Università del Salento
Dipartimento di Matematica e Fisica
"*E. De Giorgi*"

Rocco Chirivì
Raffaele Vitolo

Dispense per il corso di

GEOMETRIA ED ALGEBRA

Versione del 19 marzo 2013

ANNO ACCADEMICO 2012-2013

Indirizzo degli autori.
Rocco Chirivì, Raffaele Vitolo,
Università del Salento, Dipartimento di Matematica e Fisica "*E. De Giorgi*",
via per Arnesano, 73100 Lecce
`rocco.chirivi@unisalento.it`
`raffaele.vitolo@unisalento.it`

INDICE

INTRODUZIONE

La geometria ... è evidentemente una scienza naturale:
possiamo infatti considerarla come la più antica parte della fisica.
A. Einstein

Queste note sono state sviluppate a partire dalle *Note di Geometria ed Algebra*, di G. De Cecco ed R. Vitolo, a seguito di una revisione del programma di Geometria ed Algebra per i corsi di laurea in Ingegneria dell'Università del Salento dove gli autori sono titolari della materia.

Le dispense costituiscono una guida alla preparazione dell'esame di 'Geometria ed Algebra'. Per ulteriori approfondimenti suggeriamo di consultare i testi consigliati.

Gli argomenti proposti nel corso saranno ampiamente applicati durante i corsi di tipo ingegneristico degli anni successivi. Per questo motivo è bene conseguire una buona preparazione in questa materia, come nelle altre materie di base: ciò consentirà di affrontare agevolmente i problemi e le tecniche trattate nei corsi applicativi.

Ringraziamenti. Ringraziamo tutti gli Studenti che hanno contribuito o contribuiranno a migliorare questo testo correggendone molti errori e sviste.

Queste note sono state scritte in LaTeX2e con l'estensione `amsmath` della American Mathematical Society

Rocco Chirivì
Raffaele Vitolo

CAPITOLO 1

STRUTTURE ALGEBRICHE

Le strutture algebriche svolgono nella matematica un ruolo importante, mettendo in evidenza la struttura comune a rami diversi. Le analogie non sono accidentali, ma fanno parte dell'essenza della matematica, che è un'attività umana, non un programma per computer. Si può anche dire che la matematica è lo studio delle analogie tra le analogie. Da qui l'uso di strutture algebriche per il funzionamento di dispositivi 'analogici'.

Nel seguito, gli elementi di teoria degli insiemi si riterranno già acquisiti in corsi precedenti.

1.1 Legge di composizione

Definizione 1.1. *Sia A un insieme. Una legge di composizione in A è un'applicazione definita in $A \times A$ e a valori in A*

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

Una legge di composizione fa corrispondere ad ogni coppia ordinata (a, b) di elementi di A un elemento $c = f(a, b)$, che spesso per semplicità di scrittura scriveremo anche $c = a * b$.

Esempi ed esercizi.

- Le usuali addizione, sottrazione, moltiplicazione sono leggi di composizione in $\mathbb{Z}$.
- La sottrazione non è legge di composizione in $\mathbb{N}$.
- La divisione è una legge di composizione in $\mathbb{Z}$?

La proprietà associativa per l'addizione e quella per la moltiplicazione in $\mathbb{R}$ si scrivono

$$a + (b + c) = (a + b) + c \qquad a \cdot (b \cdot c) = (a \cdot b) \cdot c\,,$$

che possono essere riassunte da

$$a * (b * c) = (a * b) * c\,,$$

dove $*$ è $+$ o $\cdot$.

Esempi ed esercizi.

- La sottrazione non è una legge di composizione associativa poiché
$$a-(b-c) \neq (a-b)-c.$$
- Se $a*b = a^b$, l'operazione $*$ è associativa?
- Considerata in $\mathbb{R}$ la seguente legge di composizione interna
$$a*b = (a+b)/2\,,$$
vedere se $*$ è associativa e calcolare $a^n = a*\ldots*a$ (n volte).

Come è noto $\forall a \in \mathbb{R}$
$$a+0 = 0+a = a \qquad a\cdot 1 = 1\cdot a = a\,,$$
quindi sinteticamente possiamo scrivere che
$$\forall a \qquad a*u = a\,,$$
dove $u=0$ se $*$ è l'addizione, $u=1$ se $*$ è la moltiplicazione.

1.2 Gruppo

Un gruppo $(G,*)$ è un insieme G con una legge di composizione $*$ in G tale che

1. Per ogni terna di elementi $a,b,c \in G$ si ha
$$a*(b*c) = (a*b)*c\,, \qquad \underline{\text{proprietà associativa.}}$$
2. Esiste un elemento $u \in G$ tale che $\forall a \in G$ si ha
$$a*u = u*a = a\,, \qquad \underline{\text{esistenza dell'elemento neutro.}}$$
3. Ogni elemento $a \in G$ possiede un elemento $a' \in G$ tale che
$$a*a' = a'*a = u\,, \qquad \underline{\text{esistenza dell'elemento simmetrico.}}$$
Si dimostra che u ed a' sono unici.
Se oltre agli assiomi (1), (2), (3) vale l'assioma
4. $\forall a,b \in G$
$$a*b = b*a\,, \qquad \underline{\text{proprietà commutativa,}}$$

allora il gruppo si dice *commutativo* o *abeliano*[1].

[1]dal nome del matematico norvegese Niels Abel (1802–1829)

Si dice che $(G', *)$ è un *sottogruppo* di $(G, *)$ se $G' \subset G$ è un gruppo rispetto alla stessa operazione in G.

Esempi ed esercizi.

- $(\mathbb{Z}, +)$, $(\mathbb{Q}, +)$ sono gruppi e $(\mathbb{Z}, +)$ è un sottogruppo di $(\mathbb{Q}, +)$;
- $(\mathbb{Z}, \cdot)$, $(\mathbb{Q}, \cdot)$ non sono gruppi;
- $(\mathbb{Q}^*, \cdot)$, dove $\mathbb{Q}^* = \mathbb{Q} \smallsetminus \{0\}$, è un gruppo.
- Sia $G = \{1, -1, i, -i\} \subset \mathbb{C}$. Provare che $(G, \cdot)$ è un gruppo, dove $\cdot$ è l'usuale moltiplicazione tra i numeri complessi.
- $(\mathbb{R}^2, +)$ è un gruppo, dove
$$(a_1, a_2) + (b_1, b_2) = (a_1 + b_1, a_2 + b_2)\,.$$
Si osservi che al primo membro $+$ è definita in $\mathbb{R}^2$, mentre al secondo membro in $\mathbb{R}$.
- L'insieme dei movimenti del piano (o dello spazio) forma un gruppo rispetto alla composizione delle applicazioni.

1.3 Anello

Assiomatizzando le proprietà dell'insieme numerico $\mathbb{Z}$ si giunge alla definizione astratta di anello.

Un *anello* $(A, +, \cdot)$ è un insieme (non vuoto) A di elementi con due leggi di composizione interna indicate con $+$ e $\cdot$ tali che

1. $(A, +)$ è un gruppo abeliano;
2. la legge $\cdot$ è associativa;
3. per ogni $a, b, c \in A$ valgono le uguaglianze $a\cdot(b+c) = a\cdot b + a\cdot c$, $(b+c)\cdot a = b\cdot a + c\cdot a$ (proprietà distributiva).

L'anello A è detto *commutativo* se $\cdot$ è commutativa, è detto *unitario* se $\cdot$ possiede un elemento neutro.

Esempi.

- $(\mathbb{Z}, +, \cdot)$ è un anello.
- Un *polinomio* $p(x)$ a coefficienti in $\mathbb{C}$ è un'espressione del tipo

$$p(x) = a_n x^n + \cdots + a_1 x + a_0, \qquad a_i \in \mathbb{C}, \quad n \in \mathbb{N}.$$

Se $a_n \neq 0$, si dice che $p(x)$ ha *grado* n. In particolare, un polinomio di grado 0 è una costante non nulla; il grado del polinomio nullo (che ha tutti i coefficienti nulli) non è definito, oppure, per convenzione, si pone uguale a $-\infty$.

Indichiamo con $\mathbb{C}[x]$ l'insieme di tutti i polinomi a coefficienti in $\mathbb{C}$. Se $p(x) = \sum_{i=0}^{n} a_i x^i$ e $q(x) = \sum_{j=0}^{m} b_j x^j$ sono due elementi di $\mathbb{C}[x]$, poniamo, per $m \geq n$,

$$p(x) + q(x) = \sum_{h=0}^{m} (a_h + b_h) x^h,$$
$$p(x) \cdot q(x) = \sum_{h=0}^{n+m} \Big(\sum_{i+j=h} a_i b_j \Big) x^h.$$

Si vede facilmente che $(\mathbb{C}[x], +, \cdot)$ è un anello.

Osservazione 1.1. *Si dice che $x = \alpha$ è una soluzione di* molteplicità algebrica k *dell'equazione algebrica $p(x) = 0$ se il polinomio $p(x)$ è divisibile per $(x - \alpha)^k$ ma non è divisibile per $(x - \alpha)^{k+1}$, cioè k è il massimo esponente per cui*

$$p(x) = (x - \alpha)^k q(x),$$

quindi $q(\alpha) \neq 0$.

1.4 Campo

Un *campo* $(\mathbb{K}, +, \cdot)$ è un insieme $\mathbb{K}$ (non vuoto) con due leggi di composizione interna indicate con $+$ e $\cdot$ tali che

1. $(\mathbb{K}, +)$ è un gruppo abeliano (il cui elemento neutro indichiamo con 0).
2. $(\mathbb{K}^*, \cdot)$ è un gruppo abeliano (dove $\mathbb{K}^* = \mathbb{K} \setminus \{0\}$).
3. $\forall a, b, c \in \mathbb{K}$ vale la proprietà distributiva di $\cdot$ rispetto a $+$:

$$a \cdot (b + c) = a \cdot b + a \cdot c\,, \qquad (b + c) \cdot a = b \cdot a + c \cdot a\,.$$

Dunque $(\mathbb{K}, +, \cdot)$ è anche un anello commutativo.

Esempi ed esercizi.

- Sono campi $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$. Il campo $\mathbb{Q}$ è quello caratteristico dell'Informatica, poiché i numeri reali che si considerano hanno sempre un numero *finito* di cifre.
- $\mathbb{C}$ è molto importante, essendo algebricamente chiuso, cioè ogni polinomio (non costante) a coefficienti in $\mathbb{C}$ ammette una radice in $\mathbb{C}$ (e di conseguenza tutte le sue radici appartengono a $\mathbb{C}$). È questo il contenuto del Teorema Fondamentale dell'Algebra (dovuto a Gauss)
- Se $\mathbb{K}$ è un campo, l'insieme dei polinomi a coefficienti in $\mathbb{K}$ è un anello, indicato con $\mathbb{K}[x]$, se denotiamo con x l'indeterminata.

CAPITOLO 2

SPAZI VETTORIALI

L'Algebra Lineare ha caratteristiche più astratte rispetto alla Geometria Analitica. Essa, nata alla fine del 1700 per lo studio dei sistemi di equazioni lineari, è poi diventata una branca autonoma della Matematica, che trova numerose applicazioni anche in Ingegneria.

2.1 Spazi vettoriali

Questa sezione è dedicata alla definizione di spazio vettoriale.

Definizione 2.1. *Si definisce* spazio vettoriale *sul campo* $\mathbb{K}$ *una terna* $(V, +, \cdot)$*, dove* $V \neq \varnothing$ *è un insieme e*

$$+ \colon V \times V \to V, \qquad \cdot \colon \mathbb{K} \times V \to V,$$

sono due applicazioni che soddisfano le seguenti proprietà.

1. *La coppia* $(V, +)$ *è un gruppo abeliano, ossia per ogni* $v, w, z \in V$
   - *(a)* $(v + w) + z = v + (w + z)$
   - *(b)* $\exists\, 0$ *tale che* $v + 0 = v$
   - *(c)* $\exists\, (-v)$ *tale che* $v + (-v) = 0$
   - *(d)* $v + w = w + v$
2. $\lambda(\mu v) = (\lambda\mu)v$ *per ogni* $v \in V$*,* $\lambda, \mu \in \mathbb{K}$*.*
3. $(\lambda + \mu)v = \lambda v + \mu v$ *per ogni* $v \in V$*,* $\lambda, \mu \in \mathbb{K}$*.*
4. $\lambda(v + w) = \lambda v + \lambda w$ *per ogni* $v, w \in V$*,* $\lambda \in \mathbb{K}$*.*
5. $1v = v$ *per ogni* $v \in V$*.*

Gli elementi di V sono detti *vettori* e quelli di $\mathbb{K}$ *scalari*. In ciò che segue si useranno lettere latine v, w, z, ... per i vettori e lettere greche λ, μ, ν, ... per gli scalari.

Con abuso di notazione si scriverà V al posto di $(V, +, \cdot)$. L'elemento $0_V \overset{\text{def}}{=} 0$, elemento neutro di $(V, +)$, si chiama *vettore nullo* e si vede facilmente che $0v = 0$ per ogni $v \in V$.

Esempi ed esercizi.

- L'insieme $\mathbb{K}^n \stackrel{\text{def}}{=} \{(x_1, \ldots, x_n) \mid x_i \in \mathbb{K}, i = 1, \ldots, n\}$ è uno spazio vettoriale rispetto alle operazioni di seguito definite. Siano $(x_1, \ldots, x_n)$, $(y_1, \ldots, y_n) \in \mathbb{K}^n$, $\lambda \in \mathbb{K}$. Allora si pone

  $$(x_1, \ldots, x_n) + (y_1, \ldots, y_n) \stackrel{\text{def}}{=} (x_1 + y_1, \ldots, x_n + y_n),$$
  $$\lambda \cdot (x_1, \ldots, x_n) \stackrel{\text{def}}{=} (\lambda x_1, \ldots, \lambda x_n).$$

  Si noti che le operazioni al primo membro sono definite tramite le operazioni componente per componente al secondo membro, e che queste ultime sono le operazioni di $\mathbb{K}$.

  Si verifica facilmente che le operazioni sopra definite assegnano una struttura di spazio vettoriale su $\mathbb{K}^n$

- Esempi notevoli nella classe dell'esempio precedente sono $\mathbb{R}^n$ e $\mathbb{C}^n$, ed in particolare $\mathbb{R}(= \mathbb{R}^1)$, $\mathbb{R}^2$, $\mathbb{R}^3$.

- Sia $\mathbb{K}$ un campo. Si chiama *matrice ad m righe ed n colonne* a coefficienti in $\mathbb{K}$ una tabella del tipo seguente:

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

  dove gli elementi appartengono a $\mathbb{K}$.

  L'insieme
  $$\mathbb{K}^{m,n} = \{A \mid A \text{ matrice di tipo } m, n \text{ a coeff. in } \mathbb{K}\}$$
  delle matrici di tipo m, n a coefficienti in $\mathbb{K}$ è uno spazio vettoriale rispetto alle seguenti operazioni. Se $\lambda \in \mathbb{K}$ e $A \in \mathbb{K}^{m,n}$, la matrice λA, *moltiplicazione di A per lo scalare λ*, è la matrice
  $$\lambda A = (\lambda a_{ij}) \quad \forall i, j$$
  Date A, $B \in \mathbb{K}^{m,n}$, la matrice *somma* $C = A + B$ è per definizione $C = (c_{ij}) \in \mathbb{K}^{m,n}$ con
  $$c_{ij} = a_{ij} + b_{ij}\,.$$
  La matrice O avente tutti gli elementi 0 è la matrice *nulla*, e soddisfa
  $$A + O = A \qquad \forall A\,,$$
  e l'*opposta* di A è la matrice $A' = -A$, dove $a'_{ij} = -a_{ij}$ $\forall i, j$. Quindi $(\mathbb{K}^{m,n}, +)$ è un gruppo abeliano, e si può facilmente dimostrare che $(\mathbb{K}^{m,n}, +, \cdot)$ è uno spazio vettoriale.

- L'insieme $\mathbb{K}[t]$ dei polinomi a coefficienti in $\mathbb{K}$ è uno spazio vettoriale rispetto alle operazioni di somma di polinomi e moltiplicazione di un polinomio per uno scalare.

- Si consideri lo spazio euclideo S_3, ovvero l'insieme di 'punti' dotato degli assiomi della geometria euclidea dello spazio tradizionale (estensione della geometria euclidea del piano tradizionale). In S_3 è possibile definire il concetto di 'segmento orientato' AB, di primo estremo $A \in S_3$ e secondo estremo $B \in S_3$. Due segmenti orientati sono equivalenti se giacciono su rette parallele (hanno la stessa direzione), se hanno la stessa lunghezza (hanno lo stesso modulo) ed hanno la stessa orientazione. Dato un segmento orientato, l'insieme di tutti i segmenti orientati ad esso equivalenti costituisce un *vettore*

  $$v = \vec{AB} = \{PQ \mid PQ \text{ equivalente a } AB\}.$$

  L'insieme dei vettori dello spazio euclideo tradizionale si indica con V_3.

  Si può definire un'operazione di somma tra vettori usando la *regola del parallelogramma* ed un'operazione di moltiplicazione per scalari in $\mathbb{R}$. Si verifica che queste operazioni soddisfano le proprietà richieste a fare dell'insieme V_3 uno spazio vettoriale.

2.2 Sottospazi vettoriali

Definizione 2.2. *Sia $(V, +, \cdot)$ uno spazio vettoriale. Un sottoinsieme $W \subset V$ si dice* sottospazio vettoriale *di V se $(W, +, \cdot)$ è uno spazio vettoriale rispetto alle operazioni di V ristrette a W.*

Si prova facilmente che

Proposizione 2.1. *Sia $(V, +, \cdot)$ uno spazio vettoriale, e $W \subset V$. Allora si equivalgono*

1. *W sottospazio vettoriale di V;*
2. *per ogni $u, v \in W$ e $\lambda \in \mathbb{K}$ si ha $u + v \in W$ e $\lambda u \in W$.*

Ne segue che *tutti i sottospazi di V contengono il vettore nullo* $0 \in V$. Ma questa condizione *non è sufficiente* per concludere che un insieme è un sottospazio vettoriale! Si consideri, come controesempio, il seguente insieme: $\mathcal{P} = \{(x, y) \in \mathbb{R}^2 \mid y = x^2\}$.

Esempi ed esercizi.

- Sia $W \stackrel{\text{def}}{=} \{x \in \mathbb{R}^n \mid x_1 = 0\}$. Dimostriamo che W è un sottospazio vettoriale di $\mathbb{R}^n$. Siano $x, y \in W$ e $\lambda \in \mathbb{R}$. Allora

  $$x + y = (x_1 + y_1, \ldots, x_n + y_n), \quad \lambda x = (\lambda x_1, \ldots, \lambda x_n),$$

con

$$x_1 + y_1 = 0 + 0 = 0, \qquad \lambda x_1 = \lambda 0 = 0.$$

quindi W è un sottospazio vettoriale di $\mathbb{R}^n$.

- Sia $\mathbb{R}_n[t] \subset \mathbb{R}[t]$ il sottoinsieme dei polinomi di grado $\leq n$. Allora $\mathbb{R}_n[t]$ è un sottospazio di $\mathbb{R}[t]$. Infatti la somma di due polinomi di grado minore o uguale ad n ha ancora grado minore uguale ad n, ed altrettanto per la moltiplicazione di un polinomio per uno scalare.

- Sia $\bar{\mathbb{R}}_n[t] \subset \mathbb{R}[t]$ il sottoinsieme dei polinomi di grado *uguale* ad n. Allora $\bar{\mathbb{R}}_n[t]$ *non è* un sottospazio vettoriale di $\mathbb{R}[t]$. Infatti, può succedere che sommando due polinomi di grado uguale ad n si ottenga un polinomio di grado minore di n. Del resto, $\bar{\mathbb{R}}_n[t]$ non contiene il polinomio nullo, quindi non può essere uno spazio vettoriale rispetto alle operazioni indotte da $\mathbb{R}[t]$ per restrizione.

- Sia $Z_1 \stackrel{\text{def}}{=} \{x \in \mathbb{R}^2 \mid x_1 + x_2 = 1\}$. Dimostrare che Z_1 non è un sottospazio vettoriale di $\mathbb{R}^2$.

- Sia $Z_2 \stackrel{\text{def}}{=} \{x \in \mathbb{R}^2 \mid x_1^2 + x_2^2 = 1\}$. Dimostrare che Z_2 non è un sottospazio vettoriale di $\mathbb{R}^2$.

- Sia $U \stackrel{\text{def}}{=} \{x \in \mathbb{R}^3 \mid x_1^2 + x_2^2 = 0\}$; U è un sottospazio vettoriale di $\mathbb{R}^3$? Si osservi che i sottospazi vettoriali di $\mathbb{R}^3$ sono solo i seguenti: l'*origine*, le *rette per l'origine* ed i *piani per l'origine*.

Sia V uno spazio vettoriale, e siano S, T due suoi sottospazi. Si dimostra facilmente che $S \cap T$ è un sottospazio vettoriale di V, mentre, in generale, $S \cup T$ *non è* un sottospazio vettoriale di V. Posto

$$S + T \stackrel{\text{def}}{=} \{x + y \mid x \in S, y \in T\} \supset S \cup T,$$

è immediato verificare che $S + T$ è un sottospazio vettoriale di V, detto *somma* di S e di T. Inoltre:

Proposizione 2.2. *Siano S, T sottospazi di uno spazio vettoriale V. Allora:*

$S + T$ è il più piccolo sottospazio di V contenente $S \cup T$;

$S \cap T$ è il più grande sottospazio di V contenuto in S e T.

DIMOSTRAZIONE. Sia W un sottospazio di V contenente $S \cup T$. Allora W contiene gli elementi del tipo $s + t$, dove $s \in S$ e $t \in T$, dunque $S + T \in W$. L'altra affermazione si dimostra in modo analogo. $\boxed{QED}$

Posto $W \stackrel{\text{def}}{=} S + T$, se in W la decomposizione di ogni vettore come somma di due elementi in S e T è univocamente determinata, allora W si dice *somma diretta* e si scrive $W = S \oplus T$. Se $V = W = S \oplus T$, allora i due spazi S e T si dicono *supplementari*.

Proposizione 2.3. *Sia V uno spazio vettoriale, e siano S, T due suoi sottospazi. Allora, posto $W \stackrel{def}{=} S + T$, si equivalgono*

1. *$W = S \oplus T$;*
2. *$W = S + T$ e $S \cap T = \{0\}$.*

DIMOSTRAZIONE. Sia $W = S \oplus T$. Allora se $w \in W$ esiste un'unica coppia $(\vec{s}, \vec{t}) \in S \times T$ tale che $w = \vec{s} + \vec{t}$. In particolare, se $w \in S \cap T$ allora si può scrivere w nei seguenti modi: $w = w + 0 = 0 + w$; ma queste scritture devono coincidere per ipotesi, dunque $w = 0$.

Viceversa, siano $W = S + T$ e $S \cap T = \{0\}$, e supponiamo che $w \in W$ si scriva come $w = s + t = s' + t'$. Sottraendo membro a membro si ottiene $s - s' = t' - t \in S \cap T = \{0\}$, dunque $s = s'$ e $t' = t$. $\boxed{QED}$

Il concetto di somma e di somma diretta si estende a più sottospazi:

$$S_1 + \cdots + S_h \stackrel{\text{def}}{=} \{v_1 + \cdots + v_h \mid v_i \in S_i, i = 1, \ldots, h\}$$
$$S_1 \oplus \cdots \oplus S_h \stackrel{\text{def}}{=} \{v = v_1 + \cdots + v_h \mid v_i \in S_i, i = 1, \ldots, h, \text{in modo unico}\}$$

Si osservi che

$$W = S_1 \oplus \cdots \oplus S_h \quad \Rightarrow \quad S_i \cap S_j = \{0\} \quad i \neq j,$$

ma non vale il viceversa. Si considerino, per un controesempio, tre rette distinte in $\mathbf{V}_2$ passanti per l'origine. Al contrario, scegliendo in $\mathbf{V}_3$ tre rette r, s, t per l'origine non complanari si ha $\mathbf{V}_3 = r \oplus s \oplus t$.

Nota. Se x_0 è un fissato vettore di V ed S un suo sottospazio vettoriale, poniamo

$$x_0 + S \stackrel{\text{def}}{=} \{\, x = x_0 + s \mid \vec{s} \in S \,\}.$$

Si osservi che, per $x_0 \neq 0$, $x_0 + S$ non è un sottospazio vettoriale di V se $x_0 \notin S$; se $x_0 \in S$ allora $x_0 + S = S$.

2.3 Dipendenza ed indipendenza lineare

Definizione 2.3. *Sia V uno spazio vettoriale su $\mathbb{K}$ e $X \subset V$ un sottoinsieme non vuoto. Si dice* combinazione lineare *di vettori di X un'espressione del tipo*

$$v = \lambda_1 v_1 + \cdots + \lambda_n v_n = \sum_{i=1}^{n} \lambda_i v_i,$$

dove $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$ sono i coefficienti *della combinazione lineare e $v_1, \ldots, v_n \in X$ sono vettori in X.*

Si noti che una combinazione lineare è una somma di un numero *finito* di vettori di X, anche se X è infinito. Una somma infinita presenta problemi di convergenza, studiati in Analisi Matematica.

Definizione 2.4. *Nelle ipotesi precedenti, l'insieme X si dice*

1. indipendente *se il vettore 0 è combinazione lineare di X in modo* unico, *e quindi con coefficienti tutti nulli, ossia se per ogni sottoinsieme finito $\{v_1, \dots, v_n\} \subset X$ vale*
$$\lambda_1 v_1 + \cdots + \lambda_n v_n = 0 \quad \Rightarrow \quad \lambda_j = 0 \quad j = 1, \dots, n;$$
2. dipendente *se non è indipendente.*

È immediato verificare che un insieme è dipendente se almeno uno dei vettori dell'insieme è combinazione lineare dei restanti.

Valgono alcune proprietà dalla dimostrazione immediata:

1. se $0 \in X$, X è dipendente;
2. se X è indipendente, allora ogni sottoinsieme non vuoto $X' \subset X$ è indipendente;
3. se un sottoinsieme non vuoto $X' \subset X$ è dipendente allora X è dipendente.

2.4 Basi e dimensione

Proposizione 2.4. *Sia $X \subset V$ un sottoinsieme non vuoto. Allora l'insieme*
$$\mathcal{L}(X) \stackrel{def}{=} \{x \in V \mid x \text{ è comb. lin. di elementi di } X\}.$$
è un sottospazio vettoriale di V, ed è il più piccolo sottospazio vettoriale di V che contiene X.

Dimostrazione. Infatti, dati v, $w \in \mathcal{L}(X)$ si avrà
$$v = \lambda_1 v_1 + \cdots + \lambda_n v_n, \quad w = \mu_1 w_1 + \cdots + \mu_n w_n$$
dove $\lambda_1, \dots, \lambda_n, \mu_1, \dots, \mu_n \in \mathbb{K}$ e $v_1, \dots, v_n, w_1, \dots, w_n \in V$. Quindi la somma $v + w$ è un elemento di $\mathcal{L}(X)$ in quanto combinazione lineare di elementi di $\mathcal{L}(X)$:
$$v + w = \lambda_1 v_1 + \cdots + \lambda_n v_n + \mu_1 w_1 + \cdots + \mu_n w_n.$$
Analogamente per il prodotto per scalari.

Infine, ogni sottospazio W di V tale che $X \in W$ è tale che una combinazione lineare v di elementi di X sia in W in quanto v è anche una combinazione lineare di elementi di W. Dunque $\mathcal{L}(X) \subset W$. QED

Definizione 2.5. *Nelle ipotesi precedenti, si dice* sottospazio vettoriale generato da *X l'insieme $\mathcal{L}(X)$.*

È facile dimostrare che, se S, T sono sottospazi di V, allora

$$S + T = \mathcal{L}(S \cup T).$$

Definizione 2.6. *Sia V uno spazio vettoriale, $G \in V$ un sottoinsieme non vuoto. Allora si dice che G è un insieme di* generatori *per V se $\mathcal{L}(G) = V$. Se G è un insieme finito, V si dice* finitamente generato*.*

Definizione 2.7. *Sia V uno spazio vettoriale. Un sottoinsieme $\mathcal{B} \subset V$ si dice* base *di V se $\mathcal{B}$ è un insieme di* generatori indipendenti*.*

Esempi ed esercizi.

- Ogni spazio vettoriale V è generato da se stesso: $\mathcal{L}(V) = V$.

- L'insieme $\mathbb{K}^n$ è finitamente generato dall'insieme

  $$\mathcal{C} = \{e_1 \overset{\text{def}}{=} (1, 0, \dots, 0), \dots, e_n \overset{\text{def}}{=} (0, \dots, 0, 1)\}.$$

  Si può facilmente dimostrare che $\mathcal{C}$ è un insieme di vettori indipendenti, dunque è una base.

- L'insieme $\mathbb{K}^{m,n}$ è finitamente generato dalle matrici $\mathcal{C} = \{E_{ij}\}$, dove

  $$E_{ij} = (e_{hk}), \quad e_{hk} = 1 \quad \text{se e solo se} \quad h = i, k = j, \quad e_{hk} = 0 \quad \text{altrimenti}.$$

  Si può facilmente dimostrare che $\mathcal{C}$ è un insieme di vettori indipendenti, dunque è una base.

- L'insieme $\mathbb{R}[t]$ è (non finitamente!) generato dai polinomi $\mathcal{C} = \{1, t, t^2, t^3, \dots\}$. Tali polinomi sono indipendenti per costruzione (principio di indentità dei polinomi), qundi $\mathcal{C}$ è una base.

- Lo spazio dei vettori euclidei V_3 è generato da tre vettori ortogonali di lunghezza 1, indicati con $\{\vec{\imath}, \vec{\jmath}, \vec{k}\}$. I vettori sono indipendenti perché non sono complanari, dunque costituiscono una base.

L'importanza di poter scegliere una base in uno spazio vettoriale è data dalla seguente proposizione.

Proposizione 2.5. *Sia V uno spazio vettoriale, e sia $\mathcal{B} \subset V$ una base. Allora ogni vettore $v \in V$ si rappresenta in modo unico come combinazione lineare di vettori di $\mathcal{B}$.*

Dimostrazione. Supponiamo che $\mathcal{B}$ sia una base finita, e poniamo $\mathcal{B} = \{v_1, \ldots, v_n\}$. $\mathcal{B}$ è un insieme di generatori, dunque v si esprime come combinazione lineare di v. Supponiamo che questo avvenga in due modi:

$$v = \lambda_1 v_1 + \cdots + \lambda_n v_n, \quad v = \mu_1 v_1 + \cdots + \mu_n v_n.$$

Sottraendo membro a membro si ha

$$0 = (\lambda_1 - \mu_1)v_1 + \cdots + (\lambda_n - \mu_n)v_n.$$

Dall'indipendenza di $\mathcal{B}$ segue $\lambda_1 - \mu_1 = 0, \ldots, \lambda_n - \mu_n = 0$.

Il caso in cui $\mathcal{B}$ è infinita si tratta in modo simile; la dimostrazione è lasciata al lettore. $\boxed{QED}$

Se $\mathcal{B} = \{v_1, \ldots, v_n\}$, allora per $v \in V$

$$v = \lambda_1 v_1 + \cdots + \lambda_n v_n,$$

con $(\lambda_1, \ldots, \lambda_n)$ n-pla univocamente individuata. Gli scalari $\lambda_1, \ldots, \lambda_n$ sono chiamati *coordinate* di x rispetto alla base $\mathcal{B}$. La precedente corrispondenza definisce una funzione biunivoca

$$c_{\mathcal{B}} \colon V \to \mathbb{K}^n, \quad v \mapsto (\lambda_1, \ldots, \lambda_n),$$

che dipende dalla scelta di una base $\mathcal{B}$.

L'obiettivo principale di questo paragrafo è dimostrare che ogni spazio vettoriale finitamente generato ammette una base, e che il numero degli elementi di due basi distinte di uno stesso spazio vettoriale è lo stesso. Tale numero è una caratteristica fondamentale di ogni spazio vettoriale finitamente generato.

Lemma 2.1. *Sia V uno spazio vettoriale. Sia $\{v_1, \ldots, v_n\} \subset V$ tale che $v_1 \neq 0$, v_2 non è combinazione lineare di $\{v_1\}$, v_3 non è combinazione lineare di $\{v_1, v_2\}$, ..., v_n non è combinazione lineare di $\{v_1, \ldots, v_{n-1}\}$. Allora $\{v_1, \ldots, v_n\}$ è indipendente.*

Dimostrazione. Infatti, $\lambda_1 v_1 + \cdots + \lambda_n v_n = 0$ implica $\lambda_n = 0$ poiché v_n non è combinazione lineare dei rimanenti vettori, ed iterando il ragionamento, $\lambda_{n-1} = 0, \ldots, \lambda_1 = 0$. $\boxed{QED}$

Proposizione 2.6 (Metodo degli scarti successivi)**.** *Sia V uno spazio vettoriale e $W \neq \{0\}$ un suo sottospazio finitamente generato: $W = \mathcal{L}(v_1, \ldots, v_n)$. Allora esiste un sottoinsieme $B = \{v_{i_1}, \ldots, v_{i_p}\} \subset \{v_1, \ldots, v_n\}$ tale che*

1. *B è un insieme di vettori indipendenti;*

2. *W è generato da B.*

Dimostrazione. La dimostrazione è costruttiva. Al primo passo, si ponga $v_{i_1} = v_{p_1}$ dove v_{p_1} è il primo vettore non nullo tra i vettori $v_1, \ldots, v_n$ (necessariamente esiste, poiché $W \neq \{0\}$).

Al secondo passo, si ponga $v_{i_2} = v_{p_2}$ dove v_{p_2} è il primo vettore tra $v_{p_1+1}, \ldots, v_n$ che non è combinazione lineare di v_{i_1}.

Al passo k, si prenda v_{i_k} come il primo vettore tra $v_{p_{k-1}+1}, \ldots, v_n$ che non è combinazione lineare di $v_{i_1}, \ldots, v_{i_{k-1}}$.

L'algoritmo termina poiché l'insieme di partenza è finito. Inoltre, l'algoritmo produce un insieme di vettori indipendenti per il lemma 2.1.

Infine, l'algoritmo produce un insieme di generatori; infatti, se $v \in W$, $v = \lambda_1 v_1 + \cdots + \lambda_n v_n$, per ogni vettore $v_k \notin \{v_{i_1}, \ldots, v_{i_p}\}$ risulta $v_k = \lambda_{i_1} v_{i_1} + \cdots + \lambda_{i_p} v_{i_p}$. Per questo motivo si può sostituire ogni vettore del tipo di v_k nell'espressione di v ottenendo così una combinazione lineare dei soli vettori in B. Si è dunque dimostrato che ogni combinazione lineare dei vettori $v_1, \ldots, v_n$ si può esprimere come una combinazione lineare dei soli vettori di B. $\boxed{QED}$

Dalla precedente proposizione deriva immediatamente l'esistenza di basi per ogni spazio vettoriale finitamente generato.

Corollario 2.1. *Sia V uno spazio vettoriale finitamente generato dall'insieme G. Allora V ammette una base $B \subset G$.*

Vale anche un risultato che va in direzione opposta al precedente.

Proposizione 2.7 (Completamento ad una base)**.** *Sia V uno spazio vettoriale finitamente generato. Sia $I \subset V$ un insieme finito di vettori indipendenti. Allora esiste una base B di V tale che $I \subset B$.*

Dimostrazione. Si consideri un insieme finito G di generatori per V. Se $I = \{v_1, \ldots, v_p\}$ e $G = \{w_1, \ldots, w_q\}$ si applichi il metodo degli scarti successivi all'insieme $I \cup G = \{v_1, \ldots, v_p, w_1, \ldots, w_q\}$. Si ottiene una base B di V che contiene i vettori di I perché $v_1 \neq 0$, v_2 non dipendente da v_1, ecc. in quanto I è un insieme di vettori indipendenti. $\boxed{QED}$

Si può dimostrare che

Teorema 2.1. *Sia $X = \{v_1, \ldots, v_n\}$ e $\{w_1, \ldots, w_m\} \subset \mathcal{L}(X)$. Allora, se $m > n$, i vettori $\{w_1, \ldots, w_m\}$ sono dipendenti.*

Dimostrazione. Si applica il metodo dell'induzione su n. Per $n = 1$ i vettori $\{w_1, \ldots, w_m\}$ sono proporzionali a v_1, quindi sono anche proporzionali tra loro, dunque sono dipendenti.

Assumiamo la tesi vera per $n - 1$. Si ponga

$$w_j = \alpha_{1j} v_1 + \cdots + \alpha_{nj} v_n, \quad j = 1, \ldots, m.$$

Se $w_1 = 0$ la tesi è vera per n. Se $w_1 \neq 0$ si può supporre $\alpha_{n1} \neq 0$. Si può eliminare v_n dai vettori $w_2, \dots, w_n$:

$$w_2 - \frac{\alpha_{n2}}{\alpha_{n1}} w_1, \dots, w_m - \frac{\alpha_{nm}}{\alpha_{n1}} \in \mathcal{L}(\{v_1, \dots, v_{n-1}\}).$$

Applicando l'ipotesi induttiva, i vettori $w_2 - \frac{\alpha_{n2}}{\alpha_{n1}} w_1, \dots, w_m - \frac{\alpha_{nm}}{\alpha_{n1}} w_1$ sono dipendenti. Esistono dunque scalari $\lambda_2, \dots, \lambda_m$ non tutti nulli tali che

$$\lambda_2\big(w_2 - \frac{\alpha_{n2}}{\alpha_{n1}} w_1\big) + \cdots + \lambda_m\big(w_m - \frac{\alpha_{nm}}{\alpha_{n1}} w_1\big) = 0.$$

Ponendo $\lambda_1 = -(\lambda_2 \frac{\alpha_{n2}}{\alpha_{n1}} + \cdots + \lambda_m \frac{\alpha_{nm}}{\alpha_{n1}})$ si ottiene la combinazione lineare

$$\lambda_1 w_1 + \lambda_2 w_2 + \cdots + \lambda_m w_m = 0,$$

con scalari $\lambda_1, \lambda_2, \dots, \lambda_m$ non tutti nulli. Quindi la tesi è vera per n. QED

Corollario 2.2. *Sia V uno spazio vettoriale finitamente generato. Allora tutte le sue basi hanno la stessa cardinalità finita.*

DIMOSTRAZIONE. Dimostriamo innanzitutto che la cardinalità di una base $\mathcal{B}$ è finita. Preso, infatti, un insieme di generatori G contenente m vettori, un qualunque sottoinsieme di $m+k$ vettori in $\mathcal{B}$ sarebbe indipendente, contraddicendo il precedente teorema. Dunque $\mathcal{B}$ ha cardinalità finita.

Ora, se $\mathcal{B}$ e $\mathcal{B}'$ sono due basi di V di n ed n' elementi, rispettivamente, allora il teorema 2.1 implica

$$\begin{aligned} \mathcal{B}' \subset \mathcal{L}(\mathcal{B}) \quad &\Rightarrow \quad n' \leq n, \\ \mathcal{B} \subset \mathcal{L}(\mathcal{B}') \quad &\Rightarrow \quad n \leq n', \end{aligned}$$

quindi $n = n' = \dim V$. QED

Si può dimostrare che tutti gli spazi vettoriali, anche non finitamente generati, hanno basi della stessa cardinalità; per far questo è necessario introdurre il concetto di cardinalità di un insieme infinito, che non verrà trattato in questo corso.

Definizione 2.8. *Se V è uno spazio vettoriale su $\mathbb{K}$, si definisce* dimensione *di V, indicata con* $\dim_{\mathbb{K}} V$ *o* $\dim V$*, la cardinalità di una sua base.*

Si pone $\dim\{0\} = 0$.

In particolare, uno spazio vettoriale ha

- *dimensione finita* se una sua base ha cardinalità finita;
- *dimensione infinita* se una sua base ha cardinalità infinita.

Proposizione 2.8. *Se* $\dim V = n$ *(ossia, se la cardinalità di una base di V è n), allora*

- *n è il <u>massimo</u> numero di vettori* linearmente indipendenti *in V;*
- *n è il <u>minimo</u> numero di* generatori *di V.*

DIMOSTRAZIONE. 1. Se ci fosse un insieme di vettori indipendenti con $m > n$ vettori, allora questo potrebbe essere completato ad una base, ma questa avrebbe un numero di vettori maggiore di n.

2. Se ci fosse un insieme di generatori con $m < n$ vettori allora il metodo degli scarti successivi fornirebbe su di esso una base, ma questa avrebbe un numero di vettori minore di n.

$\boxed{QED}$

Dalla precedente proposizione segue che ogni sottospazio di uno spazio vettoriale di dimensione finita è di dimensione finita. Viceversa, uno spazio vettoriale di dimensione infinita può avere sottospazi di dimensione finita o infinita.

Esempi ed esercizi.

- Si ha $\dim \mathbb{K}^n = n$. Infatti, se 0 è l'elemento neutro di $\mathbb{K}$ rispetto all'addizione e 1 è l'elemento neutro di $\mathbb{K}$ rispetto alla moltiplicazione, la base standard è
$$e_1 \stackrel{\text{def}}{=} (1, 0, \ldots, 0), \ldots, e_n \stackrel{\text{def}}{=} (0, \ldots, 0, 1).$$
Si osservi che $\mathbb{K}^0 \stackrel{\text{def}}{=} \{0\}$.

- Se V è uno spazio vettoriale di dimensione n, allora ogni suo sottospazio $S \subset V$ tale che $\dim S = n$ coincide con V. Infatti, una qualunque base di S è anche una base di V.

- Si ha $\dim_{\mathbb{C}} \mathbb{C} = 1$, una base è costituita da un qualunque suo elemento non nullo (ad esempio $e_1 \stackrel{\text{def}}{=} 1$).

  Invece, $\dim_{\mathbb{R}} \mathbb{C} = 2$, poiché per definizione $\mathbb{C} = \mathbb{R}^2$ e la somma ed il prodotto per scalari su $\mathbb{C}$ come spazio vettoriale su $\mathbb{R}$ sono quelle di $\mathbb{R}^2$. Una base può essere costituita da $e_1 \stackrel{\text{def}}{=} 1$ e $e_1^* \stackrel{\text{def}}{=} i$. Ricordando l'identificazione $1 \mapsto (1, 0)$, $i \mapsto (0, 1)$ si ottiene per ogni $z \in \mathbb{C}$
$$z = (a, b) \mapsto a e_1 + b e_1^*$$
in modo unico.

- Analogamente $\dim_{\mathbb{C}} \mathbb{C}^n = n$, $\dim_{\mathbb{R}} \mathbb{C}^n = 2n$. Se $\mathcal{B} = \{e_1, \ldots, e_n\}$ è una base di $\mathbb{C}^n$ su $\mathbb{C}$, allora $\mathcal{B}' = \{e_1, \ldots, e_n, e_1^*, \ldots, e_n^*\}$ è una base di $\mathbb{C}^n$ su $\mathbb{R}$, con $e_h^* \stackrel{\text{def}}{=} i e_h$, $h = 1, \ldots, n$.

- Si ha $\dim_{\mathbb{K}} \mathbb{K}^{m,n} = m \cdot n$. Una base è data dall'insieme $\mathcal{C} \stackrel{\text{def}}{=} \{E_{ij} \mid 1 \leq i \leq m, 1 \leq j \leq n\}$ dove E_{ij} è la matrice che ha 1 al posto (ij) e 0 altrove.

- Si ha $\dim \mathbb{R}_n[t] = n+1$, una base è $\{t^h\}_{0\leq h\leq n}$.
- Si ha $\dim \mathbb{R}[t] = +\infty$. Più precisamente, una base di $\mathbb{R}[t]$ ha la cardinalità di $\mathbb{N}$, ad esempio $\{t^h \mid h \in \mathbb{N}\}$. Si noti che $\mathbb{R}[t]$ ammette sia i sottospazi di dimensione finita $\mathbb{R}_n[t]$, sia sottospazi di dimensione infinita.
- Si provi che il sottoinsieme $\mathbb{R}[t^2] \stackrel{\text{def}}{=} \mathcal{L}(\{t^{2h}\}_{h\in\mathbb{N}})$ è un sottospazio di dimensione infinita di $\mathbb{R}[t]$.

2.5 Relazione di Grassmann

Teorema 2.2. *Sia V uno spazio vettoriale di dimensione finita, e siano S, T due suoi sottospazi. Allora*

$$\dim(S+T) + \dim(S\cap T) = \dim S + \dim T.$$

La formula precedente prende il nome di *relazione di Grassmann.*

Dimostrazione. Sia $\{u_1, \ldots, u_h\}$ una base di $S \cap T$. Essa si può completare ad una base $\{u_1, \ldots, u_h, v_1, \ldots, v_p\}$ di S e ad una base $\{u_1, \ldots, u_h, w_1, \ldots, w_q\}$ di T. Se si dimostra che

$$\mathcal{B} \stackrel{\text{def}}{=} \{u_1, \ldots, u_h, v_1, \ldots, v_p, w_1, \ldots, w_q\}$$

è una base di $S+T$, la relazione di Grassmann sarà provata (con un semplice calcolo delle dimensioni).

Si dimostra facilmente che $\mathcal{B}$ è un insieme di generatori per $S+T$.

Dimostriamo ora l'indipendenza di $\mathcal{B}$, ossia che se

$$a_1u_1 + \cdots + a_hu_h + b_1v_1 + \cdots + b_pv_p + c_1w_1 + \cdots + c_qw_q = 0 \tag{2.5.1}$$

allora $a_i = b_j = c_k = 0$. Scrivendo l'equazione precedente come

$$a_1u_1 + \cdots + a_hu_h + b_1v_1 + \cdots + b_pv_p = -c_1w_1 - \cdots - c_qw_q$$

si vede che $w \stackrel{\text{def}}{=} -c_1w_1 - \cdots - c_qw_q \in T \cap S$, quindi

$$w = \lambda_1u_1 + \cdots + \lambda_hu_h,$$

per opportuni λ_i. Segue

$$\lambda_1u_1 + \cdots + \lambda_hu_h + c_1w_1 + \cdots + c_qw_q = 0,$$

che implica $\lambda_i = c_j = 0$ poiché $\{u_1, \ldots, u_h, w_1, \ldots, w_q\}$ è una base. Quindi la (2.5.1) diventa

$$a_1u_1 + \cdots + a_hu_h + b_1v_1 + \cdots + b_pv_p = 0,$$

che implica $a_i = b_j = 0$, da cui la tesi. $\boxed{QED}$

Nel caso di somma diretta, poiché $S \cap T = \{0\}$, la relazione di Grassmann dà

$$\dim(S \oplus T) = \dim S + \dim T.$$

Viceversa, se $\dim(S+T) = \dim S + \dim T$ segue $\dim(S \cap T) = 0$ quindi $S + T = S \oplus T$.

Esercizio. Sia U il sottospazio di $\mathbb{R}^4$ generato da

$$u_1 = (1,1,1,1), \quad u_2 = (1,-1,1,-1), \quad u_3 = (1,3,1,3),$$

e W il sottospazio generato da

$$w_1 = (1,2,0,2), \quad w_2 = (1,2,1,2), \quad w_3 = (3,1,3,1).$$

Trovare $\dim(U \cap W)$, $\dim(U + W)$ e descrivere $U \cap W$ e $U + W$.

CAPITOLO 3

LE MATRICI

3.1 Definizioni

Sia $\mathbb{K}$ un campo. Si chiama *matrice ad m righe ed n colonne* a coefficienti in $\mathbb{K}$ una tabella del tipo seguente:

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

dove gli elementi appartengono a $\mathbb{K}$.

L'elemento generico di A è a_{ij}, cioè l'elemento che si trova sull'i-esima riga e j-esima colonna. In breve si scrive

$$A = (a_{ij}) \qquad i = 1, 2, \dots, m; \quad j = 1, 2, \dots, n.$$

Se $m \neq n$ la matrice si dice *rettangolare*, se $m = n$ si chiama *quadrata*, ed n è detto *ordine* di A. Se $m = 1$ la matrice si dice *vettore riga*, se $n = 1$ la matrice si chiama anche *vettore colonna*.

Indichiamo con $\mathbb{K}^{m,n}$ l'insieme di tutte le matrici ad m righe ed n colonne a coefficienti in $\mathbb{K}$. Se $A = (a_{ij})$ e $B = (b_{ij}) \in \mathbb{K}^{m,n}$, allora

$$A = B \quad \Leftrightarrow \quad a_{ij} = b_{ij} \quad \forall i, j\ .$$

Una *sottomatrice* $B \in \mathbb{K}^{p,q}$ di una matrice $A = (a_{ij}) \in \mathbb{K}^{m,n}$ è una matrice i cui elementi appartengono a p righe e a q colonne prefissate di A.

Un *minore* è una sottomatrice quadrata.

Esempio.

$$A = \begin{pmatrix} 3 & 5 & 0 \\ 1 & 2 & 3 \\ 0 & -1 & -5 \end{pmatrix}, \quad B = \begin{pmatrix} 5 & 0 \\ -1 & -5 \end{pmatrix} \quad \begin{matrix} i = 1, 3 \\ j = 2, 3 \end{matrix}.$$

Si chiama *trasposta* di $A \in \mathbb{K}^{m,n}$ la matrice A^T o ${}^tA \in \mathbb{K}^{n,m}$ ottenuta da A scambiando ordinatamente le righe con le colonne:

$$A = (a_{ij}) \quad \Rightarrow \quad {}^tA = (a_{ji})\,.$$

Esempio.

$$A = \begin{pmatrix} 2 & 3 \\ 1 & 0 \\ 5 & \pi \end{pmatrix}, \qquad {}^tA = \begin{pmatrix} 2 & 1 & 5 \\ 3 & 0 & \pi \end{pmatrix}.$$

Casi particolari di matrici quadrate sono

A *simmetrica*	se $a_{ij} = a_{ji}$
A *antisimmetrica*	se $a_{ij} = -a_{ji}$
A *triangolare superiore*	se $a_{ij} = 0$, $i > j$
A *triangolare inferiore*	se $a_{ij} = 0$, $i < j$
A *diagonale*	se $a_{ij} = 0$, $i \neq j$
A *unità*, o *identica*	se $a_{ij} = \delta_{ij}$, dove $\delta_{ij} = 0$ se $i \neq j$ e $\delta_{ii} = 1$.

La matrice identità sarà indicata con $I \in \mathbb{K}^{n,n}$, ed i suoi coefficienti sono collettivamente indicati con (δ_{ij}); δ_{ij} è detto anche *simbolo di Kronecker.*

Una matrice di *permutazione* P è una matrice che si ottiene dalla matrice identità con scambi di righe e colonne.

3.2 Operazioni su vettori e matrici

Se $\lambda \in \mathbb{K}$ e $A \in \mathbb{K}^{m,n}$, la matrice λA, *moltiplicazione di A per lo scalare λ*, è la matrice

$$\lambda A = (\lambda a_{ij}) \quad \forall i, j$$

Due matrici A e B sono sommabili se entrambe appartengono a $\mathbb{K}^{m,n}$. La matrice *somma* $C = A + B$ è per definizione $C = (c_{ij}) \in \mathbb{K}^{m,n}$ con

$$c_{ij} = a_{ij} + b_{ij}\,.$$

La matrice O avente tutti gli elementi 0 è la matrice *nulla*, e soddisfa

$$A + O = A \qquad \forall A\,,$$

e l'*opposta* di A è la matrice $A' = -A$, dove $a'_{ij} = -a_{ij}$ $\forall i, j$. Quindi $(\mathbb{K}^{m,n}, +)$ è un gruppo abeliano.

Esercizi.

- Dimostrare che ${}^t(A+B) = {}^tA + {}^tB$ e ${}^t(\lambda A) = \lambda\, {}^tA$.
- Se $A \in \mathbb{K}^{n,n}$, provare che

$$\begin{aligned} \tilde{A} &= \frac{1}{2}(A + {}^tA) \quad \text{è simmetrica;} \\ \bar{A} &= \frac{1}{2}(A - {}^tA) \quad \text{è antisimmetrica;} \\ A &= \tilde{A} + \bar{A}\,. \end{aligned}$$

La matrice A è moltiplicabile (*righe per colonne*) per la matrice B se $A \in \mathbb{K}^{m,n}$ e $B \in \mathbb{K}^{n,r}$. La matrice *prodotto* di A e B è la matrice $C = AB \in \mathbb{K}^{m,r}$, con $C = (c_{ik})$ dove

$$c_{ik} = a_{i1}b_{1k} + a_{i2}b_{2k} + \cdots + a_{in}b_{nk} = \sum_{j=1}^{n} a_{ij}b_{jk}$$

è il prodotto della riga i-esima di A per la colonna k-esima di B.

Si noti che in generale non ha senso anche la moltiplicazione BA.

La moltiplicazione tra matrici soddisfa alle regole

$$\begin{aligned} &A(BC) = (AB)C, && \text{per ogni} \quad A \in \mathbb{K}^{m,n}, B \in \mathbb{K}^{n,p}, C \in \mathbb{K}^{p,q} \\ &(A+B)C = AC + BC, && \text{per ogni} \quad A, B \in \mathbb{K}^{m,n}, C \in \mathbb{K}^{n,p} \\ &A(B+C) = AB + AC, && \text{per ogni} \quad A \in \mathbb{K}^{m,n}, B, C \in \mathbb{K}^{n,q} \end{aligned}$$

dunque $(\mathbb{K}^{n,n}, +, \cdot)$ è un anello.

Si noti che, anche se la moltiplicazione AB è definita, la moltiplicazione BA può non essere definita. Tuttavia se $A \in \mathbb{K}^{m,n}$ e $B \in \mathbb{K}^{n,m}$ entrambe $AB \in \mathbb{K}^{m,m}$ e $BA \in \mathbb{K}^{n,n}$ sono definite ma il risultato è diverso se $m \neq n$. Tuttavia, anche nel caso quadrato può accadere

$$AB \neq BA\,.$$

Esempio.

$$\begin{aligned} A &= \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \\ AB &= \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = BA\,. \end{aligned}$$

Si dimostra facilmente che:

1. la matrice nulla O è tale che $AO = O$ e $OA = O$ (si noti che la lettera O in queste due formule indica matrici nulle di dimensioni diverse);

2. la matrice unità I è tale che

$$AI = A = IA \qquad \forall A\,.$$

Dunque $(\mathbb{K}^{n,n}, +, \cdot)$ è un anello unitario ma non è un anello commutativo, perché la proprietà commutativa non vale per tutte le matrici, anche se può valere per alcune particolari coppie di matrici. In particolare, due matrici si dicono *permutabili* se $AB = BA$.

Si osservi che si può avere $AB = O$ senza che A o B siano matrici nulle. In tal caso A e B si dicono *divisori dello zero*.

Esempi ed esercizi.

- Se $A = (1, 0, 3)$ allora

$$^tA = \begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix} \text{ e } A^tA = (10)\,, \quad ^tAA = \begin{pmatrix} 1 & 0 & 3 \\ 0 & 0 & 0 \\ 3 & 0 & 9 \end{pmatrix}\,.$$

- Provare che $^t(AB) = {}^tB^tA$.
- Se $A \in \mathbb{K}^{m,n}$, provare che $\tilde{A} = A^tA$ e $\bar{A} = {}^tAA$ sono simmetriche.
- Si osservi che se A e B sono simmetriche, in generale AB non è simmetrica:

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

Se A è una matrice quadrata, allora

$$A^2 = AA, \ldots, A^h = A^{h-1}A\,.$$

Se $AB = BA$, allora $(AB)^k = A^kB^k$. Questo non è vero, in generale, se $AB \neq BA$.

Una matrice $A \in \mathbb{K}^{n,n}$ è detta *ortogonale* se

$$^tAA = I = A^tA\,.$$

Esercizi.

- Trovare tutte le potenze della matrice $C = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 0 \end{smallmatrix}\right)$.
- Trovare tutte le potenze della matrice

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

- Provare che la matrice del punto precedente è ortogonale.
- Siano $A = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$, $B = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right)$. Vedere se
$$(A+B)^2 = A^2 + 2AB + B^2 .$$

Si dice *traccia* di una matrice $A \in \mathbb{K}^{n,n}$ lo scalare $\operatorname{tr} A = \sum_{i=1}^{n} a_{ii}$.

Una matrice $A \in \mathbb{K}^{n,n}$ è detta *nilpotente* se esiste un naturale k tale che $A^k = O$.

Una matrice $A \in \mathbb{K}^{n,n}$ è detta *idempotente* se $A^2 = A$ (e quindi $A^k = A$ per $k \in \mathbb{N} \setminus \{0\}$).

Una matrice $A \in \mathbb{K}^{n,n}$ è detta *invertibile* se esiste una matrice $A' \in \mathbb{K}^{n,n}$ tale che
$$AA' = I = A'A .$$
Si scrive in tal caso $A' = A^{-1}$. Quindi, se A è ortogonale, $A^{-1} = {}^tA$. Vedremo in seguito un criterio che ci permette di decidere quando una matrice è invertibile.

Se A e B sono matrici invertibili, allora
$$(AB)^{-1} = B^{-1}A^{-1} , \qquad (A^{-1})^{-1} = A .$$
Se A è invertibile, allora
$$\begin{aligned} AB = O \quad &\Rightarrow \quad B = O , \\ AB = AC \quad &\Rightarrow \quad B = C . \end{aligned}$$

Esercizi.

- Data la matrice $A = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$, stabilire se A è invertibile e in tal caso trovare l'inversa.
- Considerata la matrice (di permutazione)
$$J = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$
calcolare JX e XJ, dove $X \in \mathbb{K}^{3,3}$ è un'arbitraria matrice; verificare inoltre che $J^2 = I$ e trovare l'inversa di J.
- Date le matrici
$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} , \qquad U = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$
calcolare $(A+I)^3$ e U^h con $h \in \mathbb{N}$.

- Vedere se sono invertibili le matrici

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \qquad U = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

Nota. Le matrici sono molto utili in matematica: permettono di semplificare complicate espressioni considerando tutta la tabella come un unico ente. Le matrici intervengono nella schematizzazione di molti fenomeni, dipendenti da un numero finito di parametri. Sono nella pratica molto usate nei problemi di decisione, nei quali è necessario procedere ad una scelta tra diversi modi di agire, ossia tra più strategie.

Come vedremo più avanti, se vogliamo risolvere un sistema di equazioni lineari, basta conoscere la matrice associata, cioè la matrice ci dà tutte le informazioni necessarie per risolverlo. La matrice è quindi la *mater matris*, la parte essenziale. La denominazione *matrice* è dovuta al matematico inglese J. J. Sylvester (1814–1897).

3.3 Determinanti

Storicamente la teoria dei determinanti di una matrice quadrata è nata in relazione allo studio dell'eliminazione delle incognite in un sistema di equazioni lineari. Già nelle scuole secondarie è stata introdotta la nozione di determinante di una matrice quadrata:

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}; \qquad \det A = ad - bc\,.$$

Vogliamo ora estendere questo concetto a matrici quadrate di ogni ordine.

Sia S un insieme. Si chiama *permutazione* di S ogni corrispondenza biunivoca di S in sé. Se S è finito e $\mathrm{card}(S) = n$, allora tutte le permutazioni di S sono $n!$. Esse si possono pensare come permutazioni $(\sigma(1), \sigma(2), \dots, \sigma(n))$ dell'insieme numerico $\{1, 2, \dots, n\}$. La permutazione $(1, 2, \dots, n)$ è chiamata *fondamentale* (o permutazione identica). L'insieme di tutte le permutazioni di n oggetti è indicato con S_n. Ogni permutazione σ si può ottenere dalla fondamentale tramite *scambi*, ovvero permutazioni di due soli elementi. Se il numero di scambi della permutazione σ è pari porremo $\epsilon(\sigma) = 1$, se è dispari $\epsilon(\sigma) = -1$.

Esempio. $\sigma = (3, 2, 1)$:

$$(3, 2, 1) \to (3, 1, 2) \to (1, 3, 2) \to (1, 2, 3)\,.$$

Il numero di scambi è 3, quindi $\epsilon(\sigma) = -1$. Se percorriamo un'altra strada il numero di scambi potrà essere diverso, ma sempre dispari.

Introduciamo la seguente notazione: data $A \in \mathbb{K}^{n,n}$, si pone $A = (A_1, \dots, A_n)$, dove $A_1, \dots, A_n$ sono le colonne di A.

Definizione 3.1. *Una funzione* $f \colon \mathbb{K}^{n,n} \to \mathbb{K}$ *si dice* multilineare alternante *rispetto alle colonne se*

1. $f(A_1, \dots, A_i + A'_i, \dots, A_n) = f(A_1, \dots, A_i, \dots, A_n) + f(A_1, \dots, A'_i, \dots, A_n)$,

2. $f(A_1, \dots, \lambda A_i, \dots, A_n) = \lambda f(A_1, \dots, A_i, \dots, A_n)$,

3. $f(A_1, \dots, A_i, \dots, A_j, \dots, A_n) = -f(A_1, \dots, A_j, \dots, A_i, \dots, A_n)$.

Teorema 3.1. *Esiste un'unica funzione $f\colon \mathbb{K}^{n,n} \to \mathbb{K}$ multilineare alternante rispetto alle colonne tale che $f(I) = 1$.*

Dimostrazione. Ogni colonna A_j di $A \in \mathbb{K}^{n,n}$ si può esprimere come combinazione lineare

$$A_j = a_{1j}E_1 + \cdots + a_{nj}E_n = \sum_{i=1}^{n} a_{ij}E_j,$$

dove $E_j = {}^t(0, \dots, 1_j, 0, \dots, 0)$ è il vettore colonna che ha per componente i-esima il numero δ_{ij} (simbolo di Kronecker). Dunque, applicando la proprietà di multilinearità n volte negli n argomenti di f:

$$\begin{aligned} f(A) &= f(\sum_{i_1=1}^{n} a_{i_1 1}E_{i_1}, \dots, \sum_{i_n=1}^{n} a_{i_n n}E_{i_n}) \\ &= \sum_{i_1,\dots,i_n=1}^{n} a_{i_1 1} \cdots a_{i_n n} f(E_{i_1}, \dots, E_{i_n}). \end{aligned}$$

Si noti ora che se la stessa colonna compare due volte tra gli argomenti di f la f si annulla, infatti:

$$f(\dots, A_i, \dots, A_i, \dots) = -f(\dots, A_i, \dots, A_i, \dots), \tag{3.3.1}$$

dunque $2f(\dots, A_i, \dots, A_i, \dots) = 0$. Pertanto gli unici valori di i_1, ..., i_n che danno contributo non nullo alla precedente somma sono quelli dove gli indici assumono valori a due a due distinti, ovvero, tutti i valori compresi tra 1 ed n (non necessariamente in questo ordine). In altre parole, l'espressione $f(E_{i_1}, \dots, E_{i_n})$ assume valori non nulli solo quando $i_1 = \sigma(1)$, ..., $i_n = \sigma(n)$, dove σ è una permutazione. Pertanto si può riscrivere la somma precedente come:

$$f(A) = \sum_{\sigma \in S_n}^{n} a_{\sigma(1)1} \cdots a_{\sigma(n)n} f(E_{\sigma(1)}, \dots, E_{\sigma(n)}). \tag{3.3.2}$$

Per la multilinearità si ha

$$f(E_{\sigma(1)}, \dots, E_{\sigma(n)}) = \epsilon(\sigma) f(E_1, \dots, E_n) = \epsilon(\sigma). \tag{3.3.3}$$

Dunque l'unica funzione f ha espressione

$$f(A) = \sum_{\sigma \in S_n}^{n} \epsilon(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(n)n}$$

$\boxed{QED}$

Corollario 3.1. *Se* $f\colon \mathbb{K}^{n,n} \to \mathbb{K}$ *è multilineare alternante rispetto alle colonne, allora* $f(A) = \det(A) f(I)$.

Dimostrazione. Segue dalle equazioni (3.3.2), (3.3.3). $\boxed{QED}$

Definizione 3.2. *Se* $A = (a_{ij}) \in \mathbb{K}^{n,n}$, *chiamiamo* determinante *di* A *l'unica funzione multilineare alternante delle colonne di* A *che vale* 1 *se* $A = I$*; la sua espressione è:*

$$\det A = \sum_{\sigma \in S_n} \epsilon(\sigma)\, a_{\sigma(1)1} a_{\sigma(2)2} \dots a_{\sigma(n)n}. \tag{3.3.4}$$

Molte proprietà del determinante possono essere dimostrate usando la precedente definizione.

Proposizione 3.1. *Sia* $A \in \mathbb{K}^{n,n}$. *Allora valgono le seguenti proprietà.*

1. *Se* A *ha una colonna uguale a* 0, *allora* $\det A = 0$.
2. *Se* A *ha due colonne uguali, allora* $\det A = 0$.
3. *Se* A *ha una colonna che è combinazione lineare delle altre (dunque è dipendente dalle altre), allora* $\det A = 0$.
4. $\det A \neq 0 \quad \Rightarrow \quad$ *le colonne di* A *sono indipendenti.*
5. *Per* $\lambda \in \mathbb{K}$ *si ha* $\det(\lambda A) = \lambda^n \det A$.
6. *Posto* $A = (A_1, \dots, A_n)$ *sia* $A' = (A_1, \dots, A_i + \lambda A_j, \dots, A_n)$ *la matrice ottenuta da* A *sostituendo la colonna* i*-esima con la somma della colonna* i*-esima e di un'altra qualunque colonna* A_j *di* A, *con* $j \neq i$, *moltiplicata per un qualunque scalare* $\lambda \in \mathbb{K}$. *Allora* $\det A = \det A'$.
7. $\det A = \det {}^tA$.
8. Regola di Binet*:* $\det(AB) = \det A \det B$.
9. *Le proprietà sopra enunciate per colonne valgono anche per righe.*

Dimostrazione. 1. Si ha $\det A = \det(A_1, \dots, 0, \dots, A_n) = \det(A_1, \dots, 0A_i, \dots, A_n) = 0 \det(A_1, \dots, A_i, \dots, A_n) = 0$, dove A_i è un vettore colonna arbitrario.

2. È conseguenza dell'equazione (3.3.1).
3. Infatti, per la multilinearità si può scrivere il determinante come combinazione lineare di determinanti di matrici con almeno due colonne uguali.
4. Infatti, "$\det A \neq 0 \quad \Rightarrow \quad$ colonne di A indipendenti" è equivalente a "colonne di A dipendenti $\quad \Rightarrow \quad \det A = 0$", che è stata dimostrata nel punto precedente.

5. Si ha
$$\det \lambda A = \det(\lambda A_1, \dots, \lambda A_n) = \lambda^n \det(A_1, \dots, A_n) = \lambda^n \det A$$

6. Infatti:
$$\det A' = det(A_1, \dots, A_n) + \lambda \det(A_1, \dots, A_j, \dots, A_j, \dots, A_n) = \det A$$
dalla prima proprietà.

7. Infatti:
$$\begin{aligned}\det {}^tA &= \sum_{\sigma \in S_n} \epsilon(\sigma)\, a_{1\sigma(1)} a_{2\sigma(2)} \dots a_{n\sigma(n)} \\ &= \sum_{\sigma \in S_n} \epsilon(\sigma)\, a_{\sigma^{-1}(1)1} a_{\sigma^{-1}(2)2} \dots a_{\sigma^{-1}(n)n} \\ &= \sum_{\sigma \in S_n} \epsilon(\sigma)\, a_{\sigma(1)1} a_{\sigma(2)2} \dots a_{\sigma(n)n}.\end{aligned}$$

8. Si consideri la funzione $f\colon \mathbb{K}^{n,n} \to \mathbb{K}$, $f(X) = \det(AX)$. Si può scrivere $f(X) = \det(AX_1, \dots, AX_n)$, dunque f è multilineare alternante. Per il corollario 3.1 si ha $f(B) = \det(B) f(I) = \det(B)\det(A) = \det(A)\det(B)$.

9. Segue da $\det A = \det {}^tA$.

$\boxed{QED}$

Quindi, in generale, $AB \neq BA$ ma $\det(AB) = \det(BA)$.

Naturalmente, se $n = 1$ si avrà $\det A = a_{11}$. Se $n = 2$, le permutazioni sono soltanto $(1, 2)$ e $(2, 1)$, quindi
$$\det A = a_{11}a_{22} - a_{12}a_{21}\,,$$
e se $n = 3$
$$\det A = a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} + a_{12}a_{23}a_{31} - a_{12}a_{21}a_{33}\,,$$
che molti riconosceranno come *formula di Sarrus* (*valida solo per* $n = 3$). Per n qualsiasi il calcolo è laborioso. Soccorre però una regola di calcolo dovuta a Laplace.

Fissato un elemento a_{ij} di A, si chiama *minore complementare* di a_{ij}la sottomatrice di A di ordine $n-1$, ottenuta cancellando la i-esima riga e la j-esima colonna. Si chiama *complemento algebrico* di a_{ij} o *cofattore* di a_{ij}
$$A_{ij} = (-1)^{i+j} \det(\text{minore complementare di } a_{ij})\,.$$

Teorema 3.2 (Laplace)**.** *Sia A una matrice quadrata di ordine n. Allora*
$$\det A = a_{r1}A_{r1} + \dots + a_{rn}A_{rn}\,,$$

dove r è una fissata riga (scelta arbitrariamente), oppure

$$\det A = a_{1c}A_{1c} + \cdots + a_{nc}A_{nc}\,,$$

dove c è una fissata colonna (scelta arbitrariamente).

Dimostrazione. Il metodo di dimostrazione si basa sull'osservazione che la funzione

$$f\colon \mathbb{K}^{n,n} \to \mathbb{K}, \quad A \mapsto \sum_{i=1}^{n} a_{ik}A_{ik}$$

è multilineare alternante e vale 1 sulla matrice identità. Tuttavia la dimostrazione è lunga e pertanto è omessa. QED

Si può facilmente dimostrare che la somma degli elementi di una linea per i complementi algebrici di un'altra linea è zero.

Questa regola permette di calcolare il determinante in maniera ricorsiva:

$$|A| = \det A = \begin{cases} a_{11} & \text{se } n = 1 \\ \sum_i a_{ij}A_{ij} = \sum_j a_{ij}A_{ij} & \text{se } n > 1 \end{cases}$$

Esempi ed esercizi.

- Se $I \in \mathbb{K}^{n,n}$, allora $\det I = 1$, $\det(-I) = (-1)^n$.
- Provare con un esempio che, in generale, $\det(A + B) \neq \det A + \det B$.
- Dalla Proposizione 3.1 è possibile estrarre un algoritmo per il calcolo del determinante. Ad esempio, si calcoli il determinante della matrice

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

Si può usare la proprietà 6 della Proposizione 3.1 (vista sulle righe!) per annullare tutti gli elementi sottostanti il primo elemento della prima riga, e così via:

$$\begin{vmatrix} 2 & 1 & 3 & 0 \\ -1 & 1 & 2 & 2 \\ 2 & 0 & -1 & -1 \\ -3 & 1 & 0 & 1 \end{vmatrix} = \begin{vmatrix} 2 & 1 & 3 & 0 \\ 0 & 3/2 & 7/2 & 2 \\ 0 & -1 & -4 & -1 \\ 0 & 5/2 & 9/2 & 1 \end{vmatrix},$$

dove $r_2 \mapsto r_2 + 1/2\, r_1$, $r_3 \mapsto r_3 - r_1$, $r_4 \mapsto r_4 + 3/2\, r_1$;

$$\begin{vmatrix} 2 & 1 & 3 & 0 \\ 0 & 3/2 & 7/2 & 2 \\ 0 & -1 & -4 & -1 \\ 0 & 5/2 & 9/2 & 1 \end{vmatrix} = \begin{vmatrix} 2 & 1 & 3 & 0 \\ 0 & 3/2 & 7/2 & 2 \\ 0 & 0 & -5/3 & 1/3 \\ 0 & 0 & -4/3 & -7/3 \end{vmatrix},$$

dove $r_3 \mapsto r_3 + 2/3\, r_2$, $r_4 \mapsto r_4 - 5/3\, r_2$; infine

$$\begin{vmatrix} 2 & 1 & 3 & 0 \\ 0 & 3/2 & 7/2 & 2 \\ 0 & 0 & -5/3 & 1/3 \\ 0 & 0 & -4/3 & -7/3 \end{vmatrix} = \begin{vmatrix} 2 & 1 & 3 & 0 \\ 0 & 3/2 & 7/2 & 2 \\ 0 & 0 & -5/3 & 1/3 \\ 0 & 0 & 0 & -13/5 \end{vmatrix}.$$

Dunque risulta $\det A = +13$, poiché il determinante di A è uguale al determinante dell'ultima matrice che è uguale, a sua volta, al prodotto degli elementi diagonali, in quanto matrice triangolare.

Si noti che, in caso di elemento a_{11} nullo, si può applicare la proprietà di alternanza (sulle righe!) e scambiare la prima riga con una riga che comincia per un elemento non nullo, cambiando di segno al risultato. Ovviamente, se la prima colonna è nulla, il determinante è nullo.

3.4 Matrici invertibili

Definizione 3.3. *Sia* $A \in \mathbb{K}^{n,n}$*,* $A = (a_{ij})$*. La matrice* $\mathrm{Adj}(A)$*,* aggiunta classica *di* A*, è definita come la matrice* $\mathrm{Adj}(A) = (Adj_{ij}) = (A_{ji})$*, ossia la matrice trasposta della matrice dei cofattori* A_{ij} *di* a_{ij}*.*

Proposizione 3.2. *La matrice* A *è invertibile se e solo se* $\det A \neq 0$*. In questo caso si ha:*

1. $det(A^{-1}) = 1/\det A$,
2. $A^{-1} = \frac{1}{\det A}\mathrm{Adj}(A)$.

Dimostrazione. Infatti, sia A invertibile. Allora $AA^{-1} = I \Rightarrow \det(AA^{-1}) = \det I \Rightarrow \det(AA^{-1}) = 1 \Rightarrow \det A \neq 0$ e $\det(A^{-1}) = 1/\det A$ per il teorema di Binet.

Viceversa, sia $\det A \neq 0$. Posto $C = (c_{ij}) = \frac{1}{\det A}\mathrm{Adj}(A)$, indicato $AC = (p_{ik})$, si ha

$$p_{ik} = \sum_{j=1}^{n} a_{ij}\frac{1}{\det A}A_{kj} = \frac{1}{\det A}\sum_{j=1}^{n} a_{ij}A_{kj} = \delta_{ik}.$$

Infatti, se $i = k$ la sommatoria precedente coincide con lo sviluppo di Laplace del determinante di A, e se $i \neq k$ la sommatoria precedente è nulla, poiché equivale allo sviluppo di Laplace di una matrice che ha righe uguali a tutte le righe di A tranne la riga k, che è uguale alla riga i. $\boxed{QED}$

Esempi ed esercizi.

- Se
$$A = \begin{pmatrix} 1/2 & 0 & 1 \\ 0 & 4 & 1 \\ 3 & 0 & 2 \end{pmatrix},$$
si ha $\det A = -8 \neq 0$ e
$$\mathrm{Adj}(A) = \begin{pmatrix} 8 & 0 & -4 \\ 3 & -2 & -1/2 \\ -12 & 0 & 2 \end{pmatrix} \quad \Rightarrow \quad A^{-1} = \begin{pmatrix} -1 & 0 & 1/2 \\ -3/8 & 1/4 & 1/16 \\ 3/2 & 0 & -1/4 \end{pmatrix}$$

- Sia data la matrice
$$A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}.$$
  1. Verificare che $A^2 - A - 2I = 0$.
  2. Dedurre dal punto precedente che A è invertibile e calcolare A^{-1}.

- Si considerino matrici quadrate di ordine n.
  A' si dice *simile* ad A se esiste una matrice invertibile P tale che $A' = P^{-1}AP$, cioè
$$A' \sim A \quad \Leftrightarrow \quad A' = P^{-1}AP \quad \Leftrightarrow \quad PA' = AP\,.$$
  1. Provare che $\sim$ è una relazione di equivalenza.
  2. Se $A' \sim A$ e $B' \sim B$, è vero che $A'B' \sim AB$?
  3. Se $A' \sim A$, allora $\det A = \det A'$?

- Sia $A \in \mathbb{K}^{n,n}$ invertibile. È possibile trovare un intero r tale che $A^r = 0$?

- Sia $GL(n, \mathbb{K}) = \{A \in \mathbb{K}^{n,n} \mid \det A \neq 0\}$. Provare che $(GL(n, \mathbb{K}), \cdot)$, dove $\cdot$ è la moltiplicazione tra matrici, è un gruppo non abeliano.

- Vedere che, per $\mathbb{K} = \mathbb{R}$, $GL^+(n, \mathbb{R}) = \{A \in \mathbb{R}^{n,n} \mid \det A > 0\}$ è un sottogruppo di $GL(n, \mathbb{R})$, mentre non lo è $GL^-(n, \mathbb{R})$.

3.5 Rango di una matrice

Sia $A \in \mathbb{K}^{n,m}$. Da A possiamo estrarre sottomatrici quadrate di ordine r, $1 \leq r \leq \min(n, m)$. Di queste sottomatrici quadrate, dette *minori*, si può fare il determinante e vedere se non è nullo.

Definizione 3.4. *Il rango* $\mathrm{rg}(A)$ *di una matrice* $A \in \mathbb{K}^{n,m}$ *è dato dal massimo ordine dei suoi minori con determinante non nullo.*

Quindi $\mathrm{rg}(A) = p > 0$ vuol dire

1. esiste almeno un minore di ordine p con determinante diverso da 0;

2. tutti gli eventuali minori di ordine $p + 1$ hanno determinante nullo.

Naturalmente, $\mathrm{rg}(A) = 0$ vuol dire che la matrice è nulla.
Se $A \in \mathbb{K}^{n,n}$ allora

$$\mathrm{rg}(A) = n \quad \Leftrightarrow \quad \det A \neq 0 \quad \Leftrightarrow \quad A \text{ invertibile,}$$

Ovviamente, se B è una sottomatrice di A, allora $\mathrm{rg}(B) \leq \mathrm{rg}(A)$.

Esempi ed esercizi.

- La matrice
$$A = \begin{pmatrix} -1 & 3 & 2 & 5 \\ 6 & -2 & 4 & 3 \\ -2 & 6 & 4 & 10 \end{pmatrix}$$
ha rango 2, poiché $\det\left(\begin{smallmatrix} 3 & 2 \\ -2 & 4 \end{smallmatrix}\right) \neq 0$, e tutti i minori di ordine 3 hanno determinante nullo.

- Determinare il rango delle seguenti matrici, al variare di λ,
$$A = \begin{pmatrix} 3 & -2 & \lambda \\ 2 & 1 & \lambda \\ 1 & 4 & \lambda \end{pmatrix}, \qquad B = \begin{pmatrix} \lambda & 1 & 1 \\ 1 & \lambda & 1 \\ 1 & 1 & \lambda \end{pmatrix}.$$
Si vede che $\mathrm{rg}(A) = 2\ \forall\lambda$; $\mathrm{rg}(B) = 3$ per $\lambda \neq 1$ e $\lambda \neq -2$, mentre $\mathrm{rg}(B) = 2$ per $\lambda = -2$ e $\mathrm{rg}(B) = 1$ per $\lambda = 1$.

Proposizione 3.3. *Sia $A \in \mathbb{K}^{m,n}$. Allora le seguenti trasformazioni su A dànno matrici che hanno lo stesso rango di A:*

1. *scambio di due righe;*

2. *moltiplicazione di una riga per un elemento $k \in \mathbb{K}^* = \mathbb{K} \setminus \{0\}$;*

3. *sostituzione di una riga con la somma della riga stessa con un'altra riga di A moltiplicata per un elemento $k \in \mathbb{K}$;*

4. *trasposizione della matrice A, quindi tutte le operazioni precedenti effettuate per colonne.*

DIMOSTRAZIONE. Segue facilmente dalle proprietà dei determinanti. $\boxed{QED}$

Si osservi che dalla precedente proposizione risulta $\operatorname{rg} A = \operatorname{rg} {}^tA$.

Le trasformazioni elencate nella proposizione precedente si dicono anche *trasformazioni elementari sulle righe*.

Due matrici $A, B \in \mathbb{K}^{m,n}$ si dicono *equivalenti per righe* se l'una è ottenibile dall'altra mediante un numero finito di trasformazioni elementari sulle righe. Ovviamente, in questo caso $\operatorname{rg}(A) = \operatorname{rg}(B)$.

Una matrice $S \in \mathbb{K}^{m,n}$ si dice *matrice a scalini* (per righe)se è la matrice nulla oppure se è della forma

$$\begin{pmatrix} 0 & \dots & 0 & a_{1j_1} & \dots\dots\dots\dots\dots\dots & & & a_{1n} \\ 0 & \dots\dots\dots\dots & & 0 & a_{2j_2} & \dots\dots\dots\dots & & a_{2n} \\ \dots\dots & \dots\dots & \dots\dots & \dots\dots & \dots\dots & \dots\dots & \dots\dots & \dots\dots \\ 0 & \dots\dots\dots\dots\dots\dots & & & 0 & a_{pj_p} & \dots & a_{pn} \\ 0 & \dots\dots\dots\dots\dots\dots\dots\dots & & & & & & 0 \\ 0 & \dots\dots\dots\dots\dots\dots\dots\dots & & & & & & 0 \end{pmatrix},$$

dove $a_{1j_1} \neq 0$, $a_{2j_2} \neq 0$, ..., $a_{pj_p} \neq 0$, con $1 \leq p \leq m$ e $1 \leq j_1 < j_2 < \cdots < j_p \leq n$.

Lemma 3.1. *Sia $S \in \mathbb{K}^{m,n}$ una matrice a scalini. Allora $\operatorname{rg} S = p$, numero delle righe non nulle.*

Dimostrazione. Il minore formato dalle prime p righe e dalle colonne $j_1, \dots, j_p$ è

$$\begin{pmatrix} a_{1j_1} & \dots\dots & a_{1j_p} \\ 0 & a_{2j_2} \quad . & a_{2j_p} \\ \dots\dots & \dots\dots & \dots\dots \\ 0 & \dots \quad 0 & a_{pj_p} \end{pmatrix}$$

ed il suo determinante è $a_{1j_1} \cdot \cdots \cdot a_{pj_p} \neq 0$. Qualunque minore di ordine $p+1$ conterrebbe una riga di zeri. $\boxed{QED}$

Teorema 3.3. *Ogni matrice $A \in \mathbb{K}^{m,n}$ è equivalente per righe ad una matrice a scalini.*

Dimostrazione. La dimostrazione è costruttiva.

Al primo passo si considerino gli elementi $a_{1j_1}, \dots a_{mj_m}$ non nulli di ogni riga con indice j_i minimo relativamente alla riga (ossia si considerino tutti i primi elementi non nulli di ogni riga). Se necessario, si scambi la prima riga con la riga avente il minimo indice tra gli indici $j_1, \dots, j_m$. Questo serve ad ottenere una matrice con lo stesso rango di A e con elementi di posto i, j nulli se $j < j_1$. Si effettui la trasformazione 3 annullando tutti gli elementi di posto i, j con $i > 1$ e $j = j_1$.

Si ripeta il primo passo sulle rige 2, ..., p, dove la riga $p + 1$ risulti nulla. $\boxed{QED}$

Corollario 3.2. *Ogni matrice $A \in \mathbb{K}^{n,n}$ invertibile è equivalente per righe ad una matrice triangolare superiore con elementi non nulli sulla diagonale principale.*

Il collegamento tra rango di una matrice e dipendenza od indipendenza lineare dell'insieme dei propri vettori riga o colonna è di fondamentale importanza.

Teorema 3.4. *Sia $A \in \mathbb{K}^{m,n}$, e si denotino con A^1,…,A^m i vettori riga di A. Allora:*

$$\dim \mathcal{L}(\{A^1, \dots, A^m\}) = \operatorname{rg} A.$$

Analogamente per l'insieme dei vettori colonna.

DIMOSTRAZIONE. Prima di tutto si dimostra che $\dim \mathcal{L}(\{A^1, \dots, A^m\})$ non cambia se si effettuano su A le trasformazioni per righe della proposizione 3.3. Infatti, per le prime due operazioni questo è ovvio, per la terza si dimostra che

$$\mathcal{L}(\{A^1, \dots, A^m\}) = \mathcal{L}(\{A^1, \dots, A^i + kA^j, \dots, A^m\}),$$

dove $j \neq i$ e $k \in \mathbb{K}$. Sia

$$X = \lambda_1 A^1 + \cdots + \lambda_i(A^i + kA^j) + \cdots + \lambda_m A^m \in \mathcal{L}(\{A^1, \dots, A^i + kA^j, \dots, A^m\}).$$

Assumendo $j < i$ risulta $X = \lambda_1 A^1 + \cdots + (\lambda_j + \lambda_i k)A^j + \cdots + \lambda_i A^i + \cdots + \lambda_m A^m$ (e analogamente se fosse $j > i$), quindi $X \in \mathcal{L}(\{A^1, \dots, A^m\})$. Sia ora

$$X = \lambda_1 A^1 + \cdots + \lambda_m A^m \in \mathcal{L}(\{A^1, \dots, A^m\}).$$

Si ha

$$X = \lambda_1 A^1 + \cdots + (\lambda_j - \lambda_i k)A^j + \cdots + \lambda_i(A^i + kA^j) + \cdots + \lambda_m A^m$$

pertanto $X \in \mathcal{L}(\{A^1, \dots, A^i + kA^j, \dots, A^m\})$.

Dunque, se S è una matrice a scalini equivalente per righe ad A, e se S^1, …, S^m sono le righe di S, risulta $\mathcal{L}(\{A^1, \dots, A^m\}) = \mathcal{L}(\{S^1, \dots, S^m\})$.

Una base di $\mathcal{L}(\{S^1, \dots, S^m\})$ è formata dalle righe non nulle $\{S^1, \dots, S^p\}$ di S: infatti S^1 non si può ricavare da $\{S^2, \dots, S^p\}$, S^2 non si può ricavare da $\{S^3, \dots, S^p\}$ e così via, che dimostra che le righe non nulle sono indipendenti per il lemma 2.1. Dunque:

$$\operatorname{rg}(A) = \operatorname{rg}(S) = p = \dim \mathcal{L}(\{S^1, \dots, S^m\}) = \dim \mathcal{L}(\{A^1, \dots, A^m\}).$$

Il fatto che il risultato vale anche per le colonne dipende dal fatto che $\operatorname{rg} A = \operatorname{rg} {}^t A$. $\boxed{QED}$

Corollario 3.3. *Se A e B sono due matrici equivalenti per righe, allora*

$$\mathcal{L}(\{A^1, \dots, A^m\}) = \mathcal{L}(\{B^1, \dots, B^m\}).$$

CAPITOLO 4

FUNZIONI TRA SPAZI VETTORIALI

In questa sezione saranno studiate le funzioni tra spazi vettoriali che conservano le operazioni di somma e prodotto, dunque le funzioni che conservano la struttura di spazio vettoriale.

4.1 Preliminari

Siano V, W due spazi vettoriali sullo stesso campo $\mathbb{K}$ e

$$f\colon V \to W$$

un'applicazione. Se $U \subset V$, allora chiamiamo *immagine* di U mediante f l'insieme

$$f(U) \stackrel{\text{def}}{=} \{f(x) \mid x \in U\} = \{y \in W \mid \exists x \in U : f(x) = y\}.$$

Se $Z \subset W$, allora chiamiamo *controimmagine* o *immagine inversa* di Z mediante f l'insieme

$$f^{-1}(Z) \stackrel{\text{def}}{=} \{x \in V \mid f(x) \in Z\} \subset V.$$

Se $Z = \{y\}$, allora $f^{-1}(y) = \{x \in V \mid f(x) = y\}$. Inoltre

$$y \in f(U) \quad \Rightarrow \quad f^{-1}(y) \neq \varnothing.$$

Esempio. L'applicazione $f\colon V \to W$ tale che $f(x) = y_0$ per ogni $x \in V$ è detta *applicazione costante*, e

$$f(V) = \{y_0\} \quad \Rightarrow \quad f^{-1}(y_0) = V, \quad f^{-1}(y) = \varnothing \quad \forall\, y \neq y_0.$$

Sia $\dim V = n$ e $\mathcal{B} = \{e_1, \dots, e_n\}$ una base di V. Sia $\dim W = m$ e $\mathcal{B}' = \{e'_1, \dots, e'_m\}$ una base di W. Allora

$$x \in V \quad \Rightarrow \quad x = x_1 e_1 + \cdots + x_n e_n \quad \Rightarrow \quad (x_1, \dots, x_n) \text{ coord. di } x \text{ risp. } \mathcal{B}$$
$$y \in W \quad \Rightarrow \quad y = y_1 e'_1 + \cdots + y_m e'_m \quad \Rightarrow \quad (y_1, \dots, y_m) \text{ coord. di } y \text{ risp. } \mathcal{B}'$$

Ora $f\colon V \to W$ induce un'applicazione che chiamiamo ancora f:

$$f\colon \mathbb{K}^n \to \mathbb{K}^m, (x_1, \ldots, x_n) \mapsto (y_1, \ldots, y_m).$$

Qual è il legame tra le coordinate x_i e y_j?

Poniamo $f_i\colon \mathbb{K}^n \to \mathbb{K}$, $f_i \stackrel{\text{def}}{=} \mathrm{pr}_i \circ f$, dove

$$\mathrm{pr}_i\colon \mathbb{K}^m \to \mathbb{K}, (y_1, \ldots, y_m) \mapsto y_i.$$

Allora la relazione vettoriale $f(x) = y$ si traduce nelle m relazioni scalari

$$f_i(x_1, \ldots, x_n) = y_i \qquad i = 1, \ldots, m.$$

Supponiamo ora che le f_i siano polinomi omogenei di primo grado nelle variabili $x_1, \ldots, x_n$. Allora

$$\begin{cases} y_1 = a_{11}x_1 + \cdots + a_{1n}x_n \\ \ldots \\ y_m = a_{m1}x_1 + \cdots + a_{mn}x_n \end{cases}$$

o, in forma matriciale,

$$Y = AX \quad \text{oppure} \quad f(X) = Y.$$

In tal caso,

$$f(x + x') = f(x) + f(x'), \quad f(\lambda x) = \lambda f(x).$$

Infatti,

$$A(X + X') = AX + AX', \quad A(\lambda X) = \lambda AX.$$

Le proprietà precedenti hanno suggerito la definizione astratta di applicazione lineare.

4.2 Applicazioni lineari

Definizione 4.1. *Siano V, W due spazi vettoriali sullo stesso campo $\mathbb{K}$. Sia $f\colon V \to W$. Si dice che f è* lineare *o un* omomorfismo *se*

$$f(x + xp) = f(x) + f(xp) \quad \forall x, xp \in V$$
$$f(\lambda x) = \lambda f(x) \quad \forall x \in V, \lambda \in \mathbb{K},$$

o, equivalentemente,

$$f(\lambda x + \mu xp) = \lambda f(x) + \mu f(xp) \quad \forall x, xp \in V,\ \forall \lambda, \mu \in \mathbb{K}.$$

Come si vede, le applicazioni lineari *conservano le operazioni vettoriali.* Segue facilmente che $f(0_V) = 0_W$.

Esempi.

- $0\colon V \to W$, $v \mapsto 0$, applicazione nulla è lineare. Si noti che una qualunque applicazione costante *non è* lineare a meno che la costante non sia il vettore nullo.
- $\mathrm{pr}\colon \mathbb{R}^3 \to \mathbb{R}^2$, $(x, y, z) \mapsto (x, y)$, proiezione, è lineare, suriettiva, ma non iniettiva.
- $i\colon \mathbb{R}^2 \to \mathbb{R}^3$, $(x, y) \mapsto (x, y, 0)$, inclusione, è lineare, iniettiva, ma non suriettiva.
- Sia $\vec{a} \in V$ un vettore fissato e $T_{\vec{a}}\colon V \to V$ tale che $x \mapsto x + \vec{a}$, traslazione. Allora $T_{\vec{a}}$ è lineare se e solo se $\vec{a} = 0$.
- Si prova facilmente che, se $f\colon V \to W$ e $g\colon W \to Z$ sono due applicazioni lineari, allora
  $$g \circ f\colon V \to Z, x \mapsto g \circ f(x) \overset{\text{def}}{=} g(f(x))$$
  è lineare.

Se $V = W$, allora l'applicazione lineare è detta *endomorfismo*. Un endomorfismo notevole è l'applicazione *identità*

$$\mathrm{Id}_V\colon V \to V, x \mapsto x.$$

Un endomorfismo $f\colon V \to V$ è detto *involutivo* se $f^2 = \mathrm{Id}$, è detto *proiettore* se $f^2 = f$, è detto *nilpotente* se esiste $m \in \mathbb{N}$ tale che $f^m = 0$, endomorfismo nullo.

Se $f, g\colon V \to W$, si pone per definizione

$$f + g\colon V \to W, x \mapsto (f + g)(x) \overset{\text{def}}{=} f(x) + g(x),$$
$$\lambda f\colon V \to W, x \mapsto (\lambda f)(x) \overset{\text{def}}{=} \lambda f(x).$$

Rispetto a queste operazioni, l'insieme

$$\mathrm{Lin}(V, W) \overset{\text{def}}{=} \{f\colon V \to W \mid f \text{ lineare}\}$$

è uno spazio vettoriale. Se V e W sono finitamente generati, allora (come sarà chiaro in seguito)

$$\dim(\mathrm{Lin}(V, W)) = \dim V \cdot \dim W.$$

La prossima proposizione garantisce l'esistenza di funzioni lineari tra due spazi vettoriali. Essa può essere generalizzata a spazi vettoriali di dimensione infinita.

Proposizione 4.1. *Siano V, W spazi vettoriali, con V finitamente generato. Sia $\mathcal{B} = \{v_1, \ldots, v_n\}$ una base di V, e siano $w_1, \ldots, w_n \in W$. Allora esiste un'unica applicazione lineare $f\colon V \to W$ tale che $f(v_i) = w_i$ per $i = 1, \ldots, n$.*

Dimostrazione. Infatti, per $x = x_1 v_1 + \cdots + x_n v_n$si ponga

$$f(x) \overset{\text{def}}{=} x_1 w_1 + \cdots + x_n w_n, \quad f(v_i) = w_i.$$

È immediato verificare che f è lineare. Inoltre f è unica. Infatti, se $g\colon V \to W$ è tale che $g(v_i) = w_i$, allora per ogni x

$$f(x) - g(x) = x_1 f(v_1) + \cdots + x_n f(v_n) - x_1 g(v_1) - \cdots - x_n g(v_n) = 0.$$

QED

Osservazione. In generale per conoscere un'applicazione $f\colon V \to W$ bisogna conoscere $f(x)$ per ogni $x \in V$. Il vantaggio delle applicazioni lineari è che per conoscere f basta conoscere solo $f(v_i)$ dove $\{v_i\}$ è una base di V, quindi basta un numero finito di vettori.

Un'altra notevole proprietà è la seguente (la cui dimostrazione completa è lasciata al lettore)

Proposizione 4.2. *Siano V, W due spazi vettoriali sul campo $\mathbb{K}$, ed $f\colon V \to W$ un'applicazione lineare. Allora:*

$$\begin{aligned} S \text{ sottosp. di } V \quad &\Rightarrow \quad f(S) \text{ sottosp. di } W, \\ T \text{ sottosp. di } W \quad &\Rightarrow \quad f^{-1}(T) \text{ sottosp. di } V. \end{aligned}$$

Dimostrazione. Se $y_1, y_2 \in f(S)$, allora $y_1 = f(x_1)$, $y_2 = f(x_2)$ con $x_1, x_2 \in S$. Per la linearità

$$y_1 + y_2 = f(x_1) + f(x_2) = f(x_1 + x_2) \in f(S),$$

poiché $x_1 + x_2 \in S$. Analogamente $\lambda y_1 = f(\lambda x_1) \in f(S)$ per $\lambda \in \mathbb{K}$. QED

Casi particolarmente importanti sono

$$\operatorname{Ker} f \overset{\text{def}}{=} \{x \in V \mid f(x) = 0\} = f^{-1}(0_W),$$
$$\operatorname{Im} f \overset{\text{def}}{=} \{f(x) \in W \mid x \in V\} = f(V),$$

che, quindi, risultano sottospazi vettoriali.

Esercizio. Si verifichi direttamente, usando la proposizione 2.1, che $\operatorname{Ker} f$ ed $\operatorname{Im} f$ sono sottospazi vettoriali.

Più precisamente,

- $\operatorname{Ker} f$ è un sottospazio vettoriale di V e $\dim \operatorname{Ker} f \leq \dim V$;
- $\operatorname{Im} f$ è un sottospazio vettoriale di W e $\dim \operatorname{Im} f \leq \dim W$.

Ricordiamo che un'applicazione $f\colon V \to W$ è <u>iniettiva</u> se e solo se $x \neq \vec{x'} \Rightarrow f(x) \neq f(\vec{x'})$.

Proposizione 4.3. *Sia $f\colon V \to W$ un'applicazione lineare. Allora*

$$f(x) = f(xp) \quad \Leftrightarrow \quad xp - x \in \operatorname{Ker} f \quad \Leftrightarrow \quad xp = x + \operatorname{Ker} f.$$

DIMOSTRAZIONE. Infatti,

$$f(x) = f(xp) \quad \Leftrightarrow \quad f(xp) - f(x) = 0 \quad \Leftrightarrow \quad f(xp - x) = 0 \quad \Leftrightarrow \quad xp - x \in \operatorname{Ker} f.$$

QED

Quindi, in generale,

$$f^{-1}(f(x)) = x + \operatorname{Ker} f = \{x + u \mid u \in \operatorname{Ker} f\},$$

e dunque

$$\begin{aligned} f \text{ è iniettiva} \quad &\Leftrightarrow \quad \operatorname{Ker} f = \{0_V\} \quad \Leftrightarrow \quad f^{-1}(f(x)) = x \quad \Leftrightarrow \quad f^{-1} \circ f = \mathrm{Id}_V, \\ f \text{ è suriettiva} \quad &\Leftrightarrow \quad \operatorname{Im} f = W \quad \Leftrightarrow \quad f(f^{-1}(y)) = y \quad \Leftrightarrow \quad f \circ f^{-1} = \mathrm{Id}_W\,. \end{aligned}$$

La seguente proprietà ha dimostrazione immediata.

Lemma 4.1. *Siano V, W spazi vettoriali ed $f\colon V \to W$ lineare. Allora*

$$U = \mathcal{L}(v_1, \dots, v_h) \quad \Rightarrow \quad f(U) = \mathcal{L}(f(v_1), \dots, f(v_h)),$$

da cui, se $\mathcal{B} = \{v_1, \dots, v_n\}$ è una base di V, allora $\operatorname{Im} f$ è finitamente generato e

$$\operatorname{Im} f = \mathcal{L}(f(v_1), \dots, f(v_n)).$$

Quindi, $\dim \operatorname{Im} f \leq \dim V$.

Segue che $\dim \operatorname{Im} f = \dim V$ se e solo se f è iniettiva. In tal caso, infatti,

$$\begin{aligned} \lambda_1 f(v_1) + \dots + \lambda_n f(v_n) = 0 \quad &\Leftrightarrow \quad f(\lambda_1 v_1 + \dots + \lambda_n v_n) = 0 \\ &\Leftrightarrow \quad \lambda_1 v_1 + \dots + \lambda_n v_n = 0 \quad \Leftrightarrow \quad \lambda_i = 0. \end{aligned}$$

Teorema 4.1 (Teorema fondamentale)**.** *Siano V, W due spazi vettoriali, con V finitamente generato. Sia $f\colon V \to W$ lineare. Allora*

$$\dim(\operatorname{Ker} f) + \dim(\operatorname{Im} f) = \dim V.$$

DIMOSTRAZIONE. Abbiamo già notato che $\dim(\operatorname{Ker} f) \leq \dim V$ e $\dim(\operatorname{Im} f) \leq \dim V$. Sia $\{u_1, \dots, u_p\}$ una base di $\operatorname{Ker} f$ e $w_1, \dots, w_q$ una base di $\operatorname{Im} f$. Indicati con $v_i \in V$ i vettori tali che $f(v_i) = w_i$, basta far vedere che $\{u_1, \dots, u_p, v_1, \dots, v_q\}$ costituiscono una base di V.

Essi sono *generatori*. Infatti, se $x \in V$ allora $f(x) \in \operatorname{Im} f$, dunque

$$f(x) = \lambda_1 w_1 + \cdots + \lambda_q w_q = \lambda_1 f(v_1) + \cdots + \lambda_q f(v_q) = f(\lambda_1 v_1 + \cdots + \lambda_q v_q),$$

da cui $x - \lambda_1 v_1 - \cdots - \lambda_q v_q \in \operatorname{Ker} f = \mathcal{L}(u_1, \ldots, u_p)$, quindi

$$x - \lambda_1 v_1 - \cdots - \lambda_q v_q = \mu_1 u_1 + \cdots + \mu_p u_p.$$

Essi sono *indipendenti*. Infatti se

$$a_1 u_1 + \cdots + a_p u_p + b_1 v_1 + \cdots + b_p v_p = 0_V, \tag{4.2.1}$$

applicando la f ad entrambi i membri si ha

$$a_1 f(u_1) + \cdots + a_p f(u_p) + b_1 f(v_1) + \cdots + b_q f(v_q) = 0_W,$$

dunque

$$b_1 w_1 + \cdots + b_q w_q = 0_W \quad \Rightarrow \quad b_1 = \cdots = b_q = 0$$

per l'indipendenza di $\{w_j\}$, dunque l'equazione (4.2.1) diventa

$$a_1 u_1 + \cdots + a_p u_p = 0_V,$$

da cui la tesi per l'indipendenza di $\{u_j\}$. QED

Chiamiamo

$$\dim(\operatorname{Ker} f) = \operatorname{nl}(f) = \textit{nullità di } f, \quad \dim(\operatorname{Im} f) = \operatorname{rg}(f) = \textit{rango di } f.$$

4.3 Isomorfismi

Definizione 4.2. *Siano V, W due spazi vettoriali. Se $f\colon V \to W$ è lineare e biunivoca, f si dice* isomorfismo*. (Ovviamente, in tal caso anche $f^{-1}\colon W \to V$ è un isomorfismo.)*

Se $V = W$ e $f\colon V \to V$ è un isomorfismo, allora f prende il nome di *automorfismo*.

Definizione 4.3. *Due spazi vettoriali V e W si dicono* isomorfi*, in simboli $V \cong W$, se esiste un isomorfismo $f\colon V \to W$.*

Teorema 4.2. *Siano V, W spazi vettoriali di dimensione finita. Allora*

$$V \cong W \quad \Leftrightarrow \quad \dim V = \dim W.$$

Dimostrazione. Sia $V \cong W$. Allora esiste un isomorfismo $f\colon V \to W$, e risulta $\operatorname{Ker} f = \{0\}$ e $\operatorname{Im} f = W$. Dal teorema fondamentale segue

$$\dim(W) = \dim(\operatorname{Im} f) = \dim V.$$

Viceversa, sia $\dim V = \dim W = n$. Sia $\mathcal{B} = \{e_1, \dots, e_n\}$ una base di V e $\mathcal{B}' = \{e'_1, \dots, e'_n\}$ una base di W. Definiamo un'applicazione lineare $f\colon V \to W$ ponendo $f(e_i) \stackrel{\text{def}}{=} e'_i$. La proposizione 4.1 assicura che f è univocamente definita. Per la linearità,

$$x = x_1 e_1 + \cdots + x_n e_n \quad \Rightarrow \quad f(x) = x_1 e'_1 + \cdots + x_n e'_n.$$

Se $x \neq xp$, allora $f(x) \neq f(xp)$, dunque f è iniettiva.

Infine, se $y = y_1 e'_1 + \cdots + y_n e'_n \in W$, allora posto $x = y_1 e_1 + \cdots + y_n e_n \in V$ si ha $f(x) = y$, dunque f è suriettiva. $\boxed{QED}$

Corollario 4.1. *Nelle ipotesi del precedente teorema, se $f\colon V \to W$ è un isomorfismo allora per ogni sottospazio $U \subset V$ e $S \subset W$ si ha*

$$\dim f(U) = \dim U \quad e \quad \dim f^{-1}(S) = \dim S.$$

DIMOSTRAZIONE. Infatti, le restrizioni di f ad U e $f^{-1}(S)$, rispettivamente $f|_U\colon U \to f(U)$ e $f|_{f^{-1}(S)}\colon f^{-1}(S) \to S$, sono lineari e biunivoche (se non lo fossero, così non sarebbe per f), quindi vale il teorema precedente. $\boxed{QED}$

Dal teorema precedente segue che se V è uno spazio vettoriale su $\mathbb{K}$ e $\dim V = n$, allora $V \cong \mathbb{K}^n$. Più precisamente, sia $\mathcal{B} = \{v_i\}_{1 \leq i \leq n}$ una base di V. Allora esiste un'unica applicazione lineare

$$c_{\mathcal{B}}\colon V \to \mathbb{K}^n, \quad v_i \mapsto e_i, \tag{4.3.2}$$

dove $\mathcal{C} = \{e_i\}_{1 \leq i \leq n}$ è la base canonica di $\mathbb{K}^n$. Dal lemma 4.1 si vede subito che $\operatorname{Im} c_{\mathcal{B}} = \mathbb{K}^n$, dunque per il teorema fondamentale $c_{\mathcal{B}}$ è un isomorfismo. L'applicazione $c_{\mathcal{B}}$ è detta applicazione *coordinata* di V in $\mathbb{K}^n$ rispetto a $\mathcal{B}$. Essa associa ad ogni $v \in V$ il vettore delle sue coordinate rispetto alla base $\mathcal{B}$. In altre parole, $\mathbb{K}^n$ è un *modello* per tutti gli spazi vettoriali di dimensione n.

Esempi ed esercizi.

- Lo spazio dei polinomi $\mathbb{R}_n[t]$è isomorfo allo spazio $\mathbb{R}^{n+1}$ tramite l'isomorfismo $c_{\mathcal{B}}$ indotto dalla base $\mathcal{B} = \{t^n, \dots, t, 1\}$ tale che $c_{\mathcal{B}}(a_n t^n + \cdots + a_1 t + a_0) = (a_n, \dots, a_0)$.

- Si consideri $\mathbb{R}_3[t]$ con la base $\{1, t, t^2, t^3\}$. L'applicazione

  $$D \stackrel{\text{def}}{=} \frac{d}{dt}\colon \mathbb{R}_3[t] \to \mathbb{R}_3[t]$$

  è tale che $D(t^h) = h t^{h-1}$ per $h \geq 1$ e $D(1) = 0$. Si verifica che D è lineare e che $\operatorname{Im} D = \mathbb{R}_2[t]$, quindi D non è suriettiva. D non è nemmeno iniettiva, poiché

  $$D(p(t)) = D(p(t) + k), \qquad \forall k \in \mathbb{R}.$$

Dal teorema fondamentale risulta

$$\dim(\operatorname{Ker} D) = \dim(\mathbb{R}_3[t]) - \dim(\operatorname{Im} D) = 4 - 3 = 1,$$

da cui $\operatorname{Ker} D = \mathbb{R}$.

- Si consideri l'applicazione lineare $f\colon \mathbb{R}^3 \to \mathbb{R}^3$, $(x, y, z) \mapsto (z, 0, 0)$. Si provi che $\mathbb{R}^3 \neq \operatorname{Ker} f + \operatorname{Im} f$, mentre, ovviamente, per il teorema fondamentale, $\dim \operatorname{Ker} f + \dim \operatorname{Im} f = 3$.

- Si consideri lo spazio vettoriale $\mathbb{R}^3$ e l'endomorfismo

$$f\colon \mathbb{R}^3 \to \mathbb{R}^3, (x, y, z) \mapsto (x + z, y, x + z).$$

  1. Determinare $\operatorname{Ker} f$ e $\operatorname{Im} f$.
  2. Verificare che $\operatorname{Ker}(f^2) = \operatorname{Ker} f$ e $\operatorname{Im}(f^2) = \operatorname{Im} f$.

4.4 Matrici ed applicazioni lineari

Siano V, W, due spazi vettoriali di dimensione finita, con $\dim V = n$, $\dim W = m$. Sia $f\colon V \to W$ un'applicazione lineare. Siano $\mathcal{B} = \{e_i\}_{1\leq i\leq n}$ una base di V e $\mathcal{B}' = \{e'_j\}_{1\leq j\leq m}$ una base di W. Allora

$$f(e_i) = a_{1i}e'_1 + \cdots + a_{mi}e'_m. \tag{4.4.3}$$

Viene così introdotta una matrice $A = (a_{ij})$ di tipo $m \times n$ dove la colonna i-esima è costituita dalle coordinate di $f(e_i)$ rispetto alla base $\mathcal{B}'$.

In simboli,

$$A = \mathcal{M}^{\mathcal{B}'}_{\mathcal{B}}(f),$$

detta *matrice associata* ad f rispetto alle basi $\mathcal{B}$ e $\mathcal{B}'$. Si noti che variando le basi la matrice cambia.

Viceversa, partendo da una matrice $A \in \mathbb{K}^{m,n}$ è possibile costruire un'applicazione lineare

$$f_A\colon \mathbb{K}^n \to \mathbb{K}^m,\ X \mapsto AX. \tag{4.4.4}$$

Ovviamente, $A = \mathcal{M}^{\mathcal{C}'}_{\mathcal{C}}(f_A)$, dove $\mathcal{C}$ e $\mathcal{C}'$ sono le basi canoniche di $\mathbb{K}^n$ e $\mathbb{K}^m$, rispettivamente. La relazione tra f ed f_A è schematizzata dal seguente diagramma

$$f\colon V \xrightarrow{c_{\mathcal{B}}} \mathbb{K}^n \xrightarrow{f_A} \mathbb{K}^m \xrightarrow{c^{-1}_{\mathcal{B}'}} W$$

o, in altri termini, $f = c^{-1}_{\mathcal{B}'} \circ f_A \circ c_{\mathcal{B}}$. Infatti,

$$c^{-1}_{\mathcal{B}'} \circ f_A \circ c_{\mathcal{B}}(v) = c^{-1}_{\mathcal{B}'} \circ f_A(X) = c^{-1}_{\mathcal{B}'}(AX),$$

dove X è il vettore delle componenti di $v \in V$ rispetto a $\mathcal{B}$. Ma, usando (4.4.3), risulta

$$\begin{aligned} f(v) &= f(v_1 e_1 + \cdots + v_n e_n) = v_1 f(e_1) + \cdots + v_n f(e_n) \\ &= (v_1 a_{11} + \cdots + v_n a_{1n}) e'_1 + \cdots + (v_1 a_{m1} + \cdots + v_n a_{mn}) e'_n, \end{aligned}$$

che prova che AX è il vettore delle componenti di $f(v)$ rispetto a $\mathcal{B}'$, quindi $c'^{-1}_{\mathcal{B}}(AX) = f(v)$.

Spesso si identifica f_A con f per abuso di notazione, poiché, come vedremo, vi è una corrispondenza biunivoca tra matrici ed applicazioni lineari indotta dalla scelta delle basi.

Lemma 4.2.

$$\forall\, X \quad FX = GX \quad \Rightarrow \quad F = G.$$

Dimostrazione. Basta far assumere ad X successivamente le coordinate dei vettori della base $\mathcal{B}$, oppure tener conto che per ogni X

$$(F - G)X = O \quad \Rightarrow \quad \mathrm{rg}(F - G) = 0 \quad \Rightarrow \quad F = G.$$

QED

Proposizione 4.4. *Siano V, W, Z spazi vettoriali con $\dim V = n$, $\dim W = m$, $\dim Z = p$ e $\mathcal{B}$, $\mathcal{B}'$, $\mathcal{B}''$ basi rispettive di V, W, Z. Se*

$$f \colon V \to W, \qquad g \colon W \to Z$$

sono funzioni lineari allora

$$h \overset{def}{=} g \circ f \colon V \to Z$$

è lineare e per le matrici associate vale

$$\mathcal{M}^{\mathcal{B}''}_{\mathcal{B}}(h) = \mathcal{M}^{\mathcal{B}''}_{\mathcal{B}'}(g) \cdot \mathcal{M}^{\mathcal{B}'}_{\mathcal{B}}(f).$$

Dimostrazione. Ponendo

$$\mathcal{M}^{\mathcal{B}''}_{\mathcal{B}}(h) = C \in \mathbb{K}^{p,n}, \quad \mathcal{M}^{\mathcal{B}''}_{\mathcal{B}'}(g) = B \in \mathbb{K}^{p,m}, \quad \mathcal{M}^{\mathcal{B}'}_{\mathcal{B}}(f) = A \in \mathbb{K}^{m,n},$$

allora

$$C = BA.$$

Infatti, usando un simbolismo compatto, posto

$$Y = AX, \quad Z = BY, \quad Z = CX,$$

si ha

$$Z = B(AX) = (BA)X \quad \Rightarrow \quad C = BA$$

tenendo conto del lemma 4.2. QED

Proposizione 4.5. *Siano V, W due spazi vettoriali di dimensione finita e sia $f\colon V \to W$ una funzione lineare. Siano $\mathcal{B} = \{e_i\}_{1\le i\le n}$ una base di V e $\mathcal{B}' = \{e'_j\}_{1\le j\le m}$ una base di W. Allora*

$$\dim(\operatorname{Im} f) = \operatorname{rg} \mathcal{M}_{\mathcal{B}}^{\mathcal{B}'}(f).$$

Ne segue $\dim(\operatorname{Ker} f) = \dim V - \operatorname{rg} \mathcal{M}_{\mathcal{B}}^{\mathcal{B}'}(f)$.

Dimostrazione. Infatti, $c_{\mathcal{B}'}(\operatorname{Im} f) = \operatorname{Im} f_A$, dove $A = \mathcal{M}_{\mathcal{B}}^{\mathcal{B}'}(f)$; siccome $c_{\mathcal{B}'}$ è un isomorfismo, si applica il corollario 4.1, dunque $\dim \operatorname{Im} f = \dim \operatorname{Im} f_A$. Inoltre, $Y \in \operatorname{Im} f_A$ ha la forma $Y = AX$, con

$$Y = AX = x_1A_1 + \cdots + x_nA_n \in \mathcal{L}(\{A_1, \ldots, A_n\}).$$

Pertanto $\operatorname{Im} f_A = \mathcal{L}(\{A_1, \ldots, A_n\})$. La tesi segue dal teorema 3.4, che implica

$$\dim \mathcal{L}(\{A_1, \ldots, A_n\}) = \operatorname{rg} A.$$

$\boxed{QED}$

Si osservi che se A è una matrice invertibile, allora f è un isomorfismo.

I risultati sin qui esposti possono essere riassunti dal seguente teorema.

Teorema 4.3. *Siano V, W, due spazi vettoriali di dimensione finita, con $\dim V = n$, $\dim W = m$. Sia $f\colon V \to W$ un'applicazione lineare. Siano $\mathcal{B} = \{e_i\}_{1\le i\le n}$ una base di V e $\mathcal{B}' = \{e'_j\}_{1\le j\le m}$ una base di W. Allora l'applicazione*

$$\mathcal{M}_{\mathcal{B}}^{\mathcal{B}'}\colon \operatorname{Lin}(V, W) \to \mathbb{K}^{m,n}, f \mapsto \mathcal{M}_{\mathcal{B}}^{\mathcal{B}'}(f)$$

è un isomorfismo.

Dimostrazione. È facile verificare che

$$\mathcal{M}_{\mathcal{B}}^{\mathcal{B}'}(f+g) = \mathcal{M}_{\mathcal{B}}^{\mathcal{B}'}(f) + \mathcal{M}_{\mathcal{B}}^{\mathcal{B}'}(g),$$
$$\mathcal{M}_{\mathcal{B}}^{\mathcal{B}'}(\lambda f) = \lambda\mathcal{M}_{\mathcal{B}}^{\mathcal{B}'}(f).$$

Inoltre, $\mathcal{M}_{\mathcal{B}}^{\mathcal{B}'}$ è iniettiva perché la matrice nulla è la corrispondente della sola applicazione lineare nulla tra V e W. Infine, $\mathcal{M}_{\mathcal{B}}^{\mathcal{B}'}$ è suriettiva poiché è stato visto che per ogni matrice A è possibile costruire un'applicazione lineare $f\colon V \to W$ tale che $\mathcal{M}_{\mathcal{B}}^{\mathcal{B}'}(f) = A$. $\boxed{QED}$

Esempi ed esercizi.

- Sia $f\colon \mathbb{R}^3 \to \mathbb{R}^3, (x, y, z) \mapsto (x + z, y, x + z)$. Allora la matrice di f rispetto alla base canonica $\mathcal{C}$ di $\mathbb{R}^3$ (in dominio e codominio) è

$$\mathcal{M}_{\mathcal{C}}^{\mathcal{C}}(f) = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}.$$

- Sia $D\colon \mathbb{R}_3[t] \to \mathbb{R}_3[t], t^h \mapsto ht^{h-1}$. Si consideri la base canonica $\mathcal{P} = \{1, t, t^2, t^3\}$. Allora
$$\mathcal{M}^{\mathcal{P}}_{\mathcal{P}}(D) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$
Si provi che D è nilpotente.

4.5 Cambiamenti di base

Sia V uno spazio vettoriale su $\mathbb{K}$ di dimensione n e siano
$$\mathcal{B} = \{e_1, \dots, e_n\}, \qquad \mathcal{B}' = \{e'_1, \dots, e'_n\}$$
due basi distinte (anche solo per l'ordine degli elementi). Allora
$$e'_k = \sum_{j=1}^{n} b_{jk} e_j, \qquad e_j = \sum_{r=1}^{n} c_{rj} e'_r.$$
Si vede subito che per le matrici $B = (b_{jk})$ e $C = (c_{rj})$ vale
$$C = B^{-1}.$$
La matrice B è detta *matrice del cambiamento di base* da $\mathcal{B}$ a $\mathcal{B}'$. Occupiamoci ora della legge di trasformazione delle coordinate di un vettore. Sia $x \in V$, e sia
$$x = \sum_i x_i e_i, \qquad x = \sum_i x'_i e'_i.$$
Conoscendo la matrice del cambiamento di base da $\mathcal{B}$ a $\mathcal{B}'$, qual è il legame tra le coordinate (x_i) e (x'_i)?
$$\begin{aligned} x = \sum_i x'_i e'_i = \sum_i x'_i \left(\sum_j b_{ji} e_j \right) &= \sum_j \left(\sum_i x'_i b_{ji} \right) e_j \\ \Rightarrow \quad x_j = \sum_i b_{ji} x'_i \quad &\Rightarrow \quad X = BX', \end{aligned}$$
e, naturalmente $X' = B^{-1}X$.

È possibile dare un'interpretazione della matrice del cambiamento di base come matrice di un'applicazione lineare. Infatti, si ha:
$$\mathrm{Id}(e'_k) = e'_k = \sum_{j=1}^{n} b_{jk} e_j,$$

dunque $B = \mathcal{M}^{\mathcal{B}}_{\mathcal{B}'}(\mathrm{Id})$; analogamente si prova che $C = \mathcal{M}^{\mathcal{B}'}_{\mathcal{B}}(\mathrm{Id})$. Per questo motivo, dalla proposizione 4.4 si ha

$$I = \mathcal{M}^{\mathcal{B}}_{\mathcal{B}}(\mathrm{Id}) = \mathcal{M}^{\mathcal{B}}_{\mathcal{B}'}(\mathrm{Id}) \cdot \mathcal{M}^{\mathcal{B}'}_{\mathcal{B}}(\mathrm{Id}) = \mathcal{M}^{\mathcal{B}'}_{\mathcal{B}}(\mathrm{Id}) \cdot \mathcal{M}^{\mathcal{B}}_{\mathcal{B}'}(\mathrm{Id}).$$

Osservazione 4.1. *Il cambio di base si può interpretare in due modi:*

1. *si 'vedono' gli stessi vettori rispetto a due basi (sistemi di riferimento) diversi; questo è riflesso nell'interpretazione dei cambiamenti di base come matrici associate all'identità;*

2. *si interpreta il cambio di base come una trasformazione di tutto lo spazio in sé stesso che porta i vettori di una base nei vettori di un'altra base.*

Questi due punti di vista sono detti punto di vista passivo *ed* attivo*, e ricorrono spesso nelle Scienze applicate quali la Fisica e l'Ingegneria.*

Siano V, W, due spazi vettoriali di dimensione finita, con $\dim V = n$, $\dim W = m$. Sia $f\colon V \to W$ un'applicazione lineare. Siano $\mathcal{B} = \{v_i\}_{1\leq i\leq n}$ una base di V e $\mathcal{C} = \{w_j\}_{1\leq j\leq m}$ una base di W. Allora, se $y = f(x)$ si ha

$$Y = AX, \quad \text{con} \quad A \stackrel{\mathrm{def}}{=} \mathcal{M}^{\mathcal{C}}_{\mathcal{B}}(f).$$

Consideriamo due nuove basi $\tilde{\mathcal{B}}$ di V e $\tilde{\mathcal{C}}$ di W. Allora, indicate con $\tilde{X}$ le coordinate rispetto a $\tilde{\mathcal{B}}$ e con $\tilde{Y}$ quelle rispetto a $\tilde{\mathcal{C}}$ si ha

$$X = B\tilde{X}, \qquad Y = C\tilde{Y},$$

quindi

$$C\tilde{Y} = A(B\tilde{X}) \quad \Rightarrow \quad \tilde{Y} = C^{-1}AB\tilde{X}.$$

Ma risulta $\tilde{Y} = \tilde{A}\tilde{X}$, con $\tilde{A} \stackrel{\mathrm{def}}{=} \mathcal{M}^{\tilde{\mathcal{B}}}_{\tilde{\mathcal{C}}}(f)$, quindi

$$\tilde{A} = C^{-1}AB.$$

Lo stesso risultato si può ottenere applicando la proposizione 4.4. Infatti, utilizzando la relazione

$$f = \mathrm{Id}_W \circ f \circ \mathrm{Id}_V$$

si ha

$$\mathcal{M}^{\tilde{\mathcal{C}}}_{\tilde{\mathcal{B}}}(f) = \mathcal{M}^{\tilde{\mathcal{C}}}_{\mathcal{C}}(\mathrm{Id})\mathcal{M}^{\mathcal{C}}_{\mathcal{B}}(f)\mathcal{M}^{\mathcal{B}}_{\tilde{\mathcal{B}}}(\mathrm{Id})$$

Nota. Poiché $\dim(\operatorname{Im} f)$ non dipende ovviamente dalla scelta delle basi, segue che

$$\mathrm{rg}(A) = \mathrm{rg}(\tilde{A}).$$

Nel caso in cui f è un endomorfismo, si può scegliere $\mathcal{B} = \mathcal{C}$ e $\tilde{\mathcal{B}} = \tilde{\mathcal{C}}$, quindi $B = C$ e

$$\tilde{A} = B^{-1}AB,$$

cioè A e $\tilde{A}$ sono matrici simili.

Resta così provato il seguente notevole teorema.

Teorema 4.4. *Matrici associate ad uno stesso endomorfismo rispetto a basi diverse sono simili.*

CAPITOLO 5

SISTEMI DI EQUAZIONI LINEARI

5.1 Sistemi di equazioni lineari e matrici

Definizione 5.1. *Un* sistema lineare *di m equazioni in n incognite $x_1, \dots, x_n$ è un sistema del tipo*

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

o, in forma più compatta,

$$\sum_{j=1}^{n} a_{ij}x_j = b_i\,, \qquad i = 1, 2, \dots, m\,,$$

dove a_{ij} sono detti coefficienti *e $b_i \in \mathbb{K}$* termini noti. *Se $b_i = 0$ il sistema si dice* omogeneo.

In forma matriciale:

$$AX = B \tag{5.1.1}$$

dove $A = (a_{ij}) \in \mathbb{K}^{m,n}$ è la matrice dei coefficienti, *X e B dati da*

$${}^tX = (x_1, \dots, x_n)\,, \qquad {}^tB = (b_1, \dots, b_m)\,.$$

sono, rispettivamente, il vettore delle incognite *e il* vettore dei termini noti. *La matrice $(A \mid B) \in \mathbb{K}^{m+1,n}$, formata aggiungendo ad A il vettore dei termini noti B, si dice* matrice completa *del sistema.*

Si dice soluzione del sistema (5.1.1) *una n-pla $(\tilde{x}_1, \dots, \tilde{x}_n)$ che soddisfa simultaneamente tutte le equazioni di* (5.1.1). *Si dice anche che il sistema è* compatibile *se esiste almeno una soluzione.*

Si dice sistema omogeneo associato ad $AX = B$ *il sistema $AX = O$.*

I problemi fondamentali che si presentano sono:

1. esistenza delle soluzioni o compatibilità del sistema (aspetto qualitativo);

2. determinazione del numero delle soluzioni (aspetto quantitativo);

3. calcolo esplicito di tutte le eventuali soluzioni (aspetto computazionale).

Si osservi che, introducendo l'applicazione lineare $f_A \colon \mathbb{K}^n \to \mathbb{K}^m$ tale che $f_A(X) = AX$ (equazione (4.4.4)), si ha

$$\operatorname{Ker} f_A = \{X \in \mathbb{K}^n \mid AX = 0\}, \quad f_A^{-1}(B) = \{X \in \mathbb{K}^n \mid AX = B\}. \tag{5.1.2}$$

Dunque il nucleo di f_A è uguale all'insieme delle soluzioni del sistema lineare omogeneo associato ad $AX = B$, mentre l'insieme delle soluzioni del sistema $AX = B$ è uguale a $f_A^{-1}(B)$.

Il problema dell'esistenza e del numero delle soluzioni è risolto completamente dal seguente *teorema di Rouché-Capelli*.

Teorema 5.1 (Rouché–Capelli)**.** *Siano $A \in \mathbb{K}^{m,n}$ e $B \in \mathbb{K}^{m,1}$. Allora il sistema di equazioni lineari $AX = B$ ha soluzioni se e solo se $\operatorname{rg} A = \operatorname{rg}(A \mid B)$. Se questo accade, l'insieme S delle soluzioni del sistema è*

$$S = X_0 + \operatorname{Ker} f_A = \{X + V \in \mathbb{K}^n \mid V \in \operatorname{Ker} f_A\},$$

dove X_0 è una particolare soluzione del sistema $AX = O$. Pertanto le soluzioni sono rappresentate mediante $n - p$ parametri, dove $p = \operatorname{rg} A$. Si dice anche che esistono ∞^{n-p} soluzioni.

Dimostrazione. Il sistema $AX = B$ è compatibile se e solo se $f_A^{-1}(B) \neq \varnothing$, equivalentemente $B \in \operatorname{Im} f_A$. Poniamo $\dim(\operatorname{Im} f_A) = \operatorname{rg}(A) = p$; si ha

$$\begin{aligned} AX = B \text{ compatibile} \quad &\Leftrightarrow \quad B \in \operatorname{Im} f_A \quad \Leftrightarrow \quad B \in \mathcal{L}(\{A_1, \dots, A_n\}) \\ &\Leftrightarrow \quad \{A_1, \dots, A_n, B\} \text{ dipendenti} \quad \Leftrightarrow \quad \operatorname{rg}(A \mid B) = \operatorname{rg} A = p. \end{aligned}$$

Supponiamo ora che il sistema $AX = B$ sia compatibile, o, equivalentemente, che $\operatorname{rg}(A \mid B) = \operatorname{rg} A = p$. Sia $\tilde{X} \in S$. Allora risulta $A\tilde{X} = B$. Sottraendo questa relazione con $AX_0 = B$ si ha $A(\tilde{X} - X_0) = 0$, dunque $\tilde{X} - X_0 \in \operatorname{Ker} f_A$. Pertanto, essendo $\tilde{X} = X_0 + (\tilde{X} - X_0)$, si ha $X \in X_0 + \operatorname{Ker} f_A$. Quindi $S \subset X_0 + \operatorname{Ker} f_A$. Viceversa, se $\tilde{X} = X + V \in X_0 + \operatorname{Ker} f_A$, si ha $A\tilde{X} = AX_0 + AV = B + O = B$, dunque $\tilde{X} \in S$. Quindi $S \supset X_0 + \operatorname{Ker} f_A$. Concludendo, $S = X_0 + \operatorname{Ker} f_A$.

Dal teorema fondamentale dell'algebra lineare si ha

$$\dim \mathbb{R}^n = \dim \operatorname{Ker} f_A + \dim \operatorname{Im} f_A = \dim \operatorname{Ker} f_A + \operatorname{rg} A,$$

dunque $\dim \operatorname{Ker} f_A = n - \operatorname{rg} A = n - p$. Le soluzioni dipendono dalle $n - p$ coordinate dei vettori $V \in \operatorname{Ker} f_A$. QED

Esempi.

1. I sistemi *omogenei*, ossia sistemi del tipo

$$AX = O, \tag{5.1.3}$$

ammettono *sempre* la soluzione nulla $X = O$ (ed è l'unica se $|A| \neq 0$), dunque sono sempre compatibili. I sistemi omogenei possono ammettere anche soluzioni non nulle, ovviamente.

2. Il sistema

$$\begin{cases} x+y=0 \\ x+y=1 \end{cases}, \qquad \text{con } A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \quad \tilde{A} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}$$

è chiaramente incompatibile. Infatti $1 = \operatorname{rg}(A) \neq \operatorname{rg}(\tilde{A}) = 2$.

Descriviamo ora lo spazio S delle soluzioni di $AX = B$. Presi due vettori Y, $Y' \in S = X_0 + \operatorname{Ker} f_A$ si ha

$$A(Y+Y') = AY + AY' = B + B = 2B,$$

pertanto S è uno spazio vettoriale se e solo se $2B = B$, che è vero se e solo se $B = O$. In questo caso risulta $\S = \operatorname{Ker} f_A$.

Più in generale, se U è un sottospazio vettoriale di V ed a un vettore fissato di V, lo spazio

$$S = a + U$$

è detto *varietà lineare* o *spazio affine*. Poniamo per definizione

$$\dim S \stackrel{\text{def}}{=} \dim U.$$

Ad esempio, le rette ed i piani passanti per l'origine sono sottospazi vettoriali di $\mathbb{R}^3$, tutte le rette ed i piani sono sottospazi affini di dimensione 1, 2 di $\mathbb{R}^3$.

Infine, chiamiamo *applicazione affine* un'applicazione

$$\varphi \colon V \to W, \; \varphi = a + f, \qquad a \in W,$$

dove $f \colon V \to W$ è un'applicazione lineare.

Esempio. L'applicazione $\varphi \colon \mathbb{R}^2 \to \mathbb{R}^2, (x,y) \mapsto (x',y')$ dove

$$\begin{cases} x' = 2x - y + 1 \\ y' = 3x + y + 2 \end{cases}$$

è un'applicazione affine. Infatti $\varphi = t_{\vec{a}} \circ f$, dove $\vec{a} = (1,2)$ e $f \colon \mathbb{R}^2 \to \mathbb{R}^2$ tale che $f(x,y) = (2x-y, 3x+y)$. Essendo f un isomorfismo, φ può considerarsi un cambiamento di riferimento affine da $\mathcal{RA}(Oxy)$ a $\mathcal{RA}(O'x'y')$. Il cambiamento di riferimento da $\mathcal{RA}(O'x'y')$ a $\mathcal{RA}(Oxy)$ si fa invertendo φ, cioè

$$\varphi^{-1} \colon \begin{cases} x = \frac{1}{5}x' + \frac{1}{5}y' - \frac{3}{5} \\ y = -\frac{3}{5}x' + \frac{2}{5}y' - \frac{1}{5}. \end{cases}$$

Per quanto riguarda il problema del calcolo delle soluzioni, iniziamo con il seguente teorema.

Teorema 5.2 (Teorema di Cramer)**.** *Sia $AX = B$ un sistema di n equazioni in n incognite, e sia* $\det A \neq 0$*. Allora il sistema ammette un'unica soluzione $(x_1, \ldots, x_p)$, le cui componenti sono date da*

$$x_k = \frac{|A^{(k)}|}{|A|},$$

dove $A^{(k)}$ è la matrice ottenuta da A sostituendo alla k-esima colonna di A la colonna dei termini noti.

Il metodo di risoluzione di Cramer ha interesse prevalentemente teorico, poiché bisogna conoscere preliminarmente determinante di A. In generale non conviene effettuare questo calcolo, conviene utilizzare altri metodi, primo fra tutti il metodo di eliminazione di Gauss.

Il *metodo di eliminazione di Gauss* prescinde dalla conoscenza a priori di quei ranghi ed inoltre è facilmente programmabile sui calcolatori; perciò è comunemente usato nelle applicazioni. Esso si basa sull'idea di ricondurre il sistema lineare di partenza ad un sistema lineare che ha le stesse soluzioni ma che è più semplice da risolvere, ed è descritto nel teorema 5.3.

Definizione 5.2. *Siano $A \in \mathbb{K}^{m,n}$, $A' \in \mathbb{K}^{m',n}$ e $B \in \mathbb{K}^{m,1}$, $B' \in \mathbb{K}^{m',1}$. Allora i sistemi di equazioni lineari $AX = B$ e $A'X = B'$ si dicono* equivalenti *se hanno le stesse soluzioni.*

Si noti che due sistemi equivalenti hanno le stesse incognite ma possono avere equazioni completamente differenti, anche nel numero.

Proposizione 5.1. *Sia $A \in \mathbb{K}^{m,n}$, $B \in \mathbb{K}^{m,1}$. Allora le seguenti trasformazioni su $(A \mid B)$ dànno una matrice $(A' \mid B')$ tale che il sistema lineare $AX = B$ è equivalente al sistema lineare $A'X = B'$:*

1. *scambio di due righe;*
2. *moltiplicazione di una riga per un elemento $k \in \mathbb{K}^* = \mathbb{K} \setminus \{0\}$;*
3. *sostituzione di una riga con la somma della riga stessa con un'altra riga di A moltiplicata per un elemento $k \in \mathbb{K}$.*

Dimostrazione. Infatti, scambiare due righe equivale a cambiare l'ordine delle equazioni del sistema, operazione che non ne cambia le soluzioni. Anche la moltiplicazione di una riga per una costante non nulla non aggiunge o toglie soluzioni al sistema.

Si consideri ora il sistema $A'X = B'$ dove $A = {}^t(A^1, \ldots, A^i + kA^j, \ldots, A^m)$ e $B = {}^t(B^1, \ldots, B^i + kB^j, \ldots, B^m)$. Siano S, S' gli insiemi delle soluzioni di $AX = B$, $A'X = B'$, rispettivamente. Sia $X \in S$. Allora la riga i e la riga j del sistema $AX = B$ sono equazioni verificate da X: $A^iX = B^i$ e $A^jX = B^j$. Dunque $(A^i + kA^j)X = B^i + kB^j$, quindi $X \in S'$. Viceversa sia $X \in S'$. Allora la riga i e la riga j del sistema $A'X = B'$ sono equazioni verificate da X: $(A^i + kA^j)X = B^i + kB^j$ e $A^jX = B^j$. Dunque, moltiplicando la seconda condizione per k e sottraendola dalla prima si ha $A^iX = B^i$, pertanto $X \in S$. Quindi $S = S'$. $\boxed{QED}$

Attenzione: per i sistemi di equazioni lineari *non è vero* che operare per righe o per colonne è la stessa cosa. Operare per colonne significa modificare le incognite del sistema, che è un'operazione che ne modifica le soluzioni.

Teorema 5.3. *Siano $A \in \mathbb{K}^{m,n}$ e $B \in \mathbb{K}^{m,1}$. Allora il sistema di equazioni lineari $AX = B$ è equivalente al sistema di equazioni lineari $SX = B'$, dove le matrici S e $(S \mid B')$ sono a scalini.*

Dimostrazione. Infatti, le trasformazioni elementari sulle righe possono essere usate per trasformare la matrice $(A \mid B)$ nella matrice a scalini $(S \mid B)$ per quanto dimostrato nella proposizione 5.1. Ovviamente, anche la matrice S sarà a scalini. $\boxed{QED}$

Una volta ridotto il sistema nella forma a scalini si possono presentare le seguenti situazioni:

1. S ha p scalini non nulli e $(S \mid B')$ ha $p+1$ scalini non nulli: questo implica che c'è un'equazione nel sistema della forma $0 = B^{p+1}$, dove $B^{p+1} \neq 0$. Pertanto il sistema è incompatibile, essendo presente una condizione mai verificata. Questa condizione è anche responsabile del fatto che $p = \operatorname{rg} S \neq \operatorname{rg}(S \mid B') = p+1$.

2. S e $(S \mid B')$ hanno p scalini non nulli. Il teorema di Rouché–Capelli assicura che il sistema ha soluzioni.

Se il sistema è compatibile, a partire dal sistema equivalente $(S \mid B')$ le soluzioni si determinano nel modo seguente: si individuino le p colonne $S_{j_1}, \ldots, S_{j_p}$ di S che costituiscono il minore di S con determinante diverso da zero, secondo il lemma 3.1. Le restanti colonne di S, $S_{k_1}, \ldots, S_{k_{n-p}}$ corrispondono ad altrettanti parametri che descrivono l'insieme delle soluzioni. Il sistema di equazioni lineari

$$(S_{j_1}, \ldots, S_{j_p})^t(x_{j_1}, \ldots, x_{j_p}) = -(S_{k_1}, \ldots, S_{k_{n-p}})^t(x_{k_1}, \ldots, x_{k_{n-p}}) + B' \qquad (5.1.4)$$

ha esattamente una soluzione per ogni valore attribuito ai parametri $x_{k_1}, \ldots, x_{k_{n-p}}$.

Si noti che il sistema (5.1.4) ha matrice triangolare superiore. Il sistema si può risolvere *all'indietro*: l'ultima equazione dipende da una sola incognita, che può essere ricavata e sostituita nelle equazioni precedenti, rendendo così la matrice diagonale.

Esempi.

1. Si risolva il seguente sistema

$$\begin{cases} 2x + y + 2z = 3, \\ -y - 5z = 1, \\ x + y = 0. \end{cases}$$

In notazione matriciale, si ha:

$$\left(\begin{array}{ccc|c} 2 & 1 & 2 & 3 \\ 0 & -1 & -5 & 1 \\ 1 & 1 & 0 & 0 \end{array}\right) \to \left(\begin{array}{ccc|c} 2 & 1 & 2 & 3 \\ 0 & -1 & -5 & 1 \\ 0 & 1/2 & -1 & -3/2 \end{array}\right) \to \left(\begin{array}{ccc|c} 2 & 1 & 2 & 3 \\ 0 & -1 & -5 & 1 \\ 0 & 0 & -7/2 & -1 \end{array}\right) \to$$

$$\to \left(\begin{array}{ccc|c} 2 & 1 & 0 & 17/7 \\ 0 & -1 & 0 & 17/7 \\ 0 & 0 & 1 & 2/7 \end{array}\right) \to \left(\begin{array}{ccc|c} 1 & 0 & 0 & 17/7 \\ 0 & 1 & 0 & -17/7 \\ 0 & 0 & 1 & 2/7 \end{array}\right),$$

dunque il sistema ammette l'unica soluzione $(17/7, -17/7, 2/7)$.

2. Si consideri il seguente sistema

$$\begin{cases} x + y + z + t = 1, \\ x + y + 2z = 3, \\ 2x + 2y + 3z + t = 7. \end{cases}$$

La matrice del sistema è

$$\left(\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 2 & 0 & 3 \\ 2 & 2 & 3 & 1 & 7 \end{array}\right).$$

Operando per righe

$$\left(\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 2 & 0 & 3 \\ 2 & 2 & 3 & 1 & 7 \end{array}\right) \to \left(\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & -1 & 2 \\ 0 & 0 & 1 & -1 & 5 \end{array}\right) \to \left(\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & -1 & 2 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right).$$

Quindi, la matrice incompleta ha rango 2 mentre la matrice completa ha rango 3 (si prenda la sottomatrice costituita dalle colonne 2, 3 e 5). Pertanto, per il teorema di Rouché–Capelli, il sistema non ha soluzioni. Questo si può comunque dedurre dalla presenza della condizione impossibile $0 = 1$ nella terza riga.

3. Si consideri il seguente sistema

$$\begin{cases} 2x - y + 3z = 4, \\ x + 2z = 3, \\ 3x - 2y + 4z = 5. \end{cases}$$

La matrice associata al sistema lineare è

$$\left(\begin{array}{ccc|c} 2 & -1 & 3 & 4 \\ 1 & 0 & 2 & 3 \\ 3 & -2 & 4 & 5 \end{array}\right).$$

Operando per righe si ha

$$\left(\begin{array}{ccc|c} 2 & -1 & 3 & 4 \\ 1 & 0 & 2 & 3 \\ 3 & -2 & 4 & 5 \end{array}\right) \to \left(\begin{array}{ccc|c} 2 & -1 & 3 & 4 \\ 0 & 1/2 & 1/2 & 1 \\ 0 & -1/2 & -1/2 & -1 \end{array}\right),$$

dove nell'ultima matrice la terza riga si può eliminare in quanto proporzionale alla seconda. Si ottiene quindi

$$\left(\begin{array}{ccc|c} 2 & -1 & 3 & 4 \\ 0 & 1 & 1 & 2 \end{array}\right) \to \left(\begin{array}{ccc|c} 2 & 0 & 4 & 6 \\ 0 & 1 & 1 & 2 \end{array}\right).$$

Pertanto, sia la matrice incompleta che la matrice completa del sistema hanno rango 2. Secondo il teorema di Rouché–Capelli, il sistema ammette quindi ∞^1 soluzioni. Una delle variabili è, dunque, arbitraria: diventa un *parametro*. Scegliamo come parametro z e poniamo, per questo, $z = k$. Le soluzioni sono tutti e soli i vettori del tipo $(3, 2, 0) + k(-2, -1, 1)$, con $k \in \mathbb{R}$.

Appendice: variabili, incognite e parametri. È necessario stabilire la terminologia usata per le *quantità variabili*. Queste quantità sono indicate da una lettera, e rappresentano uno o più elementi in un fissato insieme V, detto *spazio delle variabili*. Le variabili possono essere di due tipi:

- le **incognite**, che sono variabili soggette ad una o più *condizioni* (di solito *equazioni*);
- i **parametri**, che sono variabili *non* soggette a condizioni, il cui valore può essere uno degli elementi di V ad arbitrio.

Un sottoinsieme $S \subset V$ può essere rappresentato in uno dei seguenti modi:

- i suoi elementi possono essere tutti e soli gli elementi di V soddisfacenti una o più condizioni (di solito equazioni); questa si dice (impropriamente) **rappresentazione cartesiana** di S;
- i suoi elementi possono essere noti in funzione di uno o più parametri; questa si dice **rappresentazione parametrica** di S.

Esercizi.

- Discutere il seguente sistema, al variare di $\lambda \in \mathbb{R}$, e risolverlo nei casi in cui è compatibile

$$\begin{cases} x - y = 1 \\ \lambda y + z = 0 \\ 2x - \lambda z = -1 \\ x + y + z = 1 \end{cases}$$

- Determinare il polinomio

$$P(x) = ax^3 + bx^2 + cx + d$$

di grado ≤ 3 tale che

$$P(0) = 1, \quad P(1) = -2, \quad P(-1) = -6, \quad P(2) = 3.$$

Imponendo queste condizioni si ha

$$\begin{cases} d = 1 \\ a + b + c + d = -2 \\ -a + b - c + d = -6 \\ 8a + 4b + 2c + d = 3 \end{cases}$$

che ha un'unica soluzione: $(a, b, c, d) = (3, -5, -1, 1)$, e quindi $P(x) = 3x^3 - 5x^2 - x + 1$. Si verifichi che P soddisfa le richieste precedenti!

- Risolvere il sistema

$$\begin{cases} x_1 + x_2 + x_3 + 2x_4 = 5 \\ 2x_1 + 2x_2 + x_3 + 3x_4 = 8 \\ 4x_1 + 4x_2 + 3x_3 + 7x_4 = 18 \\ x_1 + x_2 + x_4 = 3 \end{cases}$$

CAPITOLO 6

AUTOVALORI ED AUTOVETTORI

Questa sezione è dedicata allo studio di alcune proprietà degli endomorfismi che sono indipendenti dalla scelta di una particolare base. Tali proprietà rivestono notevole importanza nelle applicazioni perché corrispondono a proprietà fisiche del modello studiato.

6.1 Definizioni

Sia V uno spazio vettoriale sul campo $\mathbb{K}$ e

$$f : V \to V$$

un endomorfismo. Vogliamo vedere se esistono vettori $x_0 \neq 0$ tali che

$$f(x_0) = \lambda x_0, \qquad \lambda \in \mathbb{K},$$

cioè vettori che dall'applicazione f vengono mutati di grandezza o di verso, ma non di direzione.

Definizione 6.1. *Un elemento $\lambda \in \mathbb{K}$ è detto* autovalore *(o* valore proprio*, o* valore caratteristico*) di f se esiste un vettore* <u>*non nullo*</u> *x tale che*

$$f(x) = \lambda x.$$

In tal caso, il vettore x si dice autovettore *(o* vettore proprio*, o* vettore caratteristico*) di f associato all'autovalore λ.*

Naturalmente, se $\lambda = 0$, allora $x \in \operatorname{Ker} f$. Si noti anche come la definizione perda di significato se si pone $x = 0$.

Lo studio degli autovalori è uno degli argomenti più importanti sia dal punto di vista teorico sia per le applicazioni dell'Algebra Lineare alla Fisica, all'Ingegneria ed alla Chimica.

Infatti gli autovalori intervengono ogni volta che l'<u>effetto</u> $f(x)$ di una applicazione sia proporzionale alla <u>causa</u> x.

In Fisica gli autovalori corrispondono agli assi principali di f, cioè sono assi di rotazioni privilegiati. Se S è un solido, soggetto a forze, esso, a causa degli sforzi a cui è sottoposto, si rompe lungo direzioni che sono quelle degli autovettori.

Gli autovettori corrispondono anche alle frequenze critiche di una lamina vibrante, problema importante nella teoria dei microfoni. Si tratta, naturalmente, del fenomeno delle risonanze, che si incontra con nomi diversi in molte branche della scienza.

Se x è un autovettore per f, allora $\mathcal{L}(x)$ è invariante per f, cioè

$$f(\mathcal{L}(x)) \subset \mathcal{L}(x).$$

Se λ è un autovalore per f, allora si verifica facilmente che

$$V(\lambda) \stackrel{\text{def}}{=} \{x \in V \mid f(x) = \lambda x\} = \operatorname{Ker}(f - \lambda \operatorname{Id}_V)$$

è un sottospazio vettoriale di V, detto *autospazio*[1] di f relativo all'autovalore λ. Lo spazio $V(\lambda)$ è costituito dal vettore nullo e da tutti gli autovettori relativi a λ, quindi non può ridursi a 0, perciò

$$\dim V(\lambda) \geq 1.$$

Il numero naturale $\dim V(\lambda)$ è detto *molteplicità geometrica* di λ.

Proposizione 6.1. *Se $\lambda_1, \ldots, \lambda_p$ sono autovalori di f* distinti *e se $x_1, \ldots, x_p$ sono autovettori di f corrispondenti agli autovalori dati, allora i vettori $x_1, \ldots, x_p$ sono linearmente indipendenti.*

Dimostrazione. Per induzione, quando $p = 1$ la dimostrazione è banale.

Si consideri la combinazione lineare

$$\mu_1 x_1 + \cdots + \mu_p x_p = 0. \tag{6.1.1}$$

Applicando la f ad entrambi i membri si ha

$$f(\mu_1 x_1 + \cdots + \mu_p x_p) = 0 \quad \Rightarrow \quad \lambda_1 \mu_1 x_1 + \cdots + \lambda_p \mu_p x_p = 0.$$

Si consideri il prodotto $\lambda_1(\mu_1 x_1 + \cdots + \mu_p x_p) = 0$, e si calcoli la differenza tra questa identità e la precedente:

$$(\lambda_1 - \lambda_2)\mu_2 x_2 + \cdots + (\lambda_1 - \lambda_p)\mu_p x_p = 0.$$

Per l'ipotesi del metodo di induzione, essendo i vettori $(\lambda_1 - \lambda_i)x_i$ autovettori (non nulli!) relativi all'autovalore λ_i in quanto $\lambda_1 - \lambda_i \neq 0$ $(i = 2, \ldots, p)$, si ha $\mu_2 = 0, \ldots, \mu_p = 0$. Dunque, sostituendo nell'equazione (6.1.1) si ha $\mu_1 x_1 = 0$, che implica $\mu_1 = 0$ poiché $x_1 \neq 0$. $\boxed{QED}$

[1] Altri autori usano la notazione V_λ.

Corollario 6.1. *Nelle ipotesi della precedente proposizione, se $\lambda_i \neq \lambda_j$ allora*

$$V(\lambda_i) \cap V(\lambda_j) = \{0\}.$$

6.2 Polinomio caratteristico

Se $A \in \mathbb{K}^{n,n}$ è una matrice quadrata, chiamiamo *polinomio caratteristico* di A il polinomio

$$P_A(\lambda) \stackrel{\text{def}}{=} \det(A - \lambda \,\mathrm{Id}).$$

Nota 6.1. *Altri autori pongono come polinomio caratteristico*

$$\Delta_A(\lambda) \stackrel{def}{=} \det(\lambda \,\mathrm{Id} - A).$$

Allora $\Delta_A(\lambda) = (-1)^n P_A(\lambda)$.

Sviluppando il determinante per mezzo della formula (3.3.4) si vede che

$$P_A(\lambda) = c_n(A)\lambda^n + c_{n-1}(A)\lambda^{n-1} + \cdots + c_1(A)\lambda + c_0(A),$$

dove $c_h(A)$ sono funzioni degli elementi della matrice A:

$$\begin{aligned}
&c_n(A) = (-1)^n,\\
&c_{n-1}(A) = (-1)^{n-1}\operatorname{tr}(A) = (-1)^{n-1}\sum_{i=1}^{n} a_{ii},\\
&\ldots\\
&c_0(A) = \det(A).
\end{aligned}$$

Queste relazioni si dimostrano osservando che, se $A - \lambda I = (\tilde{a}_{ij})$, nella formula (3.3.4)si ha $\tilde{a}_{ij} = a_{ij}$ se $i \neq j$, e $\tilde{a}_{ii} = a_{ii} - \lambda$ se $i = j$. In particolare, per $\lambda = 0$ si ha $P_A(0) = \det A$. Essendo $c_n(A) \neq 0$, $P_A(\lambda)$ è un polinomio di grado n.

Si chiama *autovalore* della matrice A uno scalare $\lambda \in \mathbb{K}$ in corrispondenza del quale esiste un vettore non nullo $X \in \mathbb{K}^{n,1}$ tale che

$$AX = \lambda X.$$

Teorema 6.1. *Gli autovalori di A sono le radici in $\mathbb{K}$ dell'equazione $P_A(\lambda) = 0$.*

Dimostrazione.

$$AX = \lambda X \quad \Leftrightarrow \quad (A - \lambda \,\mathrm{Id})X = O.$$

Il sistema omogeneo ammette autosoluzioni se e solo se $\det(A - \lambda \,\mathrm{Id}) = 0$. $\boxed{QED}$

L'equazione $P_A(\lambda) = 0$ è detta *equazione caratteristica.* Se $\mathbb{K} = \mathbb{C}$, campo algebricamente chiuso, l'equazione $P_A(\lambda) = 0$ ha n radici in $\mathbb{C}$, pertanto A ha in $\mathbb{C}$ n autovalori contati con la loro molteplicità. Se $\mathbb{K} = \mathbb{R}$, gli autovalori di A sono le radici *reali* di $P_A(\lambda) = 0$. Se $\lambda \in \mathbb{K}$ è uno zero di $P_A(\lambda)$ di molteplicità k (si veda l'osservazione 1.1 per la definizione), si dice che λ è un autovalore di *molteplicità algebrica* k.

Nota. Per trovare gli autovalori di una matrice, troviamo gli zeri di un polinomio. Questa metodologia è in un certo senso 'reversibile'. Il calcolo degli zeri di un qualunque polinomio

$$p(t) = t^n + a_{n-1}t^{n-1} + \cdots + a_1 t + a_0$$

si può ricondurre al calcolo degli autovalori della 'matrice compagna' di p, cioè della matrice

$$A = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \\ -a_0 & -a_1 & -a_2 & \dots & -a_{n-1} \end{pmatrix}.$$

Infatti $\det(A - \lambda \,\mathrm{Id}) = -p(\lambda)$, dunque le radici di $p(t) = 0$ sono gli autovalori di A. Tuttavia, esistono infinite matrici che hanno il polinomio $p(t)$ come polinomio caratteristico.

Per alcune matrici è facile calcolare il polinomio caratteristico. Infatti, se $A \in \mathbb{K}^{n,n}$ è triangolare (superiore o inferiore), allora $\det(A - \lambda I) = (a_{11} - \lambda)\cdots(a_{nn} - \lambda)$. Gli autovalori sono così gli elementi sulla diagonale di A.

Teorema 6.2. *Se A ed A' sono due matrici simili, allora $P_A(\lambda) = P_{A'}(\lambda)$.*

Dimostrazione. Sia $A' = B^{-1}AB$. Allora

$$\begin{aligned} P_{A'}(\lambda) &= \det(A' - \lambda \,\mathrm{Id}) = \det(B^{-1}AB - \lambda B^{-1}B) \\ &= \det(B^{-1}(A - \lambda \,\mathrm{Id})B) = \det(A - \lambda \,\mathrm{Id}) = P_A(\lambda). \end{aligned}$$

QED

Si osservi che quanto sopra esposto implica che $\operatorname{tr} A$ è invariante per similitudini.

Teorema 6.3. *Sia V uno spazio vettoriale su $\mathbb{K}$ di dimensione finita e sia $f\colon V \to V$ un endomorfismo. Allora gli autovalori di f sono le radici in $\mathbb{K}$ dell'equazione $P_A(\lambda) = 0$, dove $A = \mathcal{M}^{\mathcal{B}}_{\mathcal{B}}(f)$ e $\mathcal{B}$ è una qualunque base di V.*

Dimostrazione.

$$x \text{ autovettore} \quad \Rightarrow \quad f(x) = \lambda x \quad \Rightarrow \quad (f - \lambda \,\mathrm{Id}_V)(x) = 0.$$

Fissata una base $\mathcal{B}$ di V, si ha $\mathcal{M}^{\mathcal{B}}_{\mathcal{B}}(f - \lambda \,\mathrm{Id}_V) = A - \lambda I$, dove $\mathcal{M}^{\mathcal{B}}_{\mathcal{B}}(f) = A$, dunque $(f - \lambda \,\mathrm{Id}_V)(x) = 0$ è rappresentato dal sistema omogeneo $(A - \lambda \,\mathrm{Id})X = O$.

Il polinomio caratteristico è invariante per matrici simili, pertanto il risultato non dipende dalla base $\mathcal{B}$ scelta. QED

Nota. Si può dimostrare che per trovare gli autovalori di un endomorfismo $f\colon V \to V$ si può anche operare con una matrice $\mathcal{M}^{\mathcal{C}}_{\mathcal{B}}(f)$, dove $\mathcal{C} \neq \mathcal{B}$, ma bisogna tener conto del fatto che la matrice di Id_V in $f - \lambda\,\mathrm{Id}_V$ non sarebbe uguale alla matrice identità, ma ad una matrice di cambiamento di base, cosa che complicherebbe notevolmente i calcoli.

Per i motivi sopra esposti si può definire polinomio caratteristico di f il polinomio caratteristico di una matrice di f rispetto ad una base qualunque.

Esempio. Trovare autovalori ed autospazi di

$$f\colon \mathbb{R}^3 \to \mathbb{R}^3,\ f(x,y,z) = (-y, x, 2z).$$

Considerando la base canonica $\mathcal{C}$ di $\mathbb{R}^3$ si ha

$$A = \mathcal{M}^{\mathcal{C}}_{\mathcal{C}}(f) = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \quad \det(A - \lambda\,\mathrm{Id}) = (2-\lambda)(\lambda^2+1) = 0.$$

Dunque A ammette come unico autovalore reale $\lambda = 2$. Troviamo ora l'autospazio $V(2) = \mathrm{Ker}(A - 2\,\mathrm{Id})$.

$$(A - 2\,\mathrm{Id})X = O \quad \Rightarrow \quad \begin{pmatrix} -2 & -1 & 0 \\ 1 & -2 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\Rightarrow \quad \begin{cases} -2x - y = 0 \\ x - 2y = 0 \end{cases} \quad \Rightarrow \quad V(2) = \{t(0,0,1) \mid t \in \mathbb{R}\} = \mathcal{L}(e_3).$$

Sia ora

$$f\colon \mathbb{C}^3 \to \mathbb{C}^3,\ f(x,y,z) = (-y, x, 2z).$$

Allora f ha i tre autovalori distinti $\lambda_1 = i$, $\lambda_2 = -i$, $\lambda_3 = 2$. Si vede facilmente che

$$\begin{aligned} V(\lambda_1) &= \mathcal{L}(u_1), \qquad & u_1 &= (i,1,0) \\ V(\lambda_2) &= \mathcal{L}(u_2), \qquad & u_2 &= (-i,1,0) \\ V(\lambda_3) &= \mathcal{L}(u_3), \qquad & u_3 &= (0,0,1). \end{aligned}$$

Si osservi che, in questo caso,

$$V(2) = \{t(0,0,1) \mid t \in \mathbb{C}\}.$$

I vettori $\{u_1, u_2, u_3\}$ costituiscono una base $\mathcal{B}$ per $\mathbb{C}^3$ rispetto alla quale la matrice di f è

$$\tilde{A} = \mathcal{M}^{\mathcal{B}}_{\mathcal{B}}(f) = \begin{pmatrix} i & 0 & 0 \\ 0 & -i & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

Si verifichi che $\tilde{A}$ è simile ad A e che $\det\tilde{A} = \det A$, $\mathrm{tr}\,\tilde{A} = \mathrm{tr}\,A$.

6.3 Endomorfismi semplici

Sia V uno spazio vettoriale di dimensione finita sul campo $\mathbb{K}$.

Un endomorfismo $f\colon V \to V$ si dice *semplice* se esiste una base di V costituita da autovettori per f.

Se $\tilde{\mathcal{B}} = \{u_1, \dots, u_n\}$ è una tale base, allora $f(u_i) = \lambda_i u_i$ e

$$\mathcal{M}^{\tilde{\mathcal{B}}}_{\tilde{\mathcal{B}}}(f) = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix},$$

cioè f ha una rappresentazione 'semplice' costituita da una matrice diagonale, Per questo si dice anche che f è *diagonalizzabile.*

Se f è semplice e $\mathcal{B} = \{e_1, \dots, e_n\}$ è un'altra base, allora, come è noto, $\mathcal{M}^{\mathcal{B}}_{\mathcal{B}}(f)$ è simile a $\mathcal{M}^{\tilde{\mathcal{B}}}_{\tilde{\mathcal{B}}}(f)$.

Una matrice $A \in \mathbb{K}^{n,n}$ è detta *diagonalizzabile* se esiste una matrice D diagonale simile ad A.

Si vede subito che non tutti gli endomorfismi sono semplici e quindi che non ogni matrice è diagonalizzabile. Per esempio, la matrice $A = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$ non è diagonalizzabile: si vede subito che non esiste una matrice invertibile P tale che $O = P^{-1}AP$ e O è la matrice che ha sulla diagonale gli autovalori di A, ossia la matrice nulla.

Proposizione 6.2. *Dato uno spazio vettoriale V su $\mathbb{K}$, e dato un endomorfismo $f\colon V \to V$, sia $\lambda_0 \in \mathbb{K}$ un autovalore di f. Allora si ha*

$$1 \leq \dim V(\lambda_0) \leq m_a(\lambda_0),$$

dove $m_a(\lambda_0)$ è la molteplicità algebrica di λ_0 come radice del polinomio caratteristico $P(\lambda)$ di f.

Dimostrazione. La disuguaglianza $1 \leq \dim V(\lambda_0)$ è già stata dimostrata.

Si ricordi che la molteplicità algebrica è stata introdotta nell'osservazione 1.1.

Sia $\dim V(\lambda_0) = s_0$. Sia $\mathcal{B}' = \{v_1, \dots, v_{s_0}\}$ una base di $V(\lambda_0)$. Si completi $\mathcal{B}'$ ad una base $\mathcal{B}$ di V. Rispetto tale base si ha

$$\mathcal{M}^{\mathcal{B}}_{\mathcal{B}}(f) = A \begin{pmatrix} \lambda_0 I & B \\ O & D \end{pmatrix}$$

dove $B \in \mathbb{K}^{s_0, n-s_0}$, $D \in \mathbb{K}^{n-s_0, n-S_0}$ ed $O \in \mathbb{K}^{n-s_0, s_0}$ è la matrice nulla. Sviluppando il determinante $\det(A - \lambda I)$ con le formule di Laplace rispetto alla prima colonna si ha $\det(A - \lambda I) = (\lambda_0 - \lambda)\det(A' - \lambda I)$, dove A' è ottenuta da A togliendo la prima colonna e la prima riga; continuando a sviluppare nello stesso modo si ottiene

$$\det(A - \lambda I) = (\lambda_0 - \lambda)^{s_0} \det(D - \lambda I).$$

Siccome λ_0 può essere radice di $\det(D - \lambda I)$ si ha $s_0 \leq m_a(\lambda_0)$. $\boxed{QED}$

In particolare se $P_f(\lambda)$ ha n zeri in $\mathbb{K}$ tutti distinti, appartenenti a $\mathbb{K}$, allora $m_i = 1$ $\forall i$, quindi f è semplice poiché $\dim V(\lambda_i) = m_i = 1$ per ogni i. D'altra parte se $\lambda_1, \ldots, \lambda_n$ sono distinti allora autovettori corrispondenti $u_1, \ldots, u_n$ sono indipendenti e quindi costituiscono una base. Si noti che l'avere n autovalori distinti è condizione solo **sufficiente** ma non **necessaria**. Ad esempio, O ed I sono certamente diagonalizzabili (essendo diagonali) ma gli autovalori sono tutti coincidenti.

Teorema 6.4 (Criterio di semplicità)**.** *L'endomorfismo $f \colon V \to V$ è semplice se e solo se*

1. *tutti gli zeri $\lambda_1, \ldots, \lambda_r$ di $P_f(\lambda)$ appartengono a $\mathbb{K}$;*

2. *$m_i = \dim V(\lambda_i)$ $\forall i = 1, \ldots, r \leq n$, dove m_i è la molteplicità algebrica (osservazione 1.1) dell'autovalore λ_i.*

DIMOSTRAZIONE. Infatti, se f è semplice e $\mathcal{B} - \{v_1, \ldots, v_n\}$ è una base di autovettori per f la matrice $A = \mathcal{M}^{\mathcal{B}}_{\mathcal{B}}(f)$ è diagonale, con gli autovalori sulla diagonale, in quanto $f(v_i) = \lambda_i v_i = 0v_1 + \cdots + \lambda_i v_i + \cdots + 0v_n$. Il polinomio caratteristico è $P_A(\lambda) = (\lambda - \lambda_1)^{m_1} \cdots (\lambda - \lambda_p)^{m_p}$, dove $\lambda_1, \ldots, \lambda_p$ sono gli autovalori distinti di f e $m_1, \ldots, m_p$ le rispettive molteplicità (si sviluppi il determinante secondo le colonne), quindi vale la prima affermazione. Per la seconda, si osservi che l'autovalore λ_i si ripete m_i volte sulle colonne, Si vede subito che $(A - \lambda_i I)X = 0$ ha come soluzioni $v_i, \ldots, v_{i+m_i}$: il sistema ha $n - m_i$ righe non nulle della forma $(\lambda_j - \lambda_i)x_j = 0$ per $i \neq j$, che implica $x_j = 0$ per $j \neq i$. Dunque $\dim V(\lambda_i) = m_i$.

Viceversa, per ogni i si costruisca una base $\{v_i, \ldots, v_{i+m_i}\}$ di $V(\lambda_i)$. Poiché $\sum_{i=1}^{p} m_i = n$ e poiché autovettori di autovalori distinti sono indipendenti (proposizione 6.1) l'insieme $\{v_1, \ldots, v_n\}$ risulta indipendente, dunque è una base. $\boxed{QED}$

Se f è semplice, indicato con $V(\lambda_i)$ l'autospazio relativo all'autovalore λ_i, si ha

$$V = V(\lambda_1) \oplus \cdots \oplus V(\lambda_r).$$

Una base di autovettori di V è ottenuta scegliendo in ogni autospazio una base. Da qui segue anche

$$f \text{ semplice} \quad \Leftrightarrow \quad \dim V = \dim V(\lambda_1) + \cdots + \dim V(\lambda_r).$$

Osservazione. Se A ed A' sono matrici simili, sappiamo che $P_A(\lambda) = P_{A'}(\lambda)$. Ma se $P_A(\lambda) = P_{A'}(\lambda)$, allora A ed A' non sono necessariamente simili. Le matrici

$$A = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \qquad A' = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

hanno lo stesso polinomio caratteristico (e quindi gli stessi autovalori) ma non sono simili. Infatti, se lo fossero, dovrebbe esistere una matrice invertibile $B = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ tale che $BA = A'B$. Da qui segue $a = 0 = b$, ma in tal caso $\det B = 0$, ossia B non è invertibile.

Osservazione. Se A è diagonalizzabile, è facile calcolare le potenze di A. Infatti

$$A = B^{-1}DB \quad \Rightarrow \quad A^k = B^{-1}D^kB,$$

e D^k è facile da calcolarsi essendo D diagonale.

Esempi ed esercizi.

- Sia $V = \mathbb{R}^3$ e $W = \mathcal{L}(v_1, v_2)$ dove $v_1 = (0, 3, 1)$ e $v_2 = (1, 4, 1)$.
  1. Provare che $f\colon W \to \mathbb{R}^3$ definita da
     $$f(v_1) = (-1, 5, 2), \qquad f(v_2) = (0, 6, 2),$$
     è un endomorfismo.
  2. Trovare gli autovalori di $f\colon W \to W$.

  1. Bisogna dimostrare che $f(W) \subset W$; basta che $f(v_1)$, $f(v_2) \in \mathcal{L}(v_1, v_2)$, cioè esistono $a, b, c, d \in \mathbb{R}$ tali che
     $$f(v_1) = av_1 + bv_2, \qquad f(v_2) = cv_1 + dv_2.$$
     Segue facilmente che $a = 3$, $b = -1$, $c = 2$, $d = 0$.
  2. Considerando $\mathcal{B} = \{v_1, v_2\}$ come base di W si ha
     $$\mathcal{M}^{\mathcal{B}}_{\mathcal{B}}(f) = \begin{pmatrix} 3 & 2 \\ -1 & 0 \end{pmatrix} = A,$$
     $$\det(A - \lambda\,\mathrm{Id}) = \lambda^2 - 3\lambda + 2 = 0 \quad \Rightarrow \quad \lambda = 1, 2.$$
     Dunque gli autovalori di $f\colon W \to W$ sono $\lambda = 1, 2$.

- Se $A \in \mathbb{K}^{n,n}$, provare che A e ${}^tA$ hanno gli stessi autovalori.

- Se $A, B \in \mathbb{K}^{n,n}$ ed A invertibile, provare che AB e BA hanno gli stessi autovalori (*Suggerimento: si osservi che* $A^{-1}(AB - \lambda I)A = BA - \lambda I$).

- Dimostrare che se λ è un autovalore per f, allora λ^k ($k \in \mathbb{N} \setminus \{0\}$) è un autovalore per f^k.

- Dimostrare che se f è un isomorfismo e λ un suo autovalore, allora $\lambda \neq 0$ e λ^{-1} è un autovalore per f^{-1}.

- Determinare gli autovalori di f nelle ipotesi $f^k = 0$ ed $f^k = \mathrm{Id}$.

CAPITOLO 7

SPAZI VETTORIALI EUCLIDEI

In questo capitolo sarà introdotta un'ulteriore struttura algebrica sugli spazi vettoriali. Questa struttura permette di introdurre il concetto di lunghezza di un vettore e di angolo tra due vettori in spazi vettoriali generici, generalizzando gli analoghi concetti già visti per $\mathbf{V}_3$.

Nota. In tutto il capitolo gli spazi vettoriali saranno di dimensione finita sul campo dei numeri reali $\mathbb{R}$.

7.1 Forme bilineari e forme quadratiche

In questa sezione sarà introdotto materiale preliminare alle strutture metriche.

7.1.a Forme bilineari

Sia V uno spazio vettoriale (di dimensione n sul campo $\mathbb{R}$). Un'applicazione lineare $f\colon V \to \mathbb{R}$ si dice *forma lineare*. Lo spazio vettoriale

$$V^* \stackrel{\text{def}}{=} \operatorname{Lin}(V, \mathbb{R})$$

è detto *duale* di V. Per il teorema 4.3 si ha $\dim V^* = n \cdot 1 = n$, quindi $V^* \cong V$.

Definizione 7.1. *Sia V uno spazio vettoriale,* $\dim V = n$. *Un'applicazione*

$$\beta\colon V \times V \to \mathbb{R}$$

si dice bilineare *se è lineare in entrambi gli argomenti, cioè*

$$\begin{aligned} \beta(x + xp, y) &= \beta(x, y) + \beta(xp, y), & \beta(\lambda x, y) &= \lambda\beta(x, y),\\ \beta(x, y + \vec{y'}) &= \beta(x, y) + \beta(x, \vec{y'}), & \beta(x, \lambda y) &= \lambda\beta(x, y), \end{aligned}$$

per ogni $x, xp, y, \vec{y'} \in V$ *e* $\lambda \in \mathbb{R}$.

Una forma bilineare β è detta *simmetrica* se

$$\beta(x,y) = \beta(y,x) \qquad \forall x,y \in V;$$

è detta *antisimmetrica* o *alternante* se

$$\beta(x,y) = -\beta(y,x) \qquad \forall x,y \in V.$$

In tal caso $\beta(x,x) = 0$ per ogni $x \in V$.

Esempi. In $\mathbb{R}^2$ siano $x = (x_1, x_2)$ e $y = (y_1, y_2)$.

- $\beta\colon \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$, $\beta(x,y) = x_1y_1 - x_2y_2$ è una forma bilineare simmetrica.
- $\beta\colon \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$, $\beta(x,y) = \left|\begin{smallmatrix} x_1 & x_2 \\ y_1 & y_2 \end{smallmatrix}\right|$ è una forma bilineare antisimmetrica.

Se $\mathcal{B} = \{e_1, \dots, e_n\}$ è una base di V, posto

$$a_{ij} \stackrel{\text{def}}{=} \beta(e_i, e_j) \in \mathbb{R}, \qquad (i,j = 1, \dots, n),$$

si dice che la matrice $A = (a_{ij})$ è la matrice associata a β tramite $\mathcal{B}$; in simboli

$$\mathcal{M}_{\mathcal{B}}(\beta) \stackrel{\text{def}}{=} (\beta(e_i, e_j)).$$

Si noti che, a rigore, si potrebbero impiegare due basi per rappresentare β, una per ogni argomento. Qui, per semplicità, ne sarà usata solo una.

La conoscenza di $\mathcal{M}_{\mathcal{B}}(\beta)$ permette di calcolare $\beta(x,y)$ per ogni $x, y \in V$. Infatti, se

$$x = \sum_i x_i e_i, \qquad y = \sum_j y_j e_j,$$

allora

$$\beta(x,y) = \beta(\sum_i x_i e_i, \sum_j y_j e_j) = \sum_i \sum_j x_i y_j \beta(e_i, e_j) = \sum_{ij} a_{ij} x_i y_j = {}^tXAY,$$

dove ${}^tX = (x_1, \dots, x_n)$ e analogamente per Y.

Si noti che

$$\begin{aligned} \beta \text{ simmetrica} \quad &\Leftrightarrow \quad \mathcal{M}_{\mathcal{B}}(\beta) \text{ simmetrica}, \\ \beta \text{ antisimmetrica} \quad &\Leftrightarrow \quad \mathcal{M}_{\mathcal{B}}(\beta) \text{ antisimmetrica}. \end{aligned}$$

Se cambiamo la base come cambia la matrice associata a β?

Se $\mathcal{B}' = \{e'_1, \dots, e'_n\}$ è una nuova base allora $\beta(x,y) = {}^tX'A'Y'$. Indicata con B la matrice (invertibile) che esprime il cambiamento di base da $\mathcal{B}$ a $\mathcal{B}'$ si ha $X = BX'$, $Y = BY'$, quindi

$${}^tXAY = {}^t(BX')A(BY') = {}^tX'({}^tBAB)Y',$$

da cui, tenendo conto del lemma 4.2,

$$A' = {}^tBAB.$$

Matrici A e A' legate dalla relazione precedente si dicono *congruenti*. Si verifica facilmente che la relazione di congruenza tra matrici è una relazione di equivalenza, e che se A è simmetrica (antisimmetrica), allora A' è simmetrica (antisimmetrica). Inoltre, se A e A' sono congruenti, allora $\det A' = (\det A)(\det B)^2$, quindi si conserva il segno del determinante.

Poiché $0x = 0$, si verifica facilmente che

$$\beta(0, y) = \beta(0x, y) = 0\beta(x, y) = 0 \qquad \forall y \in V,$$

e analogamente $\beta(x, 0) = 0$. Non vale il viceversa:

$$\forall y \in V \quad \beta(x, y) = 0 \quad \not\Rightarrow \quad x = 0.$$

Se

$$\forall y \in V \quad \beta(x, y) = 0 \quad \Leftrightarrow \quad x = 0$$

allora β si dice *non degenere*. Ovviamente, se $\mathcal{B} = \{e_1, \dots, e_n\}$, allora

$$\beta \text{ degenere} \quad \Leftrightarrow \quad \exists\, x \neq 0 : \ \beta(x, e_i) = 0 \quad i = 1, \dots, n.$$

Dunque, se $x = (x_1, \dots, x_n)$ rispetto a $\mathcal{B}$, allora

$$\sum_{i=1}^{n} x_i \beta(e_i, e_j) = 0 \quad \Rightarrow \quad \sum_{i=1}^{n} a_{ij} x_i = 0,$$

e cioè

$$\beta \textit{ non} \text{ degenere} \quad \Leftrightarrow \quad \det A \neq 0.$$

Se A e A' sono matrici congruenti, si prova che $\mathrm{rg}(A) = \mathrm{rg}(A')$, quindi possiamo definire $\mathrm{rg}(\beta) \stackrel{\text{def}}{=} \mathrm{rg}(A)$, dove $A = \mathcal{M}_{\mathcal{B}}(\beta)$ rispetto ad una base qualsiasi. Se $\mathrm{rg}(A) < n$, allora β è degenere.

Esempio. Si consideri la forma bilineare simmetrica

$$\beta \colon \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}, \ \beta(x, y) = x_1 y_1 - x_2 y_2$$

La matrice $A = \mathcal{M}_{\mathcal{C}}(\beta)$ rispetto alla base canonica $\mathcal{C}$ è

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

Essendo $\det A = -1 \neq 0$, la forma bilineare β è non degenere.

7.1.b Forme quadratiche

Definizione 7.2. *Sia V uno spazio vettoriale (di dimensione n sul campo $\mathbb{R}$), e $\beta\colon V\times V\to\mathbb{R}$ una forma bilineare simmetrica. Si dice* forma quadratica associata *a β l' applicazione*

$$Q\colon V\to\mathbb{R},\quad Q(x)\stackrel{def}{=}\beta(x,x).$$

In una base si ha l'espressione

$$Q(x)=\beta(x,x)=\sum_{ij}a_{ij}x_ix_j={}^tXAX,$$

quindi la matrice di β può essere considerata anche la matrice di Q.

Q è detta *quadratica* perchè

$$Q(\lambda x)=\beta(\lambda x,\lambda x)=\lambda^2\beta(x,x)=\lambda^2Q(x).$$

Viceversa, data una forma quadratica Q si può costruire subito la forma bilineare simmetrica da cui proviene:

$$\beta(x,y)=\frac{1}{2}(Q(x+y)-Q(x)-Q(y)).$$

Infatti:

$$Q(x+y)=\beta(x+y,x+y)=\beta(x,x)+\beta(x,y)+\beta(y,x)+\beta(y,y).$$

L'insieme $\{x\in V\mid Q(x)=0\}$ si dice *cono isotropo* relativo alla forma bilineare simmetrica β.

Si può provare che ogni forma bilineare simmetrica Q ammette una *forma canonica*,cioè una rappresentazione di Q mediante un polinomio omogeneo privo di termini misti. Più precisamente si dimostra che:

Teorema 7.1. *Siano V uno spazio vettoriale (di dimensione n sul campo $\mathbb{R}$), $\beta\colon V\times V\to\mathbb{R}$ una forma bilineare simmetrica e Q la corrispondente forma quadratica. Allora esiste una base $\mathcal{B}'=\{e'_1,\dots,e'_n\}$ rispetto alla quale $\beta(e'_i,e'_j)=0$ per $i\neq j$.*

Quindi, rispetto alla base $\mathcal{B}'$, la matrice $A'=\mathcal{M}_{\mathcal{B}'}(\beta)$ è diagonale del tipo

$$A'=\begin{pmatrix}\alpha'_1 & 0 & \dots\dots\dots\dots & & & \\ 0 & \alpha'_2 & \dots\dots\dots\dots & & & \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & \alpha'_p & \dots & 0\\ 0 & \dots & \dots & \dots & 0 & \dots\\ 0 & \dots & \dots & \dots & \dots & 0\end{pmatrix},$$

con $\alpha_i=\beta(e'_i,e'_i)\neq 0$ per $i=1,\dots,p$ e $\beta(e'_i,e'_i)=0$ se $i>p$. Quindi $\mathrm{rg}(\beta)=p$ e

$$Q(x)=\alpha'_1x_1^2+\dots+\alpha'_px_p^2.$$

Il teorema afferma che, considerata una matrice simmetrica A, esiste almeno una matrice invertibile B ad elementi reali tali che $A' = {}^tBAB$ sia diagonale.

Si osservi che la forma canonica non è unica. Se troviamo un'altra base $\mathcal{B}''$ per cui $\beta(\vec{e''}_i, \vec{e''}_j) = 0$ per $i \neq j$, allora la forma canonica indotta potrà avere coefficienti $\alpha''_i \neq \alpha'_i$, ma *resta costante il numero* dei coefficienti positivi, il numero dei coefficienti negativi, e quello dei coefficienti nulli (**teorema di Sylvester**). Anzi, quei coefficienti non nulli si possono scegliere in modo tale che $\alpha'_i = \pm 1$.

Innanzitutto ordiniamo i vettori di $\mathcal{B}'$ in modo che

$$\begin{aligned} &\alpha'_i > 0 && 1 \leq i \leq s, \\ &\alpha'_i < 0 && s+1 \leq i \leq p, \\ &\alpha'_i = 0 && p+1 \leq i \leq n. \end{aligned}$$

Allora, a partire da $\mathcal{B}'$ costruiamo una nuova base $\{\vec{E}_i\}$ in questo modo:

$$\begin{aligned} &\vec{E}_i \overset{\text{def}}{=} \frac{e'_i}{\sqrt{\alpha'_i}} && \text{se } \alpha'_i > 0 \\ &\vec{E}_i \overset{\text{def}}{=} \frac{e'_i}{\sqrt{-\alpha'_i}} && \text{se } \alpha'_i < 0 \\ &\vec{E}_i \overset{\text{def}}{=} e'_i && \text{se } \alpha'_i = 0 \end{aligned}$$

Rispetto alla nuova base la matrice associata a β è

$$\begin{pmatrix} \mathrm{Id}_s & O & O \\ O & -\mathrm{Id}_{p-s} & O \\ O & O & O_{n-p} \end{pmatrix}$$

dove $\mathrm{Id}_h \in \mathbb{R}^{h,h}$ è la matrice identità e O_{n-p} quella nulla. Il cambiamento di coordinate ha, dunque, la forma

$$\begin{aligned} &X_i = \sqrt{\alpha'_i} x'_i && \text{se } \alpha'_i > 0, \\ &X_i = \sqrt{-\alpha'_i} x'_i && \text{se } \alpha'_i < 0, \\ &X_i = x'_i && \text{se } \alpha'_i = 0. \end{aligned}$$

Si ha la *forma normale* (di Sylvester) di Q

$$Q(x) = X_1^2 + \cdots + X_s^2 - X_{s+1}^2 - \cdots - X_p^2.$$

Il numero s si dice *indice di positività* di Q, il numero $p - s$ *indice di negatività* di Q, il numero $n - p$ *indice di nullità* di Q. La coppia $(s, p - s)$ (o la terna $(s, p - s, n - p)$) si dice *segnatura* di Q. Una forma quadratica Q si dice

$$\begin{array}{lcl} \textit{semidefinita positiva} & \Leftrightarrow & Q(x) \geq 0 \ \forall x \in V \\ \textit{definita positiva} & \Leftrightarrow & Q(x) > 0 \ \forall x \in V \smallsetminus 0 \\ \textit{semidefinita negativa} & \Leftrightarrow & Q(x) \leq 0 \ \forall x \in V \\ \textit{definita negativa} & \Leftrightarrow & Q(x) < 0 \ \forall x \in V \smallsetminus 0 \\ \textit{indefinita} & \Leftrightarrow & \exists x, y \in V \ : Q(x) > 0, \ Q(y) < 0. \end{array}$$

Equivalentemente

$$\begin{array}{lcl} \textit{semidefinita positiva} & \Leftrightarrow & (s, p-s) = (p, 0), \ p < n, \\ \textit{definita positiva} & \Leftrightarrow & (s, p-s) = (n, 0) \\ \textit{semidefinita negativa} & \Leftrightarrow & (s, p-s) = (0, p), \ p < n \\ \textit{definita negativa} & \Leftrightarrow & (s, p-s) = (0, n) \\ \textit{indefinita} & \Leftrightarrow & (s, p-s), \ s \neq 0, \ p \neq s. \end{array}$$

In particolare, Q è non degenere se $p = n$.

Esempi.

- $\beta\colon \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$, $\beta(x, y) = x_1 y_1 + x_2 y_2$. Si ha $Q(x) = x_1^2 + x_2^2$, e

$$\mathcal{M}_{\mathcal{C}}(\beta) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

  quindi Q è definita positiva.

- $\beta\colon \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$, $\beta(x, y) = x_1 y_1 + x_2 y_2$. Si ha $Q(x) = x_1^2 + x_2^2$, dunque Q è semidefinita positiva, poiché $Q(v) = 0$ con $v = (0, 0, 1) \neq 0$.

- $\beta\colon \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$, $\beta(x, y) = x_1 y_1 - x_2 y_2$. Si ha $Q(x) = x_1^2 - x_2^2$, dunque Q è indefinita.

7.2 Prodotti scalari

In questa sezione sono oggetto di studio le proprietà dei prodotti scalari, cioè delle forme bilineari simmetriche definite positive. Questi oggetti generalizzano la nozione analoga introdotta nello spazio $\mathbf{V}_3$ ad uno spazio vettoriale arbitrario e permettono quindi la generalizzazione della geometria euclidea a spazi vettoriali arbitrari.

7.2.a Definizione

Definizione 7.3. *Sia V uno spazio vettoriale (di dimensione n) sul campo $\mathbb{R}$. Si chiama* prodotto scalare *o* metrica *su V un'applicazione*

$$g\colon V \times V \to \mathbb{R}$$

bilineare, simmetrica e definita positiva.

Se si usa la notazione più usuale $u \cdot v = g(u, v)$[1], allora le proprietà di g si traducono nelle seguenti:

1. distributività:
$$u \cdot (v + \vec{v'}) = u \cdot v + u \cdot \vec{v'}, \qquad (u + \vec{u'}) \cdot v = u \cdot v + \vec{u'} \cdot v;$$

2. omogeneità:
$$u \cdot (\lambda v) = \lambda(u \cdot v), \qquad (\lambda u) \cdot v = \lambda(u \cdot v);$$

3. commutatività: $u \cdot v = v \cdot u$;

4. essere def. positiva: $u \cdot u > 0$ per ogni $u \neq 0$, da cui $u \cdot u = 0 \Leftrightarrow u = 0$.

Definizione 7.4. *Si dice* spazio vettoriale euclideo *uno spazio vettoriale V sul campo $\mathbb{R}$, di dimensione finita, dotato di un prodotto scalare g.*

Nota 7.1. *Dalla proprietà 4 segue anche che*

$$\forall v \quad u \cdot v = 0 \quad \Rightarrow \quad u = 0;$$

per dimostrare questo fatto, si prenda $v = u$.

Esempi ed esercizi.

- $V = \mathbf{V}_3$, $u \cdot v \stackrel{\text{def}}{=} \|u\| \, \|v\| \cos \widehat{uv}$ è un prodotto scalare.
- $V = \mathbb{R}^n$, $u = (u_1, \ldots, u_n)$, $v = (v_1, \ldots, v_n)$, allora è un prodotto scalare
$$u \cdot v \stackrel{\text{def}}{=} u_1 v_1 + \cdots + u_n v_n$$
- $V = \mathbb{R}^2$, $g(u, v) \stackrel{\text{def}}{=} u_1 v_1 + 2u_1 v_2 + 5u_2 v_2$ non è un prodotto scalare.
- $V = \mathbb{R}^2$, $g(u, v) \stackrel{\text{def}}{=} u_1 v_1 - u_2 v_2$ non è un prodotto scalare.
- $V = \mathbb{R}^{n,n}$, $g(A, B) \stackrel{\text{def}}{=} \sum_{i,j=1}^{n} a_{ij} b_{ij}$ è un prodotto scalare.

[1] Altri autori usano la notazione $g(u, v) = \langle u, v \rangle$

- $V = \mathbb{R}^{n,n}$, $g'(A, B) \stackrel{\text{def}}{=} \operatorname{tr}(A^t B)$ è un prodotto scalare. Infatti g' è bilineare e
$$g'(B, A) = \operatorname{tr}(B\,{}^t\!A) = \operatorname{tr}({}^t(B\,{}^t\!A)) = \operatorname{tr}(A\,{}^t\!B) = g'(A, B).$$
Inoltre, $g'(A, A) = \sum_{ij} a_{ij}^2 \geq 0$ se $A \neq O$. Dimostrare che, se $n = 2$, $g = g'$.

Il numero
$$\|\vec{u}\|_g \stackrel{\text{def}}{=} \sqrt{g(u, u)} \geq 0$$
si dice *norma*, o *lunghezza*, o *modulo* di u rispetto a g. Si ha
$$\|\lambda u\|_g = \sqrt{g(\lambda u, \lambda u)} = |\lambda| \|u\|_g.$$
In poche parole,
$$\|u\|_g^2 = Q(u),$$
dove Q è la forma quadratica associata a g. Inoltre (paragrafo 7.1.b)
$$g(u, v) = \frac{1}{2}(\|u + v\|_g^2 - \|u\|_g^2 - \|v\|_g^2).$$

Si dice che due vettori $u, v \in V$ sono *ortogonali* (rispetto a g) se $g(u, v) = 0$, e si scrive $u \perp_g v$, o anche $u \perp v$. In tal caso, vale il seguente *teorema di Pitagora* per spazi vettoriali euclidei
$$\|u + v\|_g^2 = \|u\|_g^2 + \|v\|_g^2.$$
Se $u, v \in V$, poniamo
$$\operatorname{dist}_g(u, v) \stackrel{\text{def}}{=} \|u - v\|_g,$$
distanza di u *da* v (rispetto a g). Infatti, se $u = \vec{OP}$ e $v = \vec{OQ}$, allora $\operatorname{dist}_g(u, v) = \operatorname{dist}_g(P, Q) = \|P - Q\|_g$.

Teorema 7.2 (Disuguaglianza di Schwarz)**.** *Se V è uno spazio vettoriale euclideo con prodotto scalare g e $u, v \in V$, allora*
$$|g(u, v)| \leq \|u\|_g \|v\|_g,$$
$$|g(u, v)| = \|u\|_g \|v\|_g \quad \Leftrightarrow \quad u \parallel v.$$

Dimostrazione. Se $v = 0$ o $u = 0$, la disuguaglianza è banale. Se $u \neq 0$ e $v \neq 0$ si consideri $u + \lambda v$, per $\lambda \in \mathbb{R}$. Allora
$$\varphi(\lambda) = \|u + \lambda v\|_g^2 = \|u\|_g^2 + 2\lambda g(u, v) + \lambda^2 \|v\|_g^2 \geq 0.$$
Da proprietà delle disequazioni (o da semplici considerazioni geometriche sulle parabole) si deduce
$$g(u, v)^2 - \|u\|_g^2 \|v\|_g^2 \leq 0,$$
da cui la conclusione. $\boxed{QED}$

Dalla precedente disuguaglianza si ha il seguente teorema.

Teorema 7.3 (Disuguaglianza di Minkowski o triangolare)**.** *Se V è uno spazio vettoriale euclideo con prodotto scalare g e $u, v \in V$, allora*

$$\|u+v\|_g \leq \|u\|_g + \|v\|_g.$$

DIMOSTRAZIONE. Infatti

$$\begin{aligned}\|u+v\|_g^2 = g(u+v, u+v) = \|u\|_g^2 + 2g(u,v) + \|v\|_g^2 &\leq \|u\|_g^2 + 2|g(u,v)| + \|v\|_g^2 \\ \leq \|u\|_g^2 + 2\|u\|_g\,\|v\|_g + \|v\|_g^2 &= (\|u\|_g + \|v\|_g)^2.\end{aligned}$$

QED

Come sopra,

$$\|u+v\|_g = \|u\|_g + \|v\|_g \quad \Leftrightarrow \quad u = kv,\ k > 0.$$

Dalla disuguaglianza di Schwarz segue anche

$$\frac{g(u,v)}{\|u\|_g\,\|v\|_g} \leq \frac{|g(u,v)|}{\|u\|_g\,\|v\|_g} \leq 1,$$

quindi possiamo <u>definire</u> il *coseno dell'angolo* $\widehat{uv}$ (se $u, v \neq 0$) così

$$\cos\widehat{uv} \stackrel{\text{def}}{=} \frac{g(u,v)}{\|u\|_g\,\|v\|_g}.$$

Naturalmente la definizione è stata suggerita dal caso dei vettori geometrici dove il prodotto scalare è quello standard.

7.2.b Ortonormalizzazione

Premettiamo il seguente lemma.

Lemma 7.1. *Se $\{u_1, \ldots, u_k\}$ sono vettori di uno spazio vettoriale euclideo tra loro ortogonali e non nulli, allora essi sono indipendenti.*

DIMOSTRAZIONE. Sia $\lambda_1 u_1 + \cdots + \lambda_k u_k = 0$. Moltiplicando scalarmente per u_i, e tenendo conto che $g(u_i, u_j) = 0$ se $i \neq j$ si ha $\lambda_i = 0$ per ogni $i = 1, \ldots, k$. QED

Sia V uno spazio vettoriale euclideo con prodotto scalare g. Si dice che una base $\mathcal{B} = \{u_1, \ldots, u_n\}$ di V è *ortonormale* se

$$g(u_i, u_j) = \begin{cases} 0 & \text{se } i \neq j, \\ 1 & \text{se } i = j. \end{cases}$$

Geometricamente, ciò significa che i vettori sono perpendicolari tra loro e di lunghezza unitaria.

Teorema 7.4. *Ogni spazio vettoriale euclideo ammette una base ortonormale.*

Dimostrazione. La dimostrazione è costruttiva. Si fa vedere come da ogni base $\{u_i\}_{1\leq i\leq n}$ si può costruire una base ortonormale $\{e_i\}_{1\leq i\leq n}$ con un procedimento detto *ortonormalizzazione di Gram–Schmidt.*

Sia, dunque, $\{u_i\}_{1\leq i\leq n}$ una base in uno spazio vettoriale euclideo con prodotto scalare g. Al primo passo si pone:

$$e_1 \stackrel{\text{def}}{=} \frac{u_1}{\|u_1\|_g}.$$

Al secondo passo, si vuole costruire un vettore e_2' con queste proprietà:

1. e_2' è combinazione lineare di e_1 e u_2;
2. e_2' è ortogonale a e_1.

Si può porre $e_2' = u_2 + \lambda_1 e_1$ (il coefficiente di u_2 si può porre uguale ad 1 per il momento). Allora la condizione $g(e_2', e_1) = 0$ dà $\lambda_1 = -g(u_2, e_1)$. Tuttavia e_2' non ha ancora lunghezza 1, quindi si pone $e_2 \stackrel{\text{def}}{=} \frac{e_2'}{\|e_2'\|_g}$. Si noti che l'impropria lunghezza di e_2' è dovuta al fatto che si è posto il coefficiente di u_2 uguale a 1.

Più in generale si dimostra facilmente che

$$e_h' = u_h - \sum_{i=1}^{h-1} g(u_h, e_i)e_i, \qquad e_h \stackrel{\text{def}}{=} \frac{e_h'}{\|e_h'\|_g}.$$

QED

Si noti che il vettore $g(u_2, e_1)e_1$ è la *proiezione ortogonale di u_2 nella direzione di e_1*. Il metodo di Gram-Schmidt consiste nel trovare la componente di u_2 ortogonale ad e_1 come la differenza tra u_2 e la sua proiezione ortogonale su e_1.

Le basi ortonormali sono molto utili anche perché, rispetto a tali basi, la matrice associata è la matrice identità ed il prodotto scalare assume la forma standard

$$g(x, y) = \sum_{i=1}^{n} x_i y_i = {}^tXY, \qquad \|x\|_g^2 = \sum_{i=1}^{n} x_i^2 = {}^tXX.$$

Proposizione 7.1. *In uno spazio vettoriale euclideo V siano $\mathcal{B} = \{e_1, \dots, e_n\}$ e $\mathcal{C} = \{f_1, \dots, f_n\}$ due basi ortonormali. Allora la matrice del cambiamento di base $P = \mathcal{M}_{\mathcal{C}}^{\mathcal{B}}(\mathrm{Id})$ è una matrice ortogonale, ossia ${}^tP = P^{-1}$.*

Dimostrazione. Sia $f_j = \sum_{i=1}^n p_{ij} e_i$. Allora

$$g(f_j, f_l) = g(\sum_{i=1}^{n} p_{ij} e_i, \sum_{k=1}^{n} p_{kl} e_k) = \sum_{i,k=1}^{n} p_{ij} p_{kl} g(e_i, e_k) = \sum_{k=1}^{n} p_{kj} p_{kl} = ({}^tPP)_{jl}.$$

Pertanto $({}^tPP)_{jl} = 0$ se $j \neq l$ mentre $({}^tPP)_{jl} = 0$ se $j = l$, dunque ${}^tPP = I$. Moltiplicando quest'identità per P^{-1} a destra si ottiene ${}^tP = P^{-1}$. QED

Si noti che le colonne (le righe) delle matrici ortogonali sono una base ortonormale di $\mathbb{K}^n$; questo è equivalente a dire che $P^tP = I$. Infatti le colonne di ${}^tP$ sono le righe di P, moltiplicate scalarmente nel prodotto tra matrici con righe di P.

7.2.c Complemento ortogonale

Sia V uno spazio vettoriale euclideo con prodotto scalare g ed $U \subset V$ un suo sottospazio. Il seguente sottoinsieme di V

$$U^\perp \stackrel{\text{def}}{=} \{x \in V \mid g(x, u) = 0 \ \forall u \in U\}$$

è detto *complemento ortogonale* di U. Si lascia al lettore di provare che $U^\perp$ è un sottospazio vettoriale di V.

Teorema 7.5. *Sia V uno spazio vettoriale euclideo con prodotto scalare g. Sia U un sottospazio di V. Allora:*

$$V = U \oplus U^\perp, \tag{7.2.1}$$

Dimostrazione. Si consideri una base ortonormale $\{e_1, \dots, e_p\}$ di U. Si completi questa base ad una base $\{e_1, \dots, e_p, v_{p+1}, \dots, v_n\}$ di V. Si applichi il metodo di Gram-Schmidt a questa base; i primi p vettori non cambiano, mentre i successivi, in generale, si. Si ottiene una base ortonormale $\mathcal{B} = \{e_1, \dots, e_n\}$ di V dove i primi p vettori sono in U. Posto $W = \mathcal{L}(\{e_{p+1}, \dots, e_n\})$, se si dimostra che $W = U^\perp$ il teorema è dimostrato. Infatti è chiaro che $V = U \oplus W$: qualunque vettore non nullo in comune tra U e W contraddirebbe l'indipendenza dei vettori della base $\mathcal{B}$.

Sia $x \in W$; allora $x = \lambda_{p+1}v_{p+1} + \cdots + \lambda_n v_n$. Allora per ogni $u \in U$, posto $u = \lambda_1 v_1 + \cdots + \lambda_p v_p$ si ha, applicando la bilinearità,

$$\begin{aligned}
g(x, u) &= g(\lambda_1 e_1 + \cdots + \lambda_p e_p, \lambda_{p+1}e_{p+1} + \cdots + \lambda_n e_n) = \\
&\lambda_1 g(e_1, \lambda_{p+1}e_{p+1} + \cdots + \lambda_n e_n) + \cdots + \lambda_p g(e_p, \lambda_{p+1}e_{p+1} + \cdots + \lambda_n e_n) = \\
&\lambda_1\lambda_{p+1} g(e_1, e_{p+1}) + \cdots + \lambda_1\lambda_n g(e_1, e_n) + \cdots \\
&\cdots + \lambda_p\lambda_{p+1} g(e_p, e_{p+1}) + \cdots + \lambda_p\lambda_n g(e_p, e_n) = 0,
\end{aligned}$$

dunque $x \in U^\perp$.

Sia $x \in U^\perp$. Si supponga che $x = \lambda_1 e_1 + \cdots + \lambda_n e_n$. Per $i = 1, \dots, p$ si ha

$$\begin{aligned}
0 = g(x, e_i) = g(\lambda_1 e_1 + \cdots + \lambda_n e_n, e_i) &= \\
&\lambda_1 g(e_1, e_i) + \cdots + \lambda_n g(e_n, e_i) = \lambda_i g(e_i, e_i) = \lambda_i.
\end{aligned}$$

Dunque $x = \lambda_{p+1}e_{p+1} + \cdots + \lambda_n e_n \in W$.

Pertanto, $W = U^\perp$. QED

Si possono dimostrare le seguenti proprietà del complemento ortogonale di due sottospazi U, W di uno spazio vettoriale euclideo W:

1. $U = (U^\perp)^\perp$;
2. $(U + W)^\perp = U^\perp \cap W^\perp$;
3. $(U \cap W)^\perp = U^\perp + W^\perp$.

Come conseguenza del teorema precedente, ogni vettore $x \in V$ si decompone in modo unico come segue

$$x = x_U + x_{U^\perp}.$$

Essendo la base ortonormale si può facilmente dimostrare che

$$x = \sum_{i=1}^{n} \lambda_i e_i \quad \Rightarrow \quad \lambda_i = g(x, e_i);$$

gli scalari λ_i sono detti *coefficienti di Fourier di* x. Segue che

$$x_U = \sum_{i=1}^{p} \lambda_i e_i, \qquad x_{U^\perp} = \sum_{i=p+1}^{n} \lambda_i e_i,$$

con $\lambda_i = g(x, e_i)$. Il vettore x_U è detto *proiezione ortogonale di* x *su* U: infatti $x_U \in U$ e $x - x_U \in U^\perp$, cioè è ortogonale a U.

Si dimostra facilmente che l'applicazione

$$p_U \colon V \to V,\ x \mapsto x_U \tag{7.2.2}$$

è lineare. Questa applicazione è detta *proiezione ortogonale su* U e si vede facilmente che $p_U^2 = p_U$. Infatti se $x \in U$ allora $p_U(x) = x$, mentre se $x \in U^\perp$ allora $p_U(x) = 0$, cioè $\operatorname{Im} p_U = U$, $\operatorname{Ker} p_U = U^\perp$.

7.2.d Applicazione aggiunta

Proposizione 7.2. *Sia $f\colon V \to V$ un'applicazione lineare dallo spazio euclideo V (con prodotto scalare g). Allora esiste una ed una sola applicazione lineare $f^*\colon V \to V$ tale che*

$$g(f(x), y) = g(x, f^*(y)) \qquad \forall x \in V,\ \forall y \in W.$$

Se $\mathcal{B}$ è una <u>base ortonormale</u> per V e $\mathcal{B}'$ è una <u>base ortonormale</u> per W, allora

$$\mathcal{M}^{\mathcal{B}}_{\mathcal{B}'}(f^*) = {}^t(\mathcal{M}^{\mathcal{B}'}_{\mathcal{B}}(f)).$$

Dimostrazione. Sia $\mathcal{B} = \{e_1, \dots, e_n\}$ una base ortonormale di V. Allora, se $\mathcal{M}^{\mathcal{B}}_{\mathcal{B}}(f) = (a_{ij})$, posta $\mathcal{M}^{\mathcal{B}}_{\mathcal{B}}(f^*) = (a^*_{ij})$ la matrice da determinare, l'equazione da risolvere è

$$g(f(e_i), e_j) = g(e_i, f^*(e_j)).$$

Si ha

$$g(f(e_i), e_j) = g(\sum_{k=1}^{n} a_{ki} e_k, e_j) = \sum_{k=1}^{n} a_{ki} g(e_k, e_j) = a_{ji},$$
$$g(e_i, f^*(e_j)) = g(e_i, \sum_{k=1}^{n} a^*_{kj} e_j) = \sum_{k=1}^{n} a^*_{kj} g(e_i, e_k) = a^*_{ij},$$

da cui l'incognita a^*_{ij} è univocamente determinata come $a^*_{ij} = a_{ji}$. QED

L'applicazione lineare definita nella proposizione precedente è detta *aggiunta* di f.

Se le basi non sono ortonormali, il legame non è così semplice. Si osservi che se f_1, $f_2 \colon V \to V$ sono endomorfismi, allora

$$(f_1 \circ f_2)^* = f_2^* \circ f_1^*.$$

7.2.e Endomorfismi simmetrici

Un endomorfismo f di uno spazio vettoriale euclideo V con prodotto scalare g si dice *simmetrico* (o *autoaggiunto*) se $f = f^*$, ossia

$$g(f(x), y) = g(x, f(y)) \qquad \forall x, y \in V.$$

Se $\mathcal{B}$ è una base ortonormale, allora la matrice A associata ad f (cioè tale che $f(X) = AX$) è simmetrica, essendo ${}^tA = A$.

Esempio. Se $U \subset V$ è un sottospazio, allora l'applicazione $p_U \colon V \to V$, $x \mapsto x_U$ è un endomorfismo simmetrico. Infatti, se x, $y \in V$, allora $x = x_U + x_{U^\perp}$ e $y = y_U + y_{U^\perp}$.

$$\begin{aligned}
p_U(x) \cdot y &= p_U(x_U + x_{U^\perp}) \cdot (y_U + y_{U^\perp}) \\
&= (p_U(x_U) + p_U(x_{U^\perp})) \cdot (y_U + y_{U^\perp}) \\
&= x_U \cdot (y_U + y_{U^\perp}) \\
&= x_U \cdot y_U, \\
x \cdot p_U(y) &= (x_U + x_{U^\perp}) \cdot y_U \\
&= x_U \cdot y_U.
\end{aligned}$$

Gli endomorfismi simmetrici sono molto importanti per il seguente teorema.

Teorema 7.6. *Sia V uno spazio vettoriale euclideo (sul campo $\mathbb{R}$) con prodotto scalare g. Se $f \colon V \to V$ è un endomorfismo simmetrico, allora*

1. *le soluzioni dell'equazione caratteristica sono tutte reali;*
2. *f ammette una base ortonormale di autovettori.*

Ne segue che

1. ogni endomorfismo simmetrico f di V è semplice e V è somma diretta degli autospazi di f a due a due ortogonali;

2. ogni matrice reale simmetrica A è diagonalizzabile ortogonalmente, ossia

$$D = {}^tBAB,$$

con D matrice diagonale e B matrice ortogonale (${}^tB = B^{-1}$).

Osservazione 7.1. *Il teorema precedente dà un metodo standard per ridurre a forma canonica una forma quadratica Q. Sia A la matrice di Q; A è simmetrica, dunque $D = {}^tPAP$ dove P è ortogonale e D diagonale. D è la matrice di Q rispetto ad una nuova base, ed è anche una forma canonica di Q.*

7.2.f Caso particolare $n = 2$: le coniche

Sia V uno spazio vettoriale euclideo di dimensione 2 con prodotto scalare g, ed $f\colon V \to V$ un endomorfismo simmetrico. Allora f ha due autovalori $\lambda_1, \lambda_2 \in \mathbb{R}$.

- Sia $\lambda_1 \neq \lambda_2$. Allora $\dim V(\lambda_i) = 1$ e $V(\lambda_1) \perp V(\lambda_2)$. La base ortonormale di autovettori, rispetto cui f si presenta in forma diagonale, è costituita da un versore di $V(\lambda_1)$ e da uno di $V(\lambda_2)$.

- Sia $\lambda_1 = \lambda_2 = \lambda$. Allora $f = \lambda \operatorname{Id}_V$ e $V(\lambda) = V$. Ogni base ortonormale è una base ortonormale di autovettori di f.

Sia $\mathcal{B}$ una base ortonormale e

$$A = \mathcal{M}^{\mathcal{B}}_{\mathcal{B}}(f) = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.$$

Allora $Q_f(x) = a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2$. Per quanto sopra detto, esiste una base ortonormale di autovettori $\mathcal{B} = \{e'_1, e'_2\}$ rispetto alla quale Q_f assume la forma canonica

$$Q_f(x) = \lambda_1 {x'}_1^2 + \lambda_2 {x'}_2^2.$$

Quanto detto è utile per la riduzione a forma canonica di una conica

$$\mathcal{C}\colon a_{11}x^2 + 2a_{12}xy + a_{22}y^2 + 2a_{13}x + 2a_{23}y + a_{33} = 0.$$

Posto $v = (x, y)$ e $Q(v) = a_{11}x^2 + 2a_{12}xy + a_{22}y^2$ si può trovare una base ortonormale di autovettori rispetto alla quale Q assume la forma canonica, quindi

$$\mathcal{C}\colon \lambda_1 {x'}^2 + \lambda_2 {y'}^2 + 2a'x' + 2b'y' + c' = 0.$$

Distinguiamo due casi.

1. Sia $\lambda_1\lambda_2 \neq 0$. Tramite la traslazione
$$\tilde{x} = x' - x_0, \qquad \tilde{y} = y' - y_0,$$
dove (x_0, y_0) è il centro di $\mathcal{C}$ (punto in cui $f_x = 0 = f_y$), si può portare l'equazione di $\mathcal{C}$ alla forma canonica
$$\mathcal{C}\colon \lambda_1\tilde{x}^2 + \lambda_2\tilde{y}^2 + d = 0.$$
$$\begin{array}{lcl} \lambda_1\lambda_2 > 0 & \Rightarrow & \mathcal{C} \text{ \underline{ellisse}}, \\ \lambda_1\lambda_2 < 0 & \Rightarrow & \mathcal{C} \text{ \underline{iperbole}}. \end{array}$$
Se $d \neq 0$ si ritrova la classificazione delle scuole superiori ponendo
$$\frac{1}{a^2} = \left|\frac{\lambda_1}{-d}\right|, \quad \frac{1}{b^2} = \left|\frac{\lambda_2}{-d}\right|, \qquad \pm\frac{\tilde{x}^2}{a^2} \pm \frac{\tilde{y}^2}{b^2} = 1.$$
Si noti che se entrambi i segni sono negativi si ha un'ellisse immaginaria. Se $d = 0$ si hanno casi degeneri.

2. Sia $\lambda_1\lambda_2 = 0$. Se $\lambda_1 \neq 0$ e $\lambda_2 = 0$, tramite un'opportuna traslazione si ottiene la forma canonica
$$\lambda_1\tilde{x}^2 + d\tilde{y} = 0,$$
da cui se $d \neq 0$ la familiare forma dell'equazione di una <u>parabola</u>, $\tilde{y} = 2p\tilde{x}^2$, ponendo $2p = \lambda_1/(-d)$. Se $d = 0$ si ottengono casi degeneri.

7.2.g Caso particolare $n = 3$: le quadriche

Sia V uno spazio vettoriale euclideo di dimensione 3 con prodotto scalare g, ed $f\colon V \to V$ un endomorfismo simmetrico. Allora f ha tre autovalori $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}$.

- Siano $\lambda_1, \lambda_2, \lambda_3$ distinti. Allora $\dim V(\lambda_i) = 1$ e una base ortonormale di autovettori, rispetto cui f si presenta in forma diagonale, è costituita dai vettori $\{e'_1, e'_2, e'_3\}$ con $e'_i \in V(\lambda_i)$ e di modulo unitario.

- Sia $\lambda_1 = \lambda_2 \neq \lambda_3$. Allora $\dim V(\lambda_1) = 2$, $\dim V(\lambda_3) = 1$, e una base ortonormale di autovettori, rispetto cui f si presenta in forma diagonale, è costituita dai vettori $\{e'_1, e'_2, e'_3\}$ con $e'_1, e'_2 \in V(\lambda_1)$ ortonormali e $e'_3 \in V(\lambda_3)$.

- Siano $\lambda_1 = \lambda_2 = \lambda_3$. Allora $f = \lambda\,\mathrm{Id}_V$ e $V(\lambda_1) = V$. Ogni base ortonormale è una base ortonormale di autovettori di f.

Sia $\mathcal{B}$ una base ortonormale e $A = \mathcal{M}^{\mathcal{B}}_{\mathcal{B}}(f)$. Allora
$$Q_f(x) = a_{11}x^2 + a_{22}y^2 + a_{33}z^2 + 2a_{12}xy + 2a_{13}xz + 2a_{23}yz.$$

Per quanto sopra detto, esiste una base ortonormale di autovettori $\mathcal{B}'$ rispetto alla quale Q_f assume la forma canonica $Q_f(x') = \lambda_1 x'^2 + \lambda_2 y'^2 + \lambda_3 z'^2$.

Quanto detto è utile per la riduzione a forma canonica di una quadrica, che si può scrivere nella forma

$$\Sigma\colon Q_f(x') + 2a'_{14}x' + 2a'_{24}y' + 2a'_{34}z' + a'_{44} = 0.$$

Se $\lambda_1\lambda_2\lambda_3 \neq 0$, tramite una traslazione nel centro della quadrica (punto in cui $f_x = f_y = f_z = 0$) si ottiene la forma canonica

$$\lambda_1\tilde{x}^2 + \lambda_2\tilde{y}^2 + \lambda_3\tilde{z}^2 + d = 0. \tag{7.2.3}$$

Esaminando i diversi casi, se $d \neq 0$ si perviene ad una delle seguenti forme:

$$\frac{\tilde{x}^2}{a^2} + \frac{\tilde{y}^2}{b^2} + \frac{\tilde{z}^2}{c^2} = 1 \qquad \underline{\text{ellissoide reale}}$$
$$\frac{\tilde{x}^2}{a^2} + \frac{\tilde{y}^2}{b^2} + \frac{\tilde{z}^2}{c^2} = -1 \qquad \underline{\text{ellissoide immaginario}}$$
$$\frac{\tilde{x}^2}{a^2} + \frac{\tilde{y}^2}{b^2} - \frac{\tilde{z}^2}{c^2} = 1 \qquad \underline{\text{iperboloide ad una falda}} \text{ (o iperbolico)}$$
$$\frac{\tilde{x}^2}{a^2} - \frac{\tilde{y}^2}{b^2} - \frac{\tilde{z}^2}{c^2} = 1 \qquad \underline{\text{iperboloide a due falde}} \text{ (o ellittico)}$$

Se nell'equazione dell'ellissoide o dell'iperboloide ad una falda si ha $a = b$, il solido risulta di rotazione intorno all'asse z. Infatti l'intersezione con $z = \text{cost}$ è una circonferenza con centro sull'asse z.

Se uno degli autovalori è nullo (escludendo casi degeneri) si perviene alle forme

$$\frac{\tilde{x}^2}{a^2} + \frac{\tilde{y}^2}{b^2} = 2z \qquad \underline{\text{paraboloide ellittico}}$$
$$\frac{\tilde{x}^2}{a^2} - \frac{\tilde{y}^2}{b^2} = 2z \qquad \underline{\text{paraboloide iperbolico}}$$

Se nell'equazione (7.2.3) poniamo $d = 0$ si ottiene l'equazione canonica di <u>coni quadrici</u> (di vertice l'origine, che risulta un punto doppio). Più precisamente

$$\frac{\tilde{x}^2}{a^2} + \frac{\tilde{y}^2}{b^2} + \frac{\tilde{z}^2}{c^2} = 0 \qquad \underline{\text{cono con generatrici immaginarie}}$$
$$\frac{\tilde{x}^2}{a^2} + \frac{\tilde{y}^2}{b^2} - \frac{\tilde{z}^2}{c^2} = 0 \qquad \underline{\text{cono con generatrici reali}}$$

Se $a = b$ si ha un cono di rotazione intorno all'asse z.

I cilindri quadrici sono le quadriche rappresentate con i seguenti tipi di equazioni canoniche:

$\dfrac{\tilde{x}^2}{a^2} + \dfrac{\tilde{y}^2}{b^2} = 1$	cilindro ellittico
$\dfrac{\tilde{x}^2}{a^2} - \dfrac{\tilde{y}^2}{b^2} = 1$	cilindro iperbolico
$\tilde{x}^2 = 2p\tilde{y}$	cilindro parabolico

Altri casi con degenerazioni maggiori sono

$\dfrac{\tilde{x}^2}{a^2} + \dfrac{\tilde{y}^2}{b^2} = 0$	due piani immaginari coniugati
$\dfrac{\tilde{x}^2}{a^2} - \dfrac{\tilde{y}^2}{b^2} = 0$	due piani reali e distinti
$\tilde{x}^2 = 0$	piano contato due volte

Esercizio. Trovare l'equazione canonica della superficie

$$\Sigma\colon x^2 + 2xy + 2y^2 + 2xz + 2yz + z^2 + 4x + 3 = 0.$$

Si verifichi che si tratta di un paraboloide ellittico usando la *regola di Cartesio*[2] applicata al polinomio caratteristico.

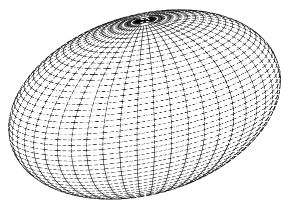

Figura – 7.1: Ellissoide

[2]Se tutti gli zeri del polinomio sono reali, allora il numero degli zeri strettamente positivi coincide con il numero delle variazioni di segno della successione ordinata dei coefficienti.

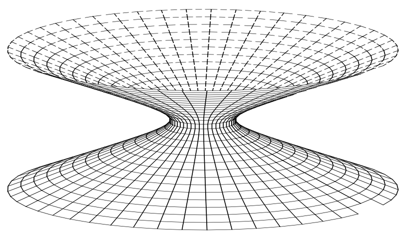

Figura – 7.2: Iperboloide ad una falda

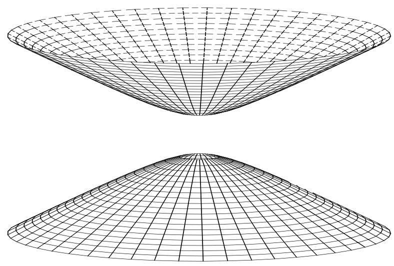

Figura – 7.3: Iperboloide a due falde

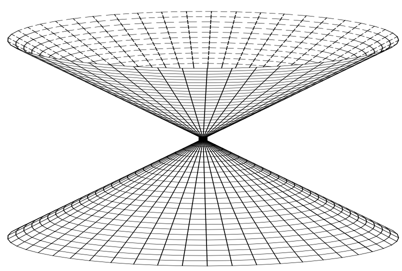

Figura – 7.4: Cono

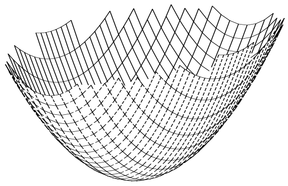

Figura – 7.5: Paraboloide ellittico

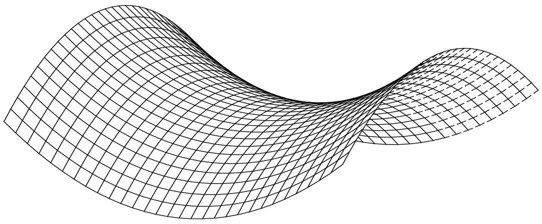

Figura – 7.6: Paraboloide iperbolico

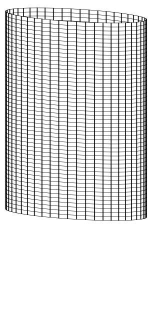

Figura – 7.7: Cilindro ellittico

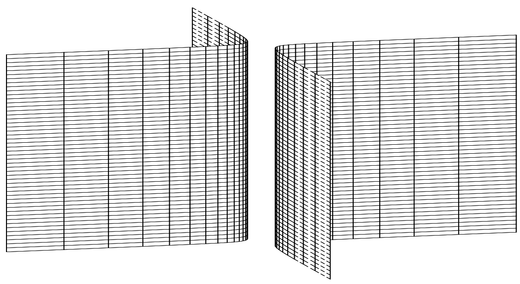

Figura – 7.8: Cilindro iperbolico

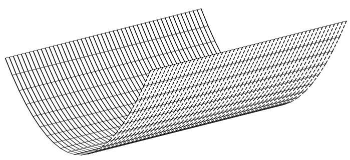

Figura – 7.9: Cilindro parabolico

7.3 Trasformazioni ortogonali

In questa sezione verranno studiate le funzioni tra spazi vettoriali euclidei che conservano i prodotti scalari. Tali funzioni conservano anche le distanze, le lunghezze e gli angoli.

7.3.a Definizione

Sia (V, g) uno spazio vettoriale euclideo. Un endomorfismo $f\colon V \to V$ si dice *trasformazione ortogonale* se

$$g(f(x), f(y)) = g(x, y) \qquad \forall x, y \in V,$$

cioè se f è un'applicazione che conserva il prodotto scalare. In particolare, f trasforma basi ortonormali in basi ortonormali.

Teorema 7.7. *Sia f un endomorfismo di uno spazio vettoriale euclideo V con prodotto scalare g. Le seguenti condizione sono equivalenti:*

1. $g(f(x), f(y)) = g(x, y)$ *per ogni* $x, y \in V$*;*
2. $\|f(x)\|_g = \|x\|_g$ *per ogni* $x \in V$*;*
3. $f^* \circ f = \mathrm{Id}_V = f \circ f^*$*, ossia* $f^* = f^{-1}$*;*
4. *per una base ortonormale* $\mathcal{B} = \{e_1, \dots, e_n\}$ *si ha che* $f(\mathcal{B}) = \{f(e_1), \dots, f(e_n)\}$ *è una base ortonormale di* V*.*

Dimostrazione. (1) $\Rightarrow$ (2). Si ponga $x = y$.
(2) $\Rightarrow$ (1). Segue tenendo conto del fatto che

$$2g(x, y) = \|x + y\|_g^2 - \|x\|_g^2 - \|y\|_g^2$$

e analogamente per $f(x)$ e $f(y)$.

(1) $\Rightarrow$ (3). Infatti $g(f(x), f(y)) = g(x, f^*(f(y)))$ per ogni $x, y \in V$. Ma $g(f(x), f(y)) = g(x, y)$, quindi

$$g(x, (f^*(f(y)) - y)) = 0 \quad \forall x, y \in V \quad \Rightarrow \quad f^*(f(y)) = y \;\; \forall y.$$

Per l'altra uguaglianza, da $g(f(z), f(w)) = g(z, w)$ e $g(f(z), f(w)) = g(z, f^*(f(w)))$ ponendo $z = f^*(x)$ e $y = f(w)$ si ha la tesi, ovvero $f(f^*(y)) = y$.

(3) $\Rightarrow$ (1). Per definizione $g(f(x), z) = g(x, f^*(z))$ per ogni $x, z \in V$. Posto $z = f(y)$ e ricordando che $f^* = f^{-1}$ si ha la tesi.

(2) $\Leftrightarrow$ (3). Segue tenendo conto le equivalenze sopra dimostrate.

(4) equivale a dire che $g(f(e_i), f(e_j)) = g(e_i, e_j)$. È facile rendersi conto del fatto che se l'uguaglianza è verificata su una base ortonormale allora è verificata su tutto V e viceversa. $\boxed{QED}$

Dal punto (2) segue che

$$\operatorname{dist}(f(x), f(y)) = \|f(x) - f(y)\|_g = \|x - y\|_g = \operatorname{dist}(x, y),$$

cioè f è una *isometria* (lineare), cioè f conserva le distanze.

Dal punto (3) segue che, indicata con $A = \mathcal{M}^{\mathcal{B}}_{\mathcal{B}}(f)$ la matrice di f rispetto ad una base ortonormale $\mathcal{B}$, si ha

$$^tAA = \mathrm{Id} = A{}^tA \quad \Rightarrow \quad {}^tA = A^{-1},$$

cioè A è una matrice *ortogonale*.

Dal punto (3) segue anche che f è un isomorfismo.

Proposizione 7.3. *Sia V uno spazio vettoriale euclideo con prodotto scalare g, e sia $f\colon V \to V$ una trasformazione ortogonale. Gli eventuali autovalori (<u>reali</u>) di f sono ± 1.*

Dimostrazione.

$$f(x) = \lambda x \quad \Rightarrow \quad \|f(x)\|_g = \|\lambda x\|_g = |\lambda|\|x\|_g.$$

Ma $\|f(x)\|_g = \|x\|_g$ da cui la tesi, essendo $\|x\|_g \neq 0$. $\boxed{QED}$

Osservazione. Una matrice reale simmetrica A si può diagonalizzare mediante una matrice ortogonale B, cioè data A esiste B ortogonale tale che

$$D = B^{-1}AB = {}^tBAB$$

sia diagonale. In tal caso A è simile a D ma anche congruente a D. Si noti che una matrice ortogonale in generale non è diagonalizzabile, come mostra l'esempio $A = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$. Infatti il polinomio caratteristico di A non ha zeri reali.

Esercizio. Sia U un piano di $\mathbb{R}^3$. La proiezione $p_U\colon \mathbb{R}^3 \to \mathbb{R}^3$ non è una trasformazione ortogonale poiché non è una isometria.

7.3.b Gruppo ortogonale

Il sottoinsieme

$$\mathbb{O}(n; \mathbb{R}) \stackrel{\text{def}}{=} \{A \in \mathbb{R}^{n,n} \mid {}^tA = A^{-1}\} \subset \mathbb{R}^{n,n}$$

è un gruppo rispetto alla moltiplicazione tra matrici (verificarlo!). Si ha

$$\mathbb{O}(n; \mathbb{R}) = \mathbb{O}^+(n; \mathbb{R}) \cup \mathbb{O}^-(n; \mathbb{R}),$$

dove

$$\mathbb{O}^+(n;\mathbb{R}) \overset{\text{def}}{=} \{A \in \mathbb{O}(n;\mathbb{R}) \mid \det A = 1\},$$
$$\mathbb{O}^-(n;\mathbb{R}) \overset{\text{def}}{=} \{A \in \mathbb{O}(n;\mathbb{R}) \mid \det A = -1\}.$$

Si vede facilmente che $\mathbb{O}^+(n;\mathbb{R})$ è un sottogruppo di $\mathbb{O}(n;\mathbb{R})$, mentre non lo è $\mathbb{O}^-(n;\mathbb{R})$. Le matrici di $\mathbb{O}^+(n;\mathbb{R})$ rappresentano cambiamenti di basi ortonormali equiverse, mentre quelle di $\mathbb{O}^-(n;\mathbb{R})$ contraverse. Le matrici di $\mathbb{O}^+(n;\mathbb{R})$ sono dette matrici di *rotazioni*, quelle di $\mathbb{O}^-(n;\mathbb{R})$ sono dette matrici di *ribaltamenti*.

Vedremo subito la motivazione di questi nomi.

Caso particolare $n = 2$. Se $A \in \mathbb{O}^+(2;\mathbb{R})$, allora essa ha la forma

$$A = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} = R(\varphi).$$

Si verifichi che $R(\varphi) \cdot R(\psi) = R(\varphi + \psi)$.

La trasformazione ortogonale $f_A \colon \mathbb{R}^2 \to \mathbb{R}^2$ tale che $f_A(X) = AX$ è una <u>rotazione</u> di angolo φ. Se $\varphi \neq k\pi$, $k \in \mathbb{Z}$, l'unico vettore fisso è il vettore nullo e non esistono autovettori per f. (Le radici dell'equazione caratteristica non sono reali.) Per $\varphi = k\pi$ si ha $R(\varphi) = \pm\,\mathrm{Id}$.

Se $A \in \mathbb{O}^-(2;\mathbb{R})$, allora essa ha la forma

$$A = \begin{pmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{pmatrix} = S(\varphi).$$

Si verifichi che $S(\varphi) \cdot S(\psi) = R(\varphi - \psi)$.

La trasformazione ortogonale $f_A \colon \mathbb{R}^2 \to \mathbb{R}^2$ tale che $f_A(X) = AX$ ha autovalori ± 1 e autospazi

$$V(1) = \{(x, y) \in \mathbb{R}^2 \mid y = x \operatorname{tg} \frac{\varphi}{2}\}$$
$$V(-1) = \{(x, y) \in \mathbb{R}^2 \mid y = x \operatorname{tg} \left(\frac{\varphi}{2} + \frac{\pi}{2}\right)\}$$

Le rette $V(1)$ e $V(-1)$ sono tra loro ortogonali, e $V(1)$ è il luogo dei punti fissi. f_A si dice *ribaltamento* o *simmetria assiale* (rispetto alla retta $V(1)$). L'endomorfismo f_A è semplice ed $A = S(\varphi)$ è simile alla matrice $S(0)$, che esprime il ribaltamento intorno all'asse x, mentre, ovviamente, $S(\varphi)$ esprime il ribaltamento rispetto alla retta $y = x \operatorname{tg}(\varphi/2)$.

Esercizi.

- Si verifichi che $R(\varphi) \cdot S(0) = S(\varphi)$, e si interpreti geometricamente il risultato.
- Calcolare $S(0) \cdot R(\varphi)$.

- Dire se $\mathbb{O}(2;\mathbb{R})$ e $\mathbb{O}^+(2;\mathbb{R})$ sono gruppi commutativi.
- È vero che $R(-\varphi) = -R(\varphi)$?
- Provare che le simmetrie assiali generano tutti gli elementi di $\mathbb{O}(2;\mathbb{R})$.

Caso particolare $n = 3$. Sia $f\colon \mathbb{R}^3 \to \mathbb{R}^3$ una trasformazione ortogonale e consideriamo il sottospazio dei vettori fissi

$$U = \{x \in \mathbb{R}^3 \mid f(x) = x\}.$$

Si può dimostrare che

$$\begin{array}{lll}
\dim U = 3 & \Rightarrow & f \text{ identità} \\
\dim U = 2 & \Rightarrow & f \text{ simm. ort. risp. } U \\
\dim U = 1 & \Rightarrow & f \text{ rotaz. intorno } U \\
\dim U = 0 & \Rightarrow & f \text{ rotosimmetria}
\end{array}$$

dove per rotosimmetria si intende la composizione di una rotazione intorno ad una retta e di una simmetria rispetto ad un piano. Se $\mathcal{B}$ è una base ortonormale contenente una base di $V(1)$ e $A = \mathcal{M}^{\mathcal{B}}_{\mathcal{B}}(f)$, allora, posto $U = V(1)$, nel caso $U \neq \{0\}$,

$$\begin{array}{lll}
\dim U = 3 & \Rightarrow & A = \mathrm{Id} \\
\dim U = 2 & \Rightarrow & A \sim \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
\dim U = 1 & \Rightarrow & A \sim \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi & -\sin\varphi \\ 0 & \sin\varphi & \cos\varphi \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & R(\varphi) \end{pmatrix} = \tilde{R}(\varphi) \\
\dim U = 0 & \Rightarrow & A \sim \begin{pmatrix} -1 & 0 \\ 0 & R(\varphi) \end{pmatrix} = \tilde{S}(\varphi)
\end{array}$$

Osservazione. Una trasformazione ortogonale con determinante positivo, cioè una rotazione, o è l'identità o è una rotazione propria intorno ad una retta (teorema di Eulero).

Nota. Sia V uno spazio vettoriale euclideo con prodotto scalare g. Il sottoinsieme

$$\mathbb{O}(V, g) \stackrel{\text{def}}{=} \{f\colon V \to V \mid f \text{ ortogonale}\} \subset \mathrm{End}(V) = \mathrm{Lin}(V, V)$$

è un gruppo rispetto alla composizione di applicazioni.

Scelta una base $\mathcal{B}$ ortonormale di V, è immediato verificare che l'applicazione

$$\mathcal{M}\colon \mathbb{O}(V, g) \to \mathbb{O}(n, \mathbb{R}),\ \ f \mapsto \mathcal{M}^{\mathcal{B}}_{\mathcal{B}}(f)$$

è un isomorfismo di gruppi.

7.3.c Movimenti

Sia V uno spazio euclideo di dimensione n con prodotto scalare g. Si dice *movimento* di V un'applicazione m ottenuta componendo una trasformazione ortogonale f di V con una traslazione $t_{\vec{b}}$, ossia

$$m = t_{\vec{b}} \circ f \colon V \to V, \; x \mapsto f(x) + \vec{b},$$

oppure $m = f \circ t_{\vec{c}}$ con $\vec{c} = f^{-1}(\vec{b})$.

Si osservi che, se $\vec{b} \neq 0$, allora m non è una trasformazione lineare.

Se $f \in \mathbb{O}^+(V, g)$, allora m si dice *movimento diretto*.

Si prova immediatamente che m è un'*isometria*, infatti

$$\mathrm{dist}(m(x), m(y)) = \|m(x) - m(y)\|_g == \|f(x-y)\|_g = \|x - y\|_g = \mathrm{dist}(x, y).$$

Questa proprietà è caratteristica dei movimenti (come mostra il seguente teorema, che non sarà dimostrato), per cui i movimenti sono chiamati anche *trasformazioni rigide*.

Teorema 7.8. *Sia V uno spazio euclideo di dimensione n con prodotto scalare g. Se m è un'isometria di V, allora m è un movimento.*

Vogliamo ora trovare la rappresentazione di un movimento in coordinate. Sia $\mathcal{B} = \{e_1, \ldots, e_n\}$ una base ortonormale di V e $\mathcal{M}^{\mathcal{B}}_{\mathcal{B}}(f) = A \in \mathbb{O}(n; \mathbb{R})$. Posto $x = (x_1, \ldots, x_n)$, $m(x) = (x'_1, \ldots, x'_n)$, $\vec{b} = (b_1, \ldots, b_n)$, si ha

$$x'_i = \sum_j a_{ij} x_j + b_i \quad \text{oppure} \quad X' = AX + B.$$

Dunque m è una *trasformazione affine*. Inoltre

$$\begin{aligned} A \in \mathbb{O}^+(n, \mathbb{R}) \quad &\Rightarrow \quad m \text{ rototraslazione}, \\ A \in \mathbb{O}^-(n, \mathbb{R}) \quad &\Rightarrow \quad m \text{ glissosimmetria}, \end{aligned}$$

dove per glissosimmetria si intende la composizione di una traslazione con una simmetria.

Osservazione. Per ogni n, le simmetrie (o ribaltamenti) rispetto ad iperpiani (sottospazi vettoriali di dimensione $n-1$) generano tutto il gruppo delle trasformazioni ortogonali di $\mathbb{R}^n$.

CAPITOLO 8

GEOMETRIA ANALITICA DELLO SPAZIO

La Geometria Analitica, nata con Cartesio (intorno al 1637), ha lo scopo di tradurre un problema geometrico in uno analitico, cioè mediante l'uso delle coordinate tradurre un problema riguardante figure geometriche in un problema riguardante numeri ed equazioni. Ma è ancora più interessante l'inverso: interpretare geometricamente equazioni e loro sistemi. In tal modo la geometria e l'analisi si illuminano a vicenda ed è possibile passare dallo spazio dell'intuizione a spazi astratti.

In questo capitolo tratteremo la geometria analitica dello spazio ordinario di dimensione 3.

Quella del piano si ritiene nota, tuttavia essa si può ottenere facilmente da quella dello spazio considerando il piano xy, immerso nello spazio, come il luogo dei punti $P(x, y, z)$ aventi la terza coordinata nulla.

8.1 Piano

Tre punti P_1, P_2, P_3 non allineati (quindi $\vec{P_1P_2}$ e $\vec{P_1P_3}$ indipendenti) individuano un piano α

$$P \in \alpha \quad \Leftrightarrow \quad \vec{P_1P}, \vec{P_1P_2}, \vec{P_1P_3} \quad \text{dipendenti.}$$

Posto $P_i(x_i, y_i, z_i)$, $P(x, y, z)$, la dipendenza lineare si può esprimere in due modi.

8.1.a Piano: equazione cartesiana

$$\begin{vmatrix} x - x_1 & y - y_1 & z - z_1 \\ x_2 - x_1 & y_2 - y_1 & z_2 - z_1 \\ x_3 - x_1 & y_3 - y_1 & z_3 - z_1 \end{vmatrix} = 0 \quad \Leftrightarrow \quad \begin{vmatrix} x & y & z & 1 \\ x_1 & y_1 & z_1 & 1 \\ x_2 & y_2 & z_2 & 1 \\ x_3 & y_3 & z_3 & 1 \end{vmatrix} = 0. \qquad (8.1.1)$$

Sviluppando il determinante, si ha un'*equazione cartesiana del piano*

$$ax + by + cz + d = 0, \qquad (a, b, c) \neq (0, 0, 0), \qquad (8.1.2)$$

che risulta individuata a meno di un fattore di proporzionalità non nullo. I parametri (a, b, c) si chiamano *coefficienti di giacitura* del piano e rappresentano le coordinate di un vettore (non nullo) perpendicolare al piano. Infatti, considerando il vettore $\vec{n} = (a, b, c)$ uscente da $P_0 \in \alpha$, si ha

$$\vec{n} \cdot \overrightarrow{P_0P} = 0 \qquad \forall P \in \alpha \tag{8.1.3}$$

da cui

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0,$$

che rappresenta il piano per P_0 con coefficienti di giacitura (a, b, c).

Se $c = 0$, allora si ha il piano passante per P_0 e parallelo all'asse z. Si osservi che la coordinata z può essere arbitraria.

8.1.b Piano: equazioni parametriche

$$\overrightarrow{P_1P} = u\overrightarrow{P_1P_2} + v\overrightarrow{P_1P_3}, \qquad u, v \in \mathbb{R}, \tag{8.1.4}$$

da cui

$$\begin{cases} x = x_1 + u(x_2 - x_1) + v(x_3 - x_1) \\ y = y_1 + u(y_2 - y_1) + v(y_3 - y_1) \\ z = z_1 + u(z_2 - z_1) + v(z_3 - z_1) \end{cases}$$

che sono dette *equazioni parametriche del piano*. Eliminando i parametri u e v si perviene all'equazione cartesiana.

Esempio. Dati i punti $P_1(1, 0, 0)$, $P_2(1, 1, 1)$, $P_3(1, 0, 1)$ troviamo le equazioni parametriche e cartesiana del piano. Si ha $\overrightarrow{P_1P_2} = (0, 1, 1)$, $\overrightarrow{P_1P_3} = (0, 0, 1)$, dunque

$$\begin{cases} x = 1 \\ y = u \\ z = u + v \end{cases},$$

da cui l'equazione cartesiana $x = 1$.

8.1.c Mutue posizioni di piani

Siano α ed α' due piani. Volendo studiare la loro mutua posizione, consideriamo il sistema lineare

$$\begin{cases} ax + by + cz + d = 0 \\ a'x + b'y + c'z + d' = 0 \end{cases}$$

Risulta

$$\begin{aligned} \text{sistema incompatibile} \quad &\Leftrightarrow \quad \mathrm{rg}(A) \neq \mathrm{rg}(\tilde{A}) \quad &\Leftrightarrow \quad \alpha \cap \alpha' = \varnothing, \\ \text{sistema compatibile} \quad &\Leftrightarrow \quad \mathrm{rg}(A) = \mathrm{rg}(\tilde{A}) \quad &\Leftrightarrow \quad \alpha \cap \alpha' \neq \varnothing. \end{aligned}$$

Inoltre

$$\begin{array}{lcll} \mathrm{rg}(A) = \mathrm{rg}(\tilde{A}) = 2 & \Leftrightarrow & \infty^1 \text{ soluzioni} & \Leftrightarrow \quad \alpha \cap \alpha' = r, \\ \mathrm{rg}(A) = \mathrm{rg}(\tilde{A}) = 1 & \Leftrightarrow & \infty^2 \text{ soluzioni} & \Leftrightarrow \quad \alpha \equiv \alpha', \end{array}$$

dove r è una retta. Ponendo

$$\alpha \parallel \alpha' \quad \Leftrightarrow \quad \alpha \cap \alpha' = \varnothing \quad \text{oppure} \quad \alpha \equiv \alpha'$$

possiamo dire che

$$\alpha \parallel \alpha' \quad \Leftrightarrow \quad (a, b, c) \sim (a', b', c'),$$

dove '$\sim$' sta per 'è proporzionale a'.

Esempi ed esercizi.

- I piani $x - y + 2z = 1$ e $3x - 3y + 6z = 1$ sono paralleli; i piani $x - y + 2z = 1$ e $3x - 3y + 6z = 3$ sono paralleli e coincidenti.
- Il piano perpendicolare al vettore $(1, -1, 2)$ e uscente dal punto $(3, -1, 5)$ è

$$1(x - 3) + (-1)(y + 1) + 2(z - 5) = 0.$$

8.2 Retta

Due punti distinti P_1 e P_2 (quindi $\vec{P_1P_2} \neq 0$)individuano una retta r

$$P \in r \quad \Leftrightarrow \quad \vec{P_1P}, \vec{P_1P_2} \text{ dipendenti.}$$

La dipendenza lineare si può esprimere nei seguenti modi.

8.2.a Retta: equazioni cartesiane

$$\mathrm{rg}\begin{pmatrix} x - x_1 & y - y_1 & z - z_1 \\ x_2 - x_1 & y_2 - y_1 & z_2 - z_1 \end{pmatrix} = 1 \quad \Leftrightarrow \quad \left\| \begin{matrix} x - x_1 & y - y_1 & z - z_1 \\ x_2 - x_1 & y_2 - y_1 & z_2 - z_1 \end{matrix} \right\| = 0,$$

cioè sono nulli i determinanti di tutti i minori di ordine 2 estratti dalla matrice. Ciò, <u>solo se ha senso la scrittura seguente</u>, equivale a

$$\frac{x - x_1}{x_2 - x_1} = \frac{y - y_1}{y_2 - y_1} = \frac{z - z_1}{z_2 - z_1},$$

(*attenzione*: al denominatore non può comparire 0) che si può porre nella forma seguente (confronta 8.1.c, caso $\alpha \cap \alpha' \neq \varnothing$).

$$r\colon \begin{cases} ax + by + cz + d = 0 \\ a'x + b'y + c'z + d' = 0. \end{cases}$$

Queste ultime sono dette *equazioni cartesiane della retta.* Quindi, ogni retta r si può scrivere come intersezione di due piani α ed α', di equazioni cartesiane

$$\alpha\colon ax + by + cz + d = 0, \qquad \alpha'\colon a'x + b'y + c'z + d' = 0,$$

e tali che

$$\operatorname{rg}\begin{pmatrix} a & b & c \\ a' & b' & c' \end{pmatrix} = 2.$$

Si osservi che la retta r non determina univocamente i piani α ed α': due altri piani *distinti* passanti per r (ce ne sono ∞^1) individuano la stessa retta.

Si chiamano *parametri direttori* di r le coordinate di un arbitrario vettore v parallelo ad r. Ovviamente, se $P_1, P_2 \in r$ e $P_1 \neq P_2$, allora $v = \vec{P_1P_2}$ è parallelo ad r e quindi parametri direttori di r sono

$$l = x_2 - x_1, \quad m = y_2 - y_1, \quad n = z_2 - z_1.$$

I parametri direttori $v = (l, m, n)$ di una retta sono individuati a meno di un fattore di proporzionalità. Poiché il vettore (l, m, n) deve essere perpendicolare ad $\vec{n}$ ed $\vec{n'}$, esso sarà parallelo a $\vec{n} \wedge \vec{n'}$, quindi

$$(l, m, n) \sim \left(\begin{vmatrix} b & c \\ b' & c' \end{vmatrix}, - \begin{vmatrix} a & c \\ a' & c' \end{vmatrix}, \begin{vmatrix} a & b \\ a' & b' \end{vmatrix} \right).$$

Se v ha modulo unitario, le sue coordinate sono dette *coseni direttori* di r. Si osservi che

$$\cos^2 \widehat{rx} + \cos^2 \widehat{ry} + \cos^2 \widehat{rz} = 1.$$

Si noti che i parametri direttori di r possono anche essere ricavati come soluzioni del sistema omogeneo associato ad r: è facile rendersi conto del fatto che il vettore $P_2 - P_1$ soddisfa le equazioni di tale sistema.

8.2.b Retta: equazioni parametriche

Si ha

$$\vec{P_1P} = t\vec{P_1P_2}, \qquad t \in \mathbb{R},$$

da cui

$$\begin{cases} x = x_1 + t(x_2 - x_1) = x_1 + lt \\ y = y_1 + t(y_2 - y_1) = y_1 + mt \\ z = z_1 + t(z_2 - z_1) = z_1 + nt \end{cases}$$

che sono dette *equazioni parametriche della retta.* Eliminando il parametro t si perviene alle equazioni cartesiane.

8.2.c Retta nel piano

Naturalmente, se P, P_1, P_2 appartengono al piano $z = 0$, si avrà

$$\begin{vmatrix} x - x_1 & y - y_1 \\ x_2 - x_1 & y_2 - y_1 \end{vmatrix} = 0 \quad \Leftrightarrow \quad \begin{vmatrix} x & y & 1 \\ x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \end{vmatrix} = 0,$$

da cui l'equazione cartesiana di una retta nel piano $z = 0$

$$r\colon ax + by + c = 0, \qquad \text{con } (a, b) \neq (0, 0).$$

Il vettore $\vec{n} = (a, b)$ del piano xy è perpendicolare ad r. Analogamente equazioni parametriche di r sono

$$\begin{cases} x = x_1 + t(x_2 - x_1) \\ y = y_1 + t(y_2 - y_1) \end{cases} \qquad t \in \mathbb{R}.$$

8.2.d Mutua posizione tra rette e piani

Si ritiene noto (se non lo è, acquisirlo) fin dalle scuole secondarie, il concetto di angolo tra due semirette e tra due rette, e quindi il concetto di angolo tra due piani e tra una retta ed un piano.

Ad ogni piano α associamo il vettore $\vec{n} = (a, b, c)$, perpendicolare ad α, di coordinate i parametri di giacitura; ad ogni retta r associamo il vettore $\vec{r} = (l, m, n)$, parallelo ad r, di coordinate i parametri direttori. Così si esprimono facilmente le condizioni di parallelismo e perpendicolarità tra rette e piani.

$$\begin{array}{rcl} \alpha \parallel \alpha' & \Rightarrow & \vec{n} \parallel \vec{n'} \qquad \Rightarrow \quad (a, b, c) \sim (a', b', c') \\ \alpha \perp \alpha' & \Rightarrow & \vec{n} \perp \vec{n'} \qquad \Rightarrow \quad aa' + bb' + cc' = 0 \\ \alpha \perp r & \Rightarrow & \vec{n} \parallel \vec{r} \qquad \Rightarrow \quad (a, b, c) \sim (l, m, n) \\ \alpha \parallel r & \Rightarrow & \vec{n} \perp \vec{r} \qquad \Rightarrow \quad al + bm + cn = 0 \\ r \parallel r' & \Rightarrow & \vec{r} \parallel \vec{r'} \qquad \Rightarrow \quad (l, m, n) \sim (l', m', n') \\ r \perp r' & \Rightarrow & \vec{r} \perp \vec{r'} \qquad \Rightarrow \quad ll' + mm' + nn' = 0 \end{array}$$

Le precedenti condizioni si possono anche ottenere studiando il sistema costituito dalle equazioni delle rette e dei piani.

Siano ora r ed r' due rette *orientate* e $\vec{r}$, $\vec{r'}$ due vettori concordemente orientati con r ed r'. Allora, sfruttando la definizione di prodotto scalare, si ha

$$\cos\widehat{rr'} = \cos\widehat{\vec{r}\vec{r'}} = \frac{\vec{r}\cdot\vec{r'}}{\|\vec{r}\|\,\|\vec{r'}\|} = \frac{ll'+mm'+nn'}{\sqrt{l^2+m^2+n^2}\sqrt{l'^2+m'^2+n'^2}}.$$

Se, invece, le due rette non sono orientate, l'angolo $\widehat{rr'}$ non è individuato, ma può assumere due valori tra loro supplementari:

$$\cos\widehat{rr'} = \pm\frac{\vec{r}\cdot\vec{r'}}{\|\vec{r}\|\,\|\vec{r'}\|} = \pm\frac{ll'+mm'+nn'}{\sqrt{l^2+m^2+n^2}\sqrt{l'^2+m'^2+n'^2}}.$$

Così, indicate con n ed n' le rette normali rispetto ad α ed α', si ha

$$\cos\widehat{\alpha\alpha'} = \cos\widehat{\vec{n}\vec{n'}} = \frac{\vec{n}\cdot\vec{n'}}{\|\vec{n}\|\,\|\vec{n'}\|} = \pm\frac{aa'+bb'+cc'}{\sqrt{a^2+b^2+c^2}\,\sqrt{a'^2+b'^2+c'^2}}$$

$$\sin\widehat{\alpha r} = |\cos\widehat{\vec{n}\vec{r}}| = \frac{|\vec{n}\cdot\vec{r}|}{\|\vec{n}\|\,\|\vec{r}\|} = \frac{|al+bm+cn|}{\sqrt{a^2+b^2+c^2}\,\sqrt{l^2+m^2+n^2}}$$

8.2.e Rette sghembe

Due rette r ed r' sono *sghembe* se non esiste alcun piano che le contiene. In tal caso si può provare che esistono due piani α e α' tra loro paralleli tali che

$$r\subset\alpha, \qquad r'\subset\alpha'.$$

Chiaramente, $\mathrm{dist}(r,r') = \mathrm{dist}(\alpha,\alpha')$, che è la cosiddetta *minima distanza* tra le rette sghembe r e r'. Ricordiamo che, se F ed F' sono due insiemi di punti dello spazio, la *distanza* tra F ed F' è per definizione

$$\mathrm{dist}(F,F') = \inf\{\mathrm{dist}(P,P');\ P\in F, P'\in F'\}.$$

Siano $\vec{r}$ ed $\vec{r'}$ vettori rispettivamente paralleli alle rette considerate. Allora $\vec{n} = \vec{r}\wedge\vec{r'}$ indica la giacitura di α ed α'. Se β (rispettivamente β') è il piano per r (rispettivamente per r') e parallelo a $\vec{n}$, allora $\beta\cap\beta' = n$ è la retta ortogonale ad r ed r' che si appoggia a queste rette.

Posto

$$n\cap\alpha = n\cap r = Q, \qquad n\cap\alpha' = n\cap r' = Q',$$

si ha $\mathrm{dist}(r,r') = \mathrm{dist}(Q,Q')$. In definitiva, la distanza tra rette, tra rette e piani, e tra piani, è sempre ricondotta alla distanza tra punti.

8.3 Sfere e circonferenze

Chiamiamo *sfera* l'insieme dei punti P dello spazio tali che $\|\vec{CP}\| = R$, dove C è un punto fisso e R un numero reale positivo. Se $C(\alpha, \beta, \gamma)$ e $P(x, y, z)$, da $\|\vec{CP}\| = R$ si ha

$$(x-\alpha)^2 + (y-\beta)^2 + (z-\gamma)^2 = R^2,$$

che dà l'equazione cartesiana di una sfera generica

$$x^2 + y^2 + z^2 - 2\alpha x - 2\beta y - 2\gamma z + \delta = 0,$$

dove $\delta = \alpha^2 + \beta^2 + \gamma^2 - R^2$. Viceversa, ogni equazione del tipo

$$x^2 + y^2 + z^2 + 2ax + 2by + 2cz + d = 0$$

rappresenta una sfera Σ di centro (α, β, γ), dove $\alpha = -a$, $\beta = -b$, $\gamma = -c$, e di raggio $R = \sqrt{a^2 + b^2 + c^2 - d}$. Si ha:

$$\begin{array}{ll} a^2 + b^2 + c^2 - d > 0 & \text{sfera ordinaria,} \\ a^2 + b^2 + c^2 - d = 0 & \text{sfera di raggio nullo,} \\ a^2 + b^2 + c^2 - d < 0 & \text{sfera immaginaria.} \end{array}$$

Se π è un piano, $\Sigma \cap \pi$ è una circonferenza. In particolare nel piano $z = 0$ una circonferenza è rappresentata da una equazione del tipo

$$\mathcal{C}\colon x^2 + y^2 + 2ax + 2by + d = 0.$$

Esempi ed esercizi.

- Nello spazio $Oxyz$ determinare le equazioni delle sfere appartenenti all'ottante $x \geq 0$, $y \geq 0$, $z \geq 0$ e tangenti ai piani coordinati ed al piano

$$\alpha\colon x + 2y + 2z - 1 = 0.$$

- Scrivere l'equazione della sfera che ha come punti diametralmente opposti $A(3, 0, 0)$ e $B(1, 1, 1)$. Determinare e discutere nel campo complesso l'intersezione della sfera con il piano coordinato yz.
- Nello spazio $Oxyz$ determinare il centro ed il raggio della circonferenza $\mathcal{C}$ passante per i punti O, $A(1, 0, 1)$, $B(0, -1, 1)$.
- Nel piano Oxy, data la circonferenza

$$\mathcal{C}\colon x^2 + y^2 - 2x - 2y = 0,$$

si determinino le rette tangenti ad essa per il punto $P_0(-1, 1)$ e si verifichi che risultano ortogonali.

- Nel piano Oxy determinare la tangente r a $\mathcal{C}\colon x^2+y^2-2x-2y=0$ nel punto $P_0(0,2)\in\mathcal{C}$.

  Si osservi che r passerà per P_0 e sarà ortogonale al vettore $\vec{CP_0}$ dove C è il centro di $\mathcal{C}$. Quindi r sarà del tipo

  $$a(x-x_0)+b(y-y_0)=0,$$

  con $(a,b)\parallel \vec{CP_0}=(x_0-\alpha, y_0-\beta)$. Nel nostro caso $C(1,1)$ e quindi $r\colon x-y+2=0$.

- Nel piano Oxy scrivere l'equazione della circonferenza che passa per l'origine O ed è tangente nel punto $P_0(1,2)$ alla retta $r\colon x-y+1=0$.

  Naturalmente il centro C apparterrà alla retta n perpendicolare ad r in P_0, e inoltre $d(C,P_0)=d(C,O)$.

- Nel piano Oxy determinare le tangenti a

  $$\mathcal{C}\colon x^2+y^2-7x+y=0$$

  parallele all'asse x e trovare i punti di tangenza.

- Scrivere tutte le circonferenze tangenti in O all'asse y.

BIBLIOGRAFIA

[1] M. Abate: Algebra lineare. McGraw-Hill Libri Italia, Milano, 2000.

[2] E. Abbena, E. Bargero, S. Garbiero: Esercizi di Algebra Lineare e Geometria Analitica, Levrotto e Bella, Torino 1990.

[3] S. Abeasis: Algebra lineare e geometria, Zanichelli 1990.

[4] I. Cattaneo Gasparini: Strutture algebriche, operatori lineari, Roma, 1989.

[5] G. Calvaruso, R. Vitolo: Esercizi di Geometria ed Algebra, Università di Lecce, 2001.

[6] P. de Bartolomeis: Algebra Lineare, La Nuova Italia, Firenze 1993.

[7] F.Fava, F. Tricerri: Geometria ed Algebra Lineare, Levrotto e Bella, Torino 1987.

[8] G. Gambarelli: Benvenuto alle matricole (e qualche consiglio), in AA. VV., L'endecasillabo matematico, Mathesis 1999.

[9] W. Greub: Linear Algebra, IV ed., Springer, 1975.

[10] P. Maroscia: Problemi di Geometria, Masson, 1994.

[11] P. Maroscia: Geometria ed Algebra Lineare, Zanichelli, 2002.

[12] V. K. Nicholson: Algebra lineare, dalle applicazioni alla teoria, McGraw–Hill, 2002.

[13] A. Sanini: Lezioni di Geometria, ed. Levrotto-Bella, Torino 1993; Esercizi di Geometria, ed. Levrotto-Bella, Torino 1993.

[14] E. Sernesi: Geometria 1, Boringhieri, Bologna 1989.

[15] G. Vaccaro, A. Carfagna, L. Piccolella: Lezioni di geometria ed algebra lineare, Masson, 1995.

INDICE ANALITICO