Università degli Studi di Lecce

Facoltà di Ingegneria

Giuseppe De Cecco

Raffaele Vitolo

NOTE

DI

GEOMETRIA ED ALGEBRA

Versione del 20 settembre 2009

ANNO ACCADEMICO 2009-2010

Indirizzo degli autori.
Giuseppe De Cecco, Raffaele Vitolo,
Università di Lecce, Dipartimento di Matematica E. De Giorgi,
via per Arnesano, 73100 Lecce
`Giuseppe.DeCecco@unile.it`
`Raffaele.Vitolo@unile.it`

INDICE

INTRODUZIONE

I confini tra matematica pura ed applicata sono labili. Alla matematica pura si domanda la coerenza interna dei suoi enunciati, alla matematica applicata la capacità di rappresentare diverse realtà esterne alla matematica stessa. La distinzione tra matematica pura ed applicata non risiede nella diversa qualità dei teoremi che vi si dimostrano, ma nei diversi criteri di interesse che inizialmente le ispirano.

E. De Giorgi

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

A. Einstein

Queste note vogliono essere solo una guida alla preparazione dell'esame di 'Geometria ed Algebra', e non intendono sostituire i testi consigliati, ai quali si rimanda per una completa preparazione.

Le presenti note si dividono in due parti. La prima parte è dedicata alla Geometria Analitica, i cui fondamenti furono elaborati da R. Descartes nel trattato *La Géometrie*, pubblicato (nel 1637) come appendice al famoso *Discours de la méthode.* In quel piccolo trattato si usa sistematicamente l'algebra simbolica e se ne dà una interpretazione geometrica.

La seconda parte è dedicata all'Algebra Lineare, nata alla fine del 1700 per lo studio dei sistemi di n equazioni lineari in n incognite. Il primo esempio di calcolo esplicito per $n = 2$ e $n = 3$ fu dato da Maclaurin nel 1748, mentre il metodo generale risale a G. Cramer nel 1750. In tutta la teoria un ruolo determinante è svolto dai *determinanti di matrici.* L'Algebra Lineare e Multilineare è diventato uno strumento potente largamente usato nelle applicazioni, dal momento che ogni fenomeno non lineare può approssimarsi ad uno lineare.

Contrariamente alle usuali presentazioni, noi abbiamo preferito premettere all'Algebra Lineare la Geometria Analitica, poiché gli studenti conoscono già dalle Scuole superiori i primi rudimenti di Geometria analitica sui quali si innesta il corso, che fin dall'inizio fa uso del metodo vettoriale, che prepara così con esempi alla seconda parte più astratta, in cui viene data la nozione generale di *spazio vettoriale* e di *applicazione lineare.* La moderna definizione di spazio vettoriale è stata introdotta da G. Peano nel 1888.

Ogni capitolo è corredato di esempi e di esercizi svolti o proposti; altri si trovano nel testo di esercizi [4].

Il testo di esercizi, a sua volta, deve essere considerato non come un corso a sé stante, ma come parte integrante della teoria. Consigliamo vivamente di svolgere tutti gli esercizi proposti per un dato argomento solo dopo aver studiato e capito quest'ultimo. Inoltre, per una piena comprensione della teoria, *è indispensabile completare tutte le semplici dimostrazioni lasciate al lettore*. Imparare a fare gli esercizi senza conoscere la teoria ("preparare lo scritto") significa solo imparare a memoria alcune tecniche, senza possedere la versatilità necessaria a dominare tutte le situazioni.

Analogamente, non basta studiare la Matematica, occorre capirla e poi saperla utilizzare. Buona parte di ciò che si studierà in questo corso verrà applicata durante i corsi di tipo ingegneristico degli anni successivi; tuttavia, non è un atteggiamento corretto quello di pensare che alcune parti del programma non servano a niente. La Matematica è essenzialmente capacità di creare una rete di risultati logici, passando anche attraverso risultati apparentemente inutili; ma se tagliamo quelli inutili, la rete si frantuma.

Potremmo anche dire qual è la matematica applicata, cioè quella che è stata *applicata*, ma non potremmo certamente dire quale sarà *applicabile*. Anzi, se guardiamo alla storia, non sbaglieremo prevedendo che tutta, anche quella più astratta, potrà essere usata per spiegare fenomeni naturali e per progettare prodotti tecnologici sofisticati.

Samuel Ting, premio Nobel per la Fisica nel 1976, così si esprime:

> *Tutti sono d'accordo che la qualità della vita e il benessere della società nei paesi industrializzati (..) si basano su ritrovati di tecnologie. Ciò che viene dimenticato sta nel fatto che le basi di questi ritrovati furono messe qualche tempo fa dagli scienziati i quali furono spinti dalla curiosità e non dalle promesse del mercato.*

(da *Research of Today is the Technology of Tomorrow*)

Concludiamo con un avvertimento, dettato dalla nostra esperienza di insegnanti, che hanno esaminato tanti studenti.

La matematica non solo va studiata e capita, ma va anche esposta correttamente: innanzitutto le definizioni e poi le dimostrazioni dei teoremi, cercando di sottolineare quando vengono usate le ipotesi.

> *Spesso si presume di aver capito, ma manca ancora la naturalezza o addirittura la coerenza dell'esposizione. Naturalmente, tale esercizio è molto più produttivo se è effettuato con un'altra persona. Lo studio isolato è comunque pericolosissimo per i timidi, perché leggono in silenzio, poi ripetono in silenzio, poi sostengono l'esame in silenzio [7].*

Giuseppe De Cecco
Raffaele Vitolo

Ringraziamenti.

