartesiane ottenute riunendo in un unico sistema i due *SLO* $A_1 X = \mathbf{0}, A_2 X = \mathbf{0}$, con

$$A_1 = \begin{pmatrix} a & 1 & 0 & -1 \\ 1 & 0 & a & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} -1 & 0 & a & 0 \\ a & 1 & 0 & 1 \end{pmatrix}.$$

Le equazioni cartesiane di $U_1 \cap U_2$ sono quindi date dal *SLO* $(4, 4, \mathbf{R})$

$$\begin{pmatrix} A_1 \\ A_2 \end{pmatrix} X = \mathbf{0}.$$

In base alla formula di Grassmann, i due sottospazi vettoriali U_1, U_2 sono supplementari $\iff dim(U_1 \cap U_2) = 0 \iff rg\left(\begin{matrix} A_1 \\ A_2 \end{matrix}\right) = 4 \iff det\left(\begin{matrix} A_1 \\ A_2 \end{matrix}\right) \neq 0$. Risulta:

$$det\left(\begin{matrix} A_1 \\ A_2 \end{matrix}\right) = \begin{vmatrix} a & 1 & 0 & -1 \\ 1 & 0 & a & 0 \\ -1 & 0 & a & 0 \\ a & 1 & 0 & 1 \end{vmatrix} = -4a.$$

Dunque U_1, U_2 sono supplementari $\iff a \neq 0$.

(ii) Se $a = 1$, U_1, U_2 sono supplementari. Vogliamo determinare una base \mathbf{F} di \mathbf{R}^4 ottenuta riunendo una base di U_1 ed una di U_2 . Basta risolvere i due *SLO* (con $a = 1$). Si ottiene ad esempio per U_1 la base

$$\underline{u}_1 = (-1, 1, 1, 0), \quad \underline{v}_1 = (0, 1, 0, 1)$$

e per U_2 la base

$$\underline{u}_2 = (1, -1, 1, 0), \quad \underline{v}_2 = (0, -1, 0, 1).$$

Allora

$$\mathbf{F} = \mathbf{E} C, \text{ con } C = \begin{pmatrix} -1 & 0 & 1 & 0 \\ 1 & 1 & -1 & -1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}.$$

Dai dati, $T(\underline{u}_1), T(\underline{v}_1) \in \langle \underline{u}_2, \underline{v}_2 \rangle$ e quindi, rispetto alla base \mathbf{F} , per opportuni $b_i \in \mathbf{R}$:

$$T(\underline{u}_1) = b_1 \underline{u}_2 + b_2 \underline{v}_2, \quad T(\underline{v}_1) = b_3 \underline{u}_2 + b_4 \underline{v}_2.$$

Analogamente, $T(\underline{u}_2), T(\underline{v}_2) \in \langle \underline{u}_1, \underline{v}_1 \rangle$ e quindi, sempre rispetto alla base \mathbf{F} , per opportuni $c_i \in \mathbf{R}$:

$$T(\underline{u}_2) = c_1 \underline{u}_1 + c_2 \underline{v}_1, \quad T(\underline{v}_2) = c_3 \underline{u}_1 + c_4 \underline{v}_1.$$

Allora

$$T(\mathbf{F}) = \mathbf{F} B, \text{ con } B = \begin{pmatrix} 0 & 0 & c_1 & c_3 \\ 0 & 0 & c_2 & c_4 \\ b_1 & b_3 & 0 & 0 \\ b_2 & b_4 & 0 & 0 \end{pmatrix}.$$

(iii) GLi automorfismi tra gli operatori T sopra ottenuti sono tutti e soli quelli per cui $\det(B) \neq 0$. In base al teorema di Lagrange, risulta

$$\det(B) \neq 0 \iff \begin{vmatrix} b_1 & b_3 \\ b_2 & b_4 \end{vmatrix} \cdot \begin{vmatrix} c_1 & c_3 \\ c_2 & c_4 \end{vmatrix} \neq 0 \iff b_1 b_4 \neq b_2 b_3 \text{ e } c_1 c_4 \neq c_2 c_3.$$

* * *

6. Sia $V = V_{\mathbf{R}}^5$ uno spazio vettoriale con base \mathbf{E} . Sono assegnati due sottospazi vettoriali W_1, W_2 , definiti rispettivamente dalle equazioni cartesiane

$$\begin{cases} x_1 - x_2 = 0 \\ x_3 - x_4 = 0, \end{cases} \quad \begin{cases} x_1 - x_3 = 0 \\ x_2 + x_4 = 0 \\ x_5 = 0. \end{cases}$$

(i) Verificare che i due sottospazi sono supplementari determinare una base di ciascuno di essi.

(ii) Sia $T : V \rightarrow V$ l'operatore lineare definito dalle seguenti condizioni:

$$T|_{W_1} = \mathbf{1}_{W_1}, \quad T|_{W_2} = \mathbf{0}.$$

Determinare la matrice di T in una base opportuna, dire se T è diagonalizzabile e indicare (anche senza eseguire i calcoli) la matrice di T in base \mathbf{E} .

Soluzione. (i) I due sottospazi W_1, W_2 hanno in base \mathbf{E} rispettivamente equazioni cartesiane $A X = \mathbf{0}$, $B X = \mathbf{0}$, con

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

In virtù della formula di Grassmann, per dimostrare che i due sottospazi sono supplementari basta osservare che hanno rispettivamente dimensioni 3, 2 e dimostrare che $\text{rg} \begin{pmatrix} A \\ B \end{pmatrix} = 5$ [ciò che corrisponde a $\dim(W_1 \cap W_2) = 0$]. Si verifica subito infatti che $\det \begin{pmatrix} A \\ B \end{pmatrix} = -2 \neq 0$.

Ora calcoliamo una base di W_1 e di W_2 . Risolvendo il primo *SLO*, si ottengono tre soluzioni linearmente indipendenti:

$$(1, 1, 0, 0, 0), (0, 0, 1, 1, 0), (0, 0, 0, 0, 1).$$

Ad esse corrisponde la base di W_1 :

$$\underline{e}_1 + \underline{e}_2, \quad \underline{e}_3 + \underline{e}_4, \quad \underline{e}_5.$$

Risolvendo il secondo *SLO*, si ottengono due soluzioni linearmente indipendenti:

$$(1, 0, 1, 0, 0), (0, 1, 0, -1, 0).$$

Ad esse corrisponde la base di W_2 :

$$\underline{e}_1 + \underline{e}_3, \quad \underline{e}_2 - \underline{e}_4.$$

(ii) Sia \mathbf{F} la base di V ottenuta riunendo tali basi. Si ottiene

$$\mathbf{F} = \mathbf{E} C, \text{ con } C = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix}.$$

La condizione $T|_{W_1} = \mathbf{1}_{W_1}$ significa che T opera come l'identità sui vettori di W_1 . Analogamente, $T|_{W_2} = \mathbf{0}$ significa che T opera come l'operatore nullo sui vettori di W_2 . Ne segue che

$$T(\mathbf{F}) = \mathbf{F} D, \text{ con } D = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Poiché D è diagonale, T è diagonalizzabile. In base \mathbf{E} , T ha la seguente formula di definizione:

$$T(\mathbf{E}) = T(\mathbf{F} C^{-1}) = T(\mathbf{F}) C^{-1} = \mathbf{F} D C^{-1} = \mathbf{E} C D C^{-1}.$$

Dunque, in base \mathbf{E} , T ha matrice $A = C D C^{-1}$.

* * *

7. In \mathbf{R}^3 , con base canonica \mathbf{E} , è assegnato un operatore lineare $T : \mathbf{R}^3 \rightarrow \mathbf{R}^3$ tale che:

$$T(\underline{v}) = (1, 0, 1), \quad T^2(\underline{v}) = (1, 0, 0), \quad T^3(\underline{v}) = (0, 1, 0), \quad T^4(\underline{v}) = (0, 0, 0)$$

[con $\underline{v} \in \mathbf{R}^3$ vettore non assegnato].

(i) Determinare la matrice di T in base \mathbf{E} .

(ii) Verificare se T è diagonalizzabile e indicare le basi di ciascun autospazio, in base \mathbf{E} .

Soluzione. (i) Dai dati dell'esercizio si nota che è assegnata l'immagine di T dei tre vettori $T(\underline{v})$, $T^2(\underline{v})$, $T^3(\underline{v})$. Se questi tre vettori formano una base di V , la matrice di T verrà facilmente determinata, rispetto a tale base. Si ha:

$$\mathbf{F} := (T(\underline{v}) \quad T^2(\underline{v}) \quad T^3(\underline{v})) = \mathbf{E} C, \text{ con } C = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}.$$

Poiché $\det(C) \neq 0$, \mathbf{F} è una base di \mathbf{R}^3 . Si ha:

$$T(\mathbf{F}) = (T^2(\underline{v}) \quad T^3(\underline{v}) \quad T^4(\underline{v})) = \mathbf{E} D, \text{ con } D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Ne segue che

$$T(\mathbf{E}) = T(\mathbf{F} C^{-1}) = T(\mathbf{F}) C^{-1} = \mathbf{E} D C^{-1}.$$

Dunque la matrice di T in base \mathbf{E}

$$A = D C^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & -1 \\ 0 & 0 & 0 \end{pmatrix}.$$

(ii) T ha polinomio caratteristico

$$P_T = \begin{vmatrix} -X & 0 & 1 \\ 1 & -X & -1 \\ 0 & 0 & -X \end{vmatrix} = -X^3.$$

T ha il solo autovalore 0 ed il relativo autospazio $\mathbf{E}_0(T) = \text{Ker}(T)$ ha dimensione 1 [infatti $d_0 = \dim(\text{Ker}(T)) = 3 - rg(A) = 3 - 2 = 1$]. Poiché $d_0 < h_0 = 3$, T non è diagonalizzabile.

Per ottenere una base di $\mathbf{E}_0(T)$ bisogna risolvere il SLO $AX = \mathbf{0}$, cioè

$$\begin{cases} z = 0 \\ x - z = 0 \quad \text{ovvero} \\ 0 = 0, \end{cases} \quad \begin{cases} z = 0 \\ x = 0. \end{cases}$$

Pertanto $\mathbf{E}_0(T) = \langle (0, 1, 0) \rangle$.

N.B. Dai dati dell'esercizio era già noto che $T^3(\underline{v}) = (0, 1, 0)$ e $T^4(\underline{v}) = \underline{0} = 0 \cdot T^3(\underline{v})$. Dunque $T^3(\underline{v})$ era a priori un autovettore associato all'autovalore 0.

* * *

8. Siano $T : \mathbf{R}^3 \rightarrow \mathbf{R}^2$ e $S : \mathbf{R}^2 \rightarrow \mathbf{R}^3$ due applicazioni lineari, dipendenti da un parametro reale a , definite, rispetto alle basi canoniche \mathbf{E} di \mathbf{R}^3 ed \mathbf{F} di \mathbf{R}^2 , in questo modo:

$$\begin{cases} T(\underline{e}_1) = \underline{f}_1 - \underline{f}_2 \\ T(\underline{e}_2) = a\underline{f}_1 + \underline{f}_2 \\ T(\underline{e}_3) = \underline{0}, \end{cases} \quad \begin{cases} T(\underline{f}_1) = \underline{e}_1 - a\underline{e}_2 \\ T(\underline{f}_2) = \underline{e}_2 + \underline{e}_3. \end{cases}$$

Si consideri l'operatore lineare $S \circ T : \mathbf{R}^3 \rightarrow \mathbf{R}^3$.

- (i) Determinare la matrice di $S \circ T$ rispetto alla base \mathbf{E} .
- (ii) Determinare, in funzione di a , la dimensione di $Im(S \circ T)$ e di $Ker(S \circ T)$.
- (iii) Stabilire, per $a = 0, 1, -1$, se $S \circ T$ è diagonalizzabile, spiegandone il motivo.

Soluzione. (i) Si ha: $T(\mathbf{E}) = \mathbf{F} A$ e $S(\mathbf{F}) = \mathbf{E} B$, con

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

Allora $S \circ T$ rispetto alla base \mathbf{E} ha matrice

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

(ii) Poiché $dim(Im(S \circ T)) = rg(BA)$, calcoliamo $rg(BA)$. A priori $1 \leq rg(BA) \leq 2$. Scegliamo in BA il minore non nullo $|BA(1|1)| = 1$ ed orliamolo. In base al principio degli orlati,

$$rg(BA) = 1 \iff \begin{vmatrix} 1 & a \\ -(a+1) & 1-a^2 \end{vmatrix} = \begin{vmatrix} 1 & a \\ -1 & 1 \end{vmatrix} = 0 \iff \begin{cases} 1-a^2+a^2+a=0 \\ 1+a=0 \end{cases} \iff a=-1.$$

Dunque

$$dim(Im(S \circ T)) = \begin{cases} 2, & \text{se } a \neq -1 \\ 1, & \text{se } a = -1. \end{cases}$$

Di conseguenza

$$dim(Ker(S \circ T)) = 3 - dim(Im(S \circ T)) = \begin{cases} 1, & \text{se } a \neq -1 \\ 2, & \text{se } a = -1. \end{cases}$$

(iii) Sia $a = 1$. $S \circ T$ ha polinomio caratteristico

$$P_{S \circ T} = \begin{vmatrix} 1-X & 1 & 0 \\ -2 & -X & 0 \\ -1 & 1 & -X \end{vmatrix} = -X(X^2 - X + 2).$$

Poiché $X^2 - X + 2$ è irriducibile in $\mathbf{R}[X]$, allora $\Lambda(S \circ T) = \{0\}$. Essendo $d_0 = h_0 = 1 < 3$, $S \circ T$ non è diagonalizzabile.

Sia $a = 0$. $S \circ T$ ha polinomio caratteristico

$$P_{S \circ T} = \begin{vmatrix} 1-X & 0 & 0 \\ -1 & 1-X & 0 \\ -1 & 1 & -X \end{vmatrix} = -X(1-X)^2.$$

Allora $\Lambda(S \circ T) = \{0, 1\}$, con $h_0 = 1$, $h_1 = 2$. Per appurare se $S \circ T$ è diagonalizzabile, bisogna calcolare $d_1 = dim(\mathbf{E}_1(S \circ T))$. Tale autospazio ha equazioni cartesiane

$$\begin{pmatrix} 0 & 0 & 0 \\ -1 & 0 & 0 \\ -1 & 1 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

La matrice ha rango 2 e dunque $d_1 = 3 - 2 = 1$. Segue che $S \circ T$ non è diagonalizzabile.

Sia $a = -1$. $S \circ T$ ha polinomio caratteristico

$$P_{S \circ T} = \begin{vmatrix} 1-X & -1 & 0 \\ 0 & -X & 0 \\ -1 & 1 & -X \end{vmatrix} = -X^2(1-X).$$

Allora $\Lambda(S \circ T) = \{0, 1\}$, con $h_0 = 2$, $h_1 = 1$. Poiché $\mathbf{E}_0(S \circ T) = Ker(S \circ T)$ ha dimensione 2, allora $d_0 = 2$. In tal caso $S \circ T$ è diagonalizzabile.

9. In \mathbf{R}^4 (con base canonica \mathbf{E}) sono definiti:

- il sottospazio vettoriale W avente equazioni cartesiane $\begin{cases} x_1 - x_4 = 0 \\ x_2 + x_3 = 0; \end{cases}$

- l'operatore lineare T , avente matrice $A = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$.

(i) Determinare equazioni cartesiane e basi di $T^{-1}(W)$ e di $T(W)$.

(ii) Si consideri l'applicazione lineare $T^2|_{T^{-1}(W)} : T^{-1}(W) \rightarrow T(W)$. Determinarne la matrice rispetto alle basi di $T^{-1}(W)$ e di $T(W)$ ottenute in (i).

Soluzione. (i) Determiniamo una base di W , risolvendo il corrispondente $SLO(2, 4, \mathbf{R})$. Si ottiene ad esempio la base

$$\underline{w}_1 = (0, -1, 1, 0), \quad \underline{w}_2 = (1, 0, 0, 1).$$

Risulta:

$\underline{v} = \mathbf{E}\underline{x} \in T^{-1}(W) \iff T(\underline{v}) \in W \iff$ i vettori $\underline{w}_1, \underline{w}_2, T(\underline{v})$ sono linearmente dipendenti.

Poiché $T(\underline{v}) = (x_1 + x_4, x_2, x_1, x_2)$, si conclude che $T^{-1}(W)$ ha equazioni cartesiane date da

$$rg \begin{pmatrix} x_1 + x_4 & 0 & 1 \\ x_2 & -1 & 0 \\ x_1 & 1 & 0 \\ x_2 & 0 & 1 \end{pmatrix} = 2 \text{ cioè } \begin{cases} \begin{vmatrix} x_1 + x_4 & 0 & 1 \\ x_2 & -1 & 0 \\ x_1 & 1 & 0 \end{vmatrix} = 0 \\ \begin{vmatrix} x_2 & -1 & 0 \\ x_1 & 1 & 0 \\ x_2 & 0 & 1 \end{vmatrix} = 0. \end{cases}$$

Sviluppando i due determinanti, si ottiene il SLO

$$\begin{cases} x_1 - x_2 + x_4 = 0 \\ x_1 + x_2 = 0. \end{cases}$$

Risolvendo tale SLO si ottiene ad esempio la seguente base di $T^{-1}(W)$:

$$\underline{u}_1 = (0, 0, 1, 0), \quad \underline{u}_2 = (-\frac{1}{2}, \frac{1}{2}, 0, 1)$$

Calcoliamo ora equazioni cartesiane e base di $T(W)$. Tale spazio vettoriale è generato da $T(\underline{w}_1), T(\underline{w}_2)$. I due vettori

$$T(\underline{w}_1) = (0, -1, 0, -1), \quad T(\underline{w}_2) = (2, 0, 1, 0)$$

sono linearmente indipendenti e quindi sono una base di $T(W)$. Per ottenere equazioni cartesiane di $T(W)$ basta imporre

$$rg \begin{pmatrix} x_1 & 0 & 2 \\ x_2 & -1 & 0 \\ x_3 & 0 & 1 \\ x_4 & -1 & 0 \end{pmatrix} = 2 \text{ cioè } \begin{cases} \begin{vmatrix} x_1 & 0 & 2 \\ x_2 & -1 & 0 \\ x_3 & 0 & 1 \end{vmatrix} = 0 \\ \begin{vmatrix} x_1 & 0 & 2 \\ x_2 & -1 & 0 \\ x_4 & -1 & 0 \end{vmatrix} = 0. \end{cases}$$

Sviluppando i due determinanti, si ottiene il SLO

$$\begin{cases} x_1 - 2x_2 = 0 \\ x_2 - x_4 = 0. \end{cases}$$

(ii) È evidente che, per ogni $\underline{u} \in T^{-1}(W)$, $T(\underline{u}) \in W$ e quindi $T^2(\underline{u}) \in T(W)$. Pertanto $T^2|_{T^{-1}(W)}$ è un'applicazione lineare da $T^{-1}(W)$ a $T(W)$.

Disponiamo già di una base di $T^{-1}(W)$ [cioè $\{\underline{u}_1, \underline{u}_2\}$] e di una base di $T(W)$ [cioè $\{\underline{v}_1 := T(\underline{w}_1), \underline{v}_2 := T(\underline{w}_2)\}$]. Si tratta allora di calcolare $T^2(\underline{u}_1)$ e $T^2(\underline{u}_2)$, esprimendoli in base $\{\underline{v}_1, \underline{v}_2\}$.

T^2 ha, in base \mathbf{E} , matrice

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

Allora

$$T^2(\underline{u}_1) = \mathbf{E} A^2 \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} = (0, 0, 0, 0) = \underline{0}, \quad T^2(\underline{u}_2) = \mathbf{E} A^2 \begin{pmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ 0 \\ 1 \end{pmatrix} = (1, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}).$$

Ovviamente $T^2(\underline{u}_1) = \underline{0} = 0\underline{v}_1 + 0\underline{v}_2$. Determiniamo ora $a, b \in \mathbf{R}$ tali che $T^2(\underline{u}_2) = (1, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}) = a\underline{v}_1 + b\underline{v}_2$. Deve risultare:

$$(1, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}) = (0, -a, 0, -a) + (2b, 0, b, 0) = (2b, -a, b, -a).$$

Ne segue: $b = \frac{1}{2}$, $a = -\frac{1}{2}$ e pertanto $T^2(\underline{u}_2) = -\frac{1}{2}\underline{v}_1 + \frac{1}{2}\underline{v}_2$. Si conclude che l'applicazione lineare $T^2|_{T^{-1}(W)}$ ha, rispetto alle basi indicate, matrice

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

* * *