Ringraziamo l'Università di Lecce ed il Dipartimento di Matematica per aver concesso l'uso delle proprie strutture ed apparecchiature. Ringraziamo anche tutti gli studenti che hanno contribuito a questo testo correggendone molti errori e sviste. R. Vitolo ringrazia A. Blandolino per aver messo a disposizione i propri appunti presi a lezione.

Queste note sono state scritte in LaTeX2e con l'estensione `amsmath` della American Mathematical Society e l'estensione `diagrams` scritta da P. Taylor.

CAPITOLO 1

PREMESSE

1.1 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 altri corsi.

1.1.a Legge di composizione

Sia A un insieme. Indichiamo con $A \times A$ il *prodotto cartesiano* A per A:

$$A \times A = \{(a, b) | a \in A, b \in A\}$$

i cui elementi sono le *coppie ordinate* (a, b); quindi in generale $(a, b) \neq (b, a)$. Perciò

$$(a, b) = (c, d) \Rightarrow a = c, b = d.$$

Una legge di composizione in A è un'applicazione definita in $A \times A$ e a valori in A, in simboli

$$f : A \times A \to A;$$

cioè, ad ogni coppia ordinata (a, b) di elementi di A fa corrispondere 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.

1.1.b 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].

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 $\mathbb{Q}^* = \mathbb{Q} \setminus \{0\}$, 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.

1.1.c Relazioni binarie

Una *relazione binaria* $\mathcal{R}$ da un insieme A ad un insieme B è un sottoinsieme S di $A \times B$. Se $a \in A$ e $b \in B$,
$$a\, \mathcal{R}\, b \quad \Leftrightarrow \quad (a, b) \in S\,.$$
Un tipico esempio di relazione da A a B è il grafico di una funzione $f : A \to B$, dove
$$S = \{(a, f(a)) \mid a \in A\} \subset A \times B.$$

Se $A = B$, una relazione binaria in A si dice *relazione d'equivalenza* e si indica col simbolo $\sim$, se $\forall a, b, c \in A$ essa soddisfa le seguenti proprietà:

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

1. $a \sim a$, <u>proprietà riflessiva</u>

2. $a \sim b \Rightarrow b \sim a$, <u>proprietà simmetrica</u>

3. $a \sim b$, $b \sim c \Rightarrow a \sim c$, <u>proprietà transitiva</u>.

Sia $\sim$ una relazione di equivalenza. Se $a \in A$, si chiama *classe di equivalenza* individuata da a l'insieme

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

Le classi di equivalenza dividono l'insieme A in sottoinsiemi mutuamente disgiunti. L'insieme formato dalle classi di equivalenza si chiama *insieme quoziente* di A e si indica $A/\sim$.

Esempi.

- Sia A l'insieme dei punti del piano ed O un fissato punto. La relazione $\mathcal{R}$ così definita

$$P\,\mathcal{R}\,Q \quad \Leftrightarrow \quad \overline{OP} = \overline{OQ}$$

è una relazione di equivalenza. La classe di equivalenza $[P]$ è la circonferenza di centro O e raggio $\overline{OP}$; l'insieme quoziente è costituito dalle circonferenze concentriche di centro O.

- Sia A l'insieme dei punti del piano e $d > 0$ un fissato numero. La seguente relazione $\mathcal{R}'$

$$P\,\mathcal{R}'\,Q \quad \Leftrightarrow \quad \overline{PQ} = d\,.$$

non è una relazione d'equivalenza.

- Sia A l'insieme delle rette del piano. Si considerino le relazioni $\mathcal{R}$, $\bar{\mathcal{R}}$ così definite

$$\begin{aligned} r\,\mathcal{R}\,s \quad &\Leftrightarrow \quad r \cap s = \varnothing\,; \\ r\,\bar{\mathcal{R}}\,s \quad &\Leftrightarrow \quad r \equiv s \text{ oppure } r \cap s = \varnothing\,. \end{aligned}$$

Le relazioni $\mathcal{R}$ e $\bar{\mathcal{R}}$ sono di equivalenza?

- Sia n un numero naturale fissato ed $a, b \in \mathbb{Z}$. Diremo che

$$a \equiv b \pmod{n} \quad \Leftrightarrow \quad b - a = kn\,, \quad k \in \mathbb{Z}\,.$$

Provare che la relazione "$\equiv \pmod{n}$" è una relazione di equivalenza su $\mathbb{Z}$ ed indicarne le classi di equivalenza.

1.1.d 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 $q(\alpha) \neq 0$.

1.1.e 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.

1.2 Matrici

1.2.a 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[2]:

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

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 $a_{ij} = 0$, $i \neq j$; $a_{ii} = 1$.

La matrice identica sarà indicata con $\mathrm{Id} \in \mathbb{K}^{n,n}$, o, più semplicemente, con I.

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

[2] Altri autori usano la notazione A^t.

1.2.b Operazioni su 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. Tuttavia, anche nel caso quadrato può accadere

$$AB \neq BA \,.$$

Due matrici si dicono *permutabili* se $AB = BA$.

Esempio.

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

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*.

Si vede facilmente che la matrice unità I è tale che

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

La moltiplicazione tra matrici soddisfa alle regole

$$A(BC) = (AB)C\,,$$
$$(A+B)C = AC + BC\,, \qquad A(B+C) = AB + AC\,,$$

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

Esempi ed esercizi.

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

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

- Provare che $^t(AB) = {}^tB\,^tA$.
- Se $A \in \mathbb{K}^{m,n}$, provare che $\tilde{A} = A\,^t\!A$ e $\bar{A} = {}^t\!AA$ 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

$$^t\!AA = I = A\,^t\!A\,.$$

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

$$(A+B)^2 = A^2 + 2AB + B^2 .$$

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).

1.3 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.

1.3.a Permutazioni

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 $(\sigma(1), \sigma(2), \ldots, \sigma(n))$ dell'insieme numerico $\{1, 2, \ldots, n\}$. La permutazione $(1, 2, \ldots, n)$ è chiamata *fondamentale*. 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.

1.3.b Determinante

Definizione 1.1. *Se* $A = (a_{ij}) \in \mathbb{K}^{n,n}$*, chiamiamo* determinante *di* A

$$\det A = \sum_{\sigma} \epsilon(\sigma)\, a_{1\sigma(1)} a_{2\sigma(2)} \dots a_{n\sigma(n)}\,,$$

dove la sommatoria è estesa a tutte le $n!$ *permutazioni dei numeri* $1, 2, \dots, n$*.*

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 1.1 (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*

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

dove c *è una fissata colonna (scelta arbitrariamente).*

Questa regola può essere assunta anche come definizione ricorsiva di determinante:

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

Quindi det è un'applicazione da $\mathbb{K}^{n,n}$ a $\mathbb{K}$.

Dal teorema di Laplace segue immediatamente che

1. $\det A = \det {}^t\!A$;
2. se la matrice B si ottiene da A moltiplicando una linea di A per un numero reale k, lasciando invariate le altre linee, allora $\det B = k \det A$.

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, $\det(A + B) \neq \det A + \det B$.
- Provare che per $k \in \mathbb{K}$ si ha $\det(kA) = k^n \det A$.
- Se A è una matrice triangolare (superiore o inferiore), allora

$$\det A = a_{11}a_{22}...a_{nn}.$$

Regole di calcolo. Procedendo per induzione, dopo averlo verificato direttamente per $n = 2$, si ha

1. se le matrici A e B differiscono soltanto per lo scambio di due linee parallele, allora $\det B = -\det A$;

   ne segue che:
2. se A ha due linee uguali, allora $\det A = 0$;
3. se A ha due linee proporzionali, allora $\det A = 0$;
4. se B si ottiene da A aggiungendo ad una certa linea di A un'altra linea di A moltiplicata per un fattore di proporzionalità, allora $\det B = \det A$;
5. la somma degli elementi di una linea per i complementi algebrici di un'altra linea è zero.

Teorema 1.2 (Regola di Binet)**.** *Se A e B sono due matrici quadrate di ordine n, si ha*

$$\det(AB) = (\det A)(\det B)\,.$$

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

1.3.c Matrici invertibili

Proposizione 1.1. *Se A è invertibile, allora*

1. $\det A \neq 0$;
2. $\det(A^{-1}) = 1/\det A$.

Dimostrazione. $AA^{-1} = I \Rightarrow \det(AA^{-1}) = \det I \Rightarrow \det(AA^{-1}) = 1 \Rightarrow \det A \neq 0$ e $\det(A^{-1}) = 1/\det A$ per il teorema di Binet. $\boxed{QED}$

Se $\det A \neq 0$, cioè A è *non singolare*, allora A è invertibile e si prova che

$$A^{-1} = \frac{1}{\det A}\operatorname{Adj}(A)\,,$$

dove $\operatorname{Adj}(A)$, *aggiunta classica di* A, è la matrice che ha al posto (i, j) il cofattore A_{ji} di a_{ji}:

$$\operatorname{Adj}(A) = \begin{pmatrix} A_{11} & A_{21} & \ldots\ldots & A_{n1} \\ A_{12} & A_{22} & \ldots\ldots & A_{n2} \\ \ldots\ldots & \ldots\ldots & \ldots\ldots & \ldots\ldots \\ A_{1n} & A_{2n} & \ldots\ldots & A_{nn} \end{pmatrix}$$

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
$$\operatorname{Adj}(A) = \begin{pmatrix} 8 & 0 & -4 \\ 3 & -2 & -1/2 \\ -12 & 0 & 2 \end{pmatrix} \quad \Rightarrow \quad A^{-1} = \begin{pmatrix} -1 & 0 & 1/2 \\ -3/8 & 1/4 & 1/16 \\ 3/2 & 0 & -1/4 \end{pmatrix}$$

- 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 $GL^-(n,\mathbb{R}) = \{A \in \mathbb{R}^{n,n} \mid \det A < 0\}$.

1.3.d 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 1.2. *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 $\lambda = 1$.

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*, ossia di trasformazioni del tipo:

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 di essa e di un'altra.

Dalle proprietà dei determinanti segue facilmente che se A e B sono due matrici equivalenti per righe, allora $\mathrm{rg}(A) = \mathrm{rg}(B)$.

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

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

dove $a_{1j_1} \neq 0$, $a_{2j_2} \neq 0$, $\dots$, $a_{pj_p} \neq 0$, con $1 \leq p \leq m$ e $1 \leq j_1 < j_2 < \dots < j_p \leq n$. Si dimostra facilmente che il rango di S è uguale al numero delle righe non nulle. Riassumendo, valgono i seguenti risultati.

Teorema 1.3. *Ogni matrice $A \in \mathbb{K}^{m,n}$ è equivalente per righe ad una matrice a scalini.*

Corollario 1.1. *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.*

1.4 Sistemi lineari

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

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

o, in forma 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:

$$AX = B \tag{1.4.1}$$

dove $A = (a_{ij}) \in \mathbb{K}^{m,n}$ è la matrice dei coefficienti, X è la colonna delle incognite e B quella dei termini noti, cioè

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

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).

Innanzitutto, una *soluzione del sistema* (1.4.1) è una n-pla $(\tilde{x}_1, \dots, \tilde{x}_n)$ che soddisfa simultaneamente tutte le equazioni di (1.4.1). Si dice anche che il sistema è *compatibile* se esiste almeno una soluzione.

Problema 1. Esso è risolto completamente dal seguente *criterio di Rouché-Capelli*, che afferma:

$$\text{il sistema è compatibile} \quad \Leftrightarrow \quad \mathrm{rg}(A) = \mathrm{rg}(\tilde{A}), \tag{1.4.2}$$

dove $\tilde{A} = (A|B)$ è la *matrice completa* del sistema, quindi $\mathrm{rg}(A) \leq \mathrm{rg}(\tilde{A})$.

Esempio. 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 $1 = \mathrm{rg}(A) \neq \mathrm{rg}(\tilde{A}) = 2$.

Problema 2. Sia $\mathrm{rg}(A) = \mathrm{rg}(\tilde{A}) = p$. Si hanno i seguenti casi.

$$\begin{array}{ll} p = n & \qquad\qquad \text{una sola soluzione,} \\ p < n & \qquad\qquad \infty^{n-p} \text{ soluzioni,} \end{array}$$

dove con '∞^{n-p} soluzioni' si intende infinite soluzioni dipendenti da $n - p$ parametri indipendenti.

Ne segue che se $rg(A) = m < n$ (*sistema normale*) il sistema è sempre compatibile.

N. B. *La risoluzione di un sistema compatibile di rango p si riconduce sempre a quella di un sistema di p equazioni in p incognite (con matrice dei coefficienti non singolare). Basta considerare come parametri le $n - p$ incognite, i cui coefficienti non concorrano a formare il minore di rango uguale a p.*

Problema 3. Si tratta dunque di risolvere un sistema di p equazioni in p incognite con $\det A \neq 0$. In tali ipotesi,

$$AX = B \quad \Leftrightarrow \quad X = A^{-1}B.$$

L'espressione esplicita delle soluzioni è fornita dal seguente teorema, di notevole importanza.

Teorema 1.4 (Teorema di Cramer)**.** *Sia $AX = B$ un sistema di p equazioni in p incognite, e sia $\det A \neq 0$. Allora il sistema ammette un'unica soluzione $(x_1, \dots, x_p)$, 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.

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

$$AX = O, \tag{1.4.3}$$

ammettono *sempre* la soluzione nulla $X = O$ (ed è l'unica se $|A| \neq 0$). Siamo perciò interessati alle soluzioni non nulle, dette anche *autosoluzioni* o *soluzioni proprie* . Se X' è una soluzione di (1.4.3), allora $\lambda X'$ è anch'essa una soluzione di (1.4.3) $\forall \lambda$; se X' e X'' sono soluzioni di (1.4.3), allora anche $X' + X''$ è una soluzione di (1.4.3). Chiaramente $\mathrm{rg}(A) = \mathrm{rg}(\tilde{A})$, e se $p = \mathrm{rg}(A)$ allora le soluzioni sono ∞^{n-p}.

Ad ogni sistema lineare non omogeneo (1.4.1) si può associare il sistema lineare omogeneo (1.4.3).

Si osservi che se X_0 è una soluzione particolare di (1.4.1) e $\tilde{X}$ una soluzione generica di (1.4.3), allora $\tilde{X} + X_0$ è una soluzione generica di (1.4.1); infatti

$$A(\tilde{X} + X_0) = A\tilde{X} + AX_0 = O + B = B.$$

Nota. Il metodo di risoluzione di Cramer ha grande interesse teorico, ma bisogna conoscere preliminarmente il rango di A e di $\tilde{A}$, la qual cosa richiede il calcolo di molti determinanti di matrici; il determinante di una matrice di ordine k comporta $k!$ operazioni.

Il *metodo di eliminazione di Gauss* prescinde dalla conoscenza a priori di quei ranghi ed inoltre è facilmente programmabile sui calcolatori; perciò è frequentemente usato nelle applicazioni. Esso si basa sull'osservazione che matrici equivalenti per righe corrispondono a sistemi di equazioni equivalenti, cioè che hanno lo stesso insieme di soluzioni. Il procedimento di Gauss consiste sostanzialmente nel ridurre la matrice completa $\tilde{A}$ ad una matrice a scalini $\tilde{S}$, equivalente per righe ad $\tilde{A}$. Ciò fatto, se il sistema è compatibile, si procede con la sostituzione successiva 'a ritroso' nel sistema. Il metodo di Gauss minimizza il numero di operazioni necessarie a risolvere un sistema lineare.

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.

Più in là sarà introdotto il concetto di *indipendenza lineare* (paragrafo 2.2). Da un punto di vista intuitivo, si può affermare che una rappresentazione cartesiana di S è costituita da condizioni indipendenti se queste sono in numero minimo; una rappresentazione parametrica è costituita da parametri indipendenti se questi sono in numero minimo. Questo numero è alla base del concetto di dimensione, come si vedrà.

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

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

PARTE I

GEOMETRIA ANALITICA

CAPITOLO 2

VETTORI DELLO SPAZIO ORDINARIO

Il concetto di vettore applicato in un punto è già noto dalla Fisica. Qui vogliamo introdurre il concetto di vettore libero, che ha suggerito la generalizzazione a spazi vettoriali astratti.

2.1 Definizioni

2.1.a Spazio $\mathbf{V}_3$

Si consideri lo spazio ordinario della geometria euclidea. Ogni segmento di estremi A e B individua due segmenti orientati AB e BA aventi orientazioni opposte; ciò è espresso scrivendo che

$$AB = -BA \quad \text{oppure} \quad \vec{AB} = -\vec{BA}.$$

Nell'insieme dei segmenti orientati dello spazio introduciamo la seguente relazione di equivalenza, detta di *equipollenza* (cioè di uguale valore ed efficacia)

$$AB \sim CD \quad \Leftrightarrow \quad \text{i punti medi di } AD,\ BC \text{ coincidono.}$$

Segue che AB è parallelo a CD (che si denota $AB \parallel CD$) e $\|AB\| = \|CD\|$, dove $\|AB\|$ indica il *modulo* o *lunghezza* del segmento AB. Le classi di equivalenza si chiamano *vettori* (liberi). Il vettore $\vec{u}$ individuato da $\vec{AB}$ e da tutti quelli ad esso equipollenti (come $\vec{CD}$) soddisfa l'uguaglianza $\vec{u} = [\vec{AB}] = [\vec{CD}]$. Il rappresentante $\vec{AB}$ di un vettore $\vec{u}$ si dice *vettore $\vec{u}$ applicato in A* e si indica $(\vec{u}, A)$. Il vettore $\vec{u}$ determina una traslazione dello spazio, da cui la parola, che proviene dal latino *vehere* = trasportare.

$$A + \vec{u} = B \quad \Leftrightarrow \quad \vec{u} = B - A = \vec{AB}.$$

La notazione $\vec{u} = B - A$, molto felice, è di W. R. Hamilton (1805–1865), il primo che ha dato il concetto preciso di vettore.

I segmenti AA, BB, ..., individuano il vettore nullo $\vec{0}$.

Un vettore non nullo è individuato dalla direzione, dal verso e dal modulo. Indichiamo con $\mathbf{V}_3$ l'insieme dei vettori liberi dello spazio e con $\mathbf{S}_3$ i punti dello spazio. Fissato un punto $O \in \mathbf{S}_3$, ad ogni punto $P \in \mathbf{S}_3$ si può associare un unico vettore $\vec{u} \in \mathbf{V}_3$, ponendo $\vec{u} = \vec{OP}$.

2.1.b Somma di vettori

Siano $\vec{u}$ e $\vec{v}$ due vettori, che vogliamo sommare. Se si considerano i rappresentanti indicati $\vec{u} = B - A$ e $\vec{v} = C - B$, poniamo

$$\vec{u} + \vec{v} = C - A$$

(che non dipende dai rappresentanti scelti). Si vede facilmente che $(\mathbf{V}_3, +)$ è un gruppo abeliano con elemento neutro $\vec{0}$ e $-\vec{u} = A - B$ se $\vec{u} = B - A$.

Si osservi che se consideriamo rappresentanti opportuni $\vec{u} = \vec{AB}$ e $\vec{v} = \vec{AD}$, allora $\vec{u} + \vec{v} = \vec{AC}$ è la diagonale del parallelogramma di lati AB e AD, in accordo con quanto si studia in Fisica.

2.1.c Differenza di vettori

Per definizione poniamo

$$\vec{u} - \vec{v} = \vec{u} + (-\vec{v}).$$

Se $\vec{u} = B - A$ e $\vec{v} = C - A$, allora $\vec{u} - \vec{v} = B - C$.

2.1.d Prodotto di un numero reale per un vettore

Sia $\lambda \in \mathbb{R}$ e $\vec{u} \in \mathbf{V_3}$. Vogliamo definire $\lambda\vec{u}$.

1. Se $\lambda = 0$, oppure $\vec{u} = \vec{0}$, poniamo $\lambda\vec{u} = \vec{0}$.
2. Se $\lambda \neq 0$ e $\vec{u} \neq \vec{0}$, il vettore $\lambda\vec{u}$ ha direzione coincidente con $\vec{u}$, verso concorde con quello di $\vec{u}$ se $\lambda > 0$, discorde se $\lambda < 0$, e inoltre

$$\|\lambda\vec{u}\| = |\lambda|\|\vec{u}\|.$$

Il numero $\lambda \in \mathbb{R}$ è detto *scalare*.

Valgono le seguenti proprietà immediate

$$\lambda(\vec{u} + \vec{v}) = \lambda\vec{u} + \lambda\vec{v}, \qquad \lambda(\mu\vec{u}) = (\lambda\mu)\vec{u},$$
$$(\lambda + \mu)\vec{u} = \lambda\vec{u} + \mu\vec{u}, \qquad 1\,\vec{u} = \vec{u}$$
$$\|\vec{u} + \vec{v}\| \leq \|\vec{u}\| + \|\vec{v}\|.$$

2.2 Dipendenza ed indipendenza lineare

2.2.a Dipendenza lineare

I vettori $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n \in \mathbf{V}_3$ si dicono *linearmente dipendenti* se e solo se esiste una n-pla $(\lambda_1, \dots, \lambda_n) \neq (0, \dots, 0)$ tale che

$$\lambda_1\vec{v}_1 + \lambda_2\vec{v}_2 + \cdots + \lambda_n\vec{v}_n = \vec{0}.$$

Infatti se $\lambda_n \neq 0$,

$$\vec{v}_n = -\frac{\lambda_1}{\lambda_n}\vec{v}_1 - \cdots - \frac{\lambda_{n-1}}{\lambda_n}\vec{v}_{n-1},$$

cioè $\vec{v}_n$ 'dipende' da $\vec{v}_1, \ldots, \vec{v}_{n-1}$. Più precisamente, si dice che $\vec{v}_n$ è combinazione lineare di $\vec{v}_1, \ldots, \vec{v}_{n-1}$. In generale, si dice che un vettore $\vec{v}$ è *combinazione lineare* di $\vec{v}_1, \ldots, \vec{v}_n$ con coefficienti $\lambda_1, \ldots \lambda_n$ se

$$\vec{v} = \lambda_1\vec{v}_1 + \cdots + \lambda_n\vec{v}_n\,. \tag{2.2.1}$$

2.2.b Indipendenza lineare

I vettori $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n \in \mathbf{V}_3$ si dicono *linearmente indipendenti* se e solo se *non* sono linearmente dipendenti, cioè

$$\lambda_1\vec{v}_1 + \cdots + \lambda_n\vec{v}_n = \vec{0} \quad \Rightarrow \quad \lambda_i = 0 \quad \forall i = 1, 2, \ldots, n.$$

Chiaramente vale sempre (sia nel caso dell'indipendenza, sia nel caso della dipendenza)

$$\lambda_i = 0 \text{ per ogni } i \quad \Rightarrow \quad \lambda_1\vec{v}_1 + \cdots + \lambda_n\vec{v}_n = \vec{0}\,.$$

2.2.c Significato geometrico

Siano $\vec{v}, \vec{v}_1, \vec{v}_2, \vec{v}_3 \in \mathbf{V}_3$. Allora

$$\begin{aligned} \vec{v} \text{ dipendente} \quad &\Leftrightarrow \quad \vec{v} = \vec{0} \\ \vec{v}_1, \vec{v}_2 \text{ dipendenti} \quad &\Leftrightarrow \quad \vec{v}_1 \parallel \vec{v}_2 \\ \vec{v}_1, \vec{v}_2, \vec{v}_3 \text{ dipendenti} \quad &\Leftrightarrow \quad \vec{v}_1, \vec{v}_2, \vec{v}_3 \text{ complanari.} \end{aligned}$$

(I vettori $\vec{v}_1, \vec{v}_2, \vec{v}_3$ sono detti complanari se i loro rappresentanti applicati in uno stesso punto appartengono ad un piani)

Quindi, se $\vec{v}_1$, $\vec{v}_2$, $\vec{v}_3$ sono indipendenti essi sono non complanari e possono considerarsi come vettori di un sistema di riferimento dello spazio. Ne segue che $n \geq 4$ vettori di $\mathbf{V}_3$ sono sempre dipendenti, quindi in $\mathbf{V}_3$ il massimo numero di vettori linearmente indipendenti è 3.

Sia $\mathbf{V}_2$ l'insieme dei vettori del piano; in $\mathbf{V}_2$ il massimo numero di vettori linearmente indipendenti è 2.

Sia $\mathbf{V}_1$ l'insieme dei vettori della retta; in $\mathbf{V}_1$ il massimo numero di vettori linearmente indipendenti è 1.

Si dice anche che la *dimensione* della retta è 1 ed una sua *base* è data da un vettore non nullo $\{\vec{e}_1\}$; la *dimensione* del piano è 2 ed una sua *base* è data da 2 vettori indipendenti $\{\vec{e}_1, \vec{e}_2\}$; la *dimensione* dello spazio è 3 ed una sua *base* è data da 3 vettori indipendenti $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$.

Sia $\mathcal{B} = \{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ una base di $\mathbf{V}_3$. Allora $\{\vec{v}, \vec{e}_1, \vec{e}_2, \vec{e}_3\}$ sono dipendenti e

$$\vec{v} = \lambda_1\vec{e}_1 + \lambda_2\vec{e}_2 + \lambda_3\vec{e}_3.$$

La terna di numeri $(\lambda_1, \lambda_2, \lambda_3)$ è univocamente individuata, e λ_1, λ_2, λ_3 sono dette le *coordinate* di $\vec{v}$ nella base $\mathcal{B}$. Naturalmente, nella base $\mathcal{B}$

$$\begin{array}{lll} \vec{e}_1 & \text{ha coordinate} & (1,0,0), \\ \vec{e}_2 & \text{ha coordinate} & (0,1,0), \\ \vec{e}_3 & \text{ha coordinate} & (0,0,1). \end{array}$$

Vediamo ora come condizioni vettoriali si traducano in problemi scalari tramite le coordinate. Siano

$$\vec{u}(u_1, u_2, u_3), \qquad \vec{v}(v_1, v_2, v_3), \qquad \vec{w}(w_1, w_2, w_3),$$

o, col simbolismo matriciale,

$$\vec{u} = {}^t(u_1\ u_2\ u_3), \quad \vec{v} = {}^t(v_1\ v_2\ v_3), \quad \vec{w} = {}^t(w_1\ w_2\ w_3).$$

Naturalmente $\vec{u} + \vec{v} = (u_1 + v_1, u_2 + v_2, u_3 + v_3)$, $\lambda\vec{u} = (\lambda u_1, \lambda u_2, \lambda u_3)$.

Allora:

$$a\vec{u} + b\vec{v} + c\vec{w} = \vec{0} \quad \Rightarrow \quad \begin{cases} au_1 + bv_1 + cw_1 = 0 \\ au_2 + bv_2 + cw_2 = 0 \\ au_3 + bv_3 + cw_3 = 0. \end{cases}$$

Si consideri

$$A = \begin{pmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{pmatrix}.$$

Se $\mathrm{rg}(A) = p$, allora si vede facilmente che p è il massimo numero di vettori indipendenti in $\{\vec{u}, \vec{v}, \vec{w}\}$, come si vede facilmente tenendo conto della risoluzione dei sistemi lineari (omogenei). Se consideriamo n vettori, la matrice A avrà 3 righe ed n colonne, quindi ancora $\mathrm{rg}(A) \leq 3$.

Se consideriamo il riferimento cartesiano (detto *affine*) $\mathcal{R}(Oxyz)$ associato a $\mathcal{B}$ tale che $\vec{e}_1, \vec{e}_2, \vec{e}_3$ siano i vettori unità sugli assi (applicati in O) si ha, con l'usuale simbolismo (usato per esempio in Fisica)

$$\vec{e}_1 = \vec{\imath}, \quad \vec{e}_2 = \vec{\jmath}, \quad \vec{e}_3 = \vec{k},$$
$$u_1 = u_x, \quad u_2 = u_y, \quad u_3 = u_z.$$

Tenendo conto della corrispondenza tra $\mathbf{V}_3$ e $\mathbf{S}_3$ indicata in (2.1), le *coordinate* (x, y, z) del punto P sono le coordinate del vettore $\overrightarrow{OP}$ nella base $\mathcal{B}$, cioè

$$\overrightarrow{OP} = x\vec{e}_1 + y\vec{e}_2 + z\vec{e}_3.$$

Se $P_i(x_i, y_i, z_i)$ per $i = 1, 2$, allora

$$\vec{P_1P_2} = \vec{OP_2} - \vec{OP_1} = (x_2 - x_1, y_2 - y_1, z_2 - z_1).$$

Esercizi.

- Siano dati i vettori $\vec{v}(1,2,3)$, $\vec{w}(1,1,1)$ e $\vec{v}_1(1,-1,0)$, $\vec{v}_2(0,1,1)$, $\vec{v}_3(2,2,4)$.
    1. Si possono scrivere $\vec{v}$ e $\vec{w}$ come combinazione lineare di $\vec{v}_1$, $\vec{v}_2$, $\vec{v}_3$? Se sì, trovare i coefficienti della combinazione lineare.
    2. $\vec{v}_2$ è combinazione lineare di $\vec{w}, \vec{v}_1, \vec{v}_3$?
- Si consideri $\mathbf{V}_2$ ed una sua base $\mathcal{B} = \{\vec{e}_1, \vec{e}_2\}$. Per quali valori di $t \in \mathbb{R}$, i vettori
$$\vec{v}_1 = (1-t)\vec{e}_1 + t\vec{e}_2, \qquad \vec{v}_2 = t\vec{e}_1 - \vec{e}_2$$
costituiscono una base di $\mathbf{V}_2$?
- Siano dati i seguenti vettori di $\mathbf{V}_3$ riferiti alla base $\mathcal{B} = \{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$:
$$\vec{v}_1(2-h, 4-2h, 2-h), \qquad \vec{v}_2(h, 3h, 2h), \qquad \vec{v}_3(1-h, 1-2h, h).$$
    1. determinare per quali valori di $h \in \mathbb{R}$ il vettore $\vec{w}(1-2h, 1-h, -5h)$ è combinazione lineare dei vettori $\vec{v}_1, \vec{v}_2, \vec{v}_3$.
    2. Esaminare il caso $h = 0$.

2.3 Cambiamenti di base

Siano $\mathcal{B} = \{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ e $\mathcal{B}' = \{\vec{e'}_1, \vec{e'}_2, \vec{e'}_3\}$ due basi di $\mathbf{V}_3$. Se (b_{1i}, b_{2i}, b_{3i}) sono le coordinate di $\vec{e'}_i$ rispetto alla base $\mathcal{B}$, la matrice $B = (b_{ij})$ $(i,j = 1,2,3)$ avrà determinante diverso da 0, ed è detta *matrice del cambiamento dalla base* $\mathcal{B}$ *alla base* $\mathcal{B}'$. Naturalmente, la matrice del cambiamento dalla base $\mathcal{B}'$ a $\mathcal{B}$ sarà B^{-1}.

Due basi si dicono *equiverse* se $\det A > 0$, *contraverse* se $\det A < 0$; "essere equiverse" è una relazione d'equivalenza.

Se $\vec{v} \in \mathbf{V}_3$, allora
$$\vec{v} = \sum_{i=1}^{3} v_i \vec{e}_i, \qquad \vec{v} = \sum_{i=1}^{3} v'_i \vec{e'}_i.$$
La matrice B permette di trovare il legame tra le coordinate v_i (relative alla base $\mathcal{B}$) e quelle v'_i (relative alla base $\mathcal{B}'$).

Esempio. Si consideri $\mathbf{V}_3$ riferito alla base $\mathcal{B} = \{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$.

1. Riconoscere che i vettori
$$\vec{v}_1 = \vec{e}_1 + \vec{e}_3, \quad \vec{v}_2 = \vec{e}_1 + 2\vec{e}_2, \quad \vec{v}_3 = 3\vec{e}_2 + \vec{e}_3$$
formano una base $\mathcal{B}'$ per $\mathbf{V}_3$ equiversa a $\mathcal{B}$.

2. Determinare le coordinate di $\vec{v}$ rispetto a $\mathcal{B}$, sapendo che $\vec{v}$ ha coordinate $(3,-1,2)$ rispetto a $\mathcal{B}'$.

(Suggerimento: per il punto 2, scrivere $\vec{v} = 3\vec{v}_1 - \vec{v}_2 + 2\vec{v}_3$ e poi sostituire $\vec{v}_i$.)

In generale, siano (a,b,c) le coordinate di $\vec{v}$ rispetto a $\mathcal{B}$ e (a',b',c') le coordinate di $\vec{v}$ rispetto a $\mathcal{B}'$. Allora risulta

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} = B \begin{pmatrix} a' \\ b' \\ c' \end{pmatrix},$$

quindi le coordinate cambiano con l'*inversa* della matrice del cambiamento di base.

2.4 Altre operazioni in $\mathbf{V}_3$

2.4.a Prodotto scalare

Il *prodotto scalare* tra due vettori di $\mathbf{V}_3$ è l'applicazione

$$g\colon \mathbf{V}_3 \times \mathbf{V}_3 \to \mathbb{R}, \quad g(\vec{u},\vec{v}) = \vec{u}\cdot\vec{v}$$

così definita:

$$\vec{u}\cdot\vec{v} = \begin{cases} 0 & \text{se } \vec{u} = \vec{0} \text{ o } \vec{v} = \vec{0} \\ \|\vec{u}\|\,\|\vec{v}\|\,\cos\widehat{\vec{u}\vec{v}} & \text{altrimenti.} \end{cases}$$

Il prodotto scalare gode delle seguenti proprietà (la cui dimostrazione è lasciata al lettore):

1. $\vec{u}\cdot\vec{v} = \vec{v}\cdot\vec{u}$, <u>commutatività</u>
2. $(\lambda\vec{u})\cdot\vec{v} = \vec{u}\cdot(\lambda\vec{v}) = \lambda(\vec{u}\cdot\vec{v})\ \forall\lambda\in\mathbb{R}$, <u>omogeneità</u>
3. $\vec{u}\cdot(\vec{v}+\vec{w}) = \vec{u}\cdot\vec{v} + \vec{u}\cdot\vec{w}$, <u>distributività</u>
4. $\vec{u}\cdot\vec{v} = 0 \quad\Leftrightarrow\quad \vec{u}\perp\vec{v}$.

Sia $\mathcal{B} = \{\vec{\imath},\vec{\jmath},\vec{k}\}$ una <u>base ortonormale</u> di $\mathbf{V}_3$ (cioè i tre vettori sono mutuamente ortogonali ed hanno modulo unitario); allora

$$\begin{aligned} &\vec{\imath}\cdot\vec{\imath} = 1, && \vec{\jmath}\cdot\vec{\jmath} = 1, && \vec{k}\cdot\vec{k} = 1, \\ &\vec{\imath}\cdot\vec{\jmath} = 0, && \vec{\jmath}\cdot\vec{k} = 0, && \vec{\imath}\cdot\vec{k} = 0. \end{aligned}$$

Posto $\vec{u} = u_1\vec{\imath} + u_2\vec{\jmath} + u_3\vec{k}$, $\vec{v} = v_1\vec{\imath} + v_2\vec{\jmath} + v_3\vec{k}$, usando le proprietà, si ha:

$$\vec{u}\cdot\vec{v} = u_1v_1 + u_2v_2 + u_3v_3. \tag{2.4.2}$$

Si osservi che se $\mathcal{B}$ non fosse ortonormale, l'espressione precedente non sarebbe così semplice. Inoltre, si vede facilmente che

$$\vec{u}\cdot\vec{u} = \|\vec{u}\|^2 = u_1^2 + u_2^2 + u_3^2, \qquad \cos\widehat{\vec{u}\vec{v}} = \frac{\vec{u}\cdot\vec{v}}{\|\vec{u}\|\,\|\vec{v}\|}. \tag{2.4.3}$$

Dunque, *conoscendo il prodotto scalare*, si può determinare la lunghezza di un vettore e l'angolo tra due vettori. Tutta la geometria euclidea si può dedurre dalla nozione di prodotto scalare (2.4.2), assumendo le (2.4.3) come *definizioni* di *modulo* e di *angolo* tra due vettori (vedi esercizi).

Poiché $\mathcal{B}$ è ortonormale segue anche che

$$v_1 = \vec{v}\cdot\vec{\imath} = \|\vec{v}\|\,\cos\widehat{\vec{v}\vec{\imath}}, \quad v_2 = \vec{v}\cdot\vec{\jmath} = \|\vec{v}\|\,\cos\widehat{\vec{v}\vec{\jmath}}, \quad v_3 = \vec{v}\cdot\vec{k} = \|\vec{v}\|\,\cos\widehat{\vec{v}\vec{k}}.$$

Si chiamano *coseni direttori di* $\vec{v}$ le componenti rispetto a $\mathcal{B}$ del *versore* di $\vec{v}$, essendo per definizione

$$\operatorname{vers}\vec{v} = \frac{\vec{v}}{\|\vec{v}\|},$$

quindi se $\vec{v} = (v_1, v_2, v_3)$, i suoi coseni direttori sono

$$\frac{v_1}{\|\vec{v}\|} = \cos\widehat{\vec{v}\vec{\imath}}, \quad \frac{v_2}{\|\vec{v}\|} = \cos\widehat{\vec{v}\vec{\jmath}}, \quad \frac{v_3}{\|\vec{v}\|} = \cos\widehat{\vec{v}\vec{k}},$$

il che spiega il nome (danno, infatti, la *direzione* di $\vec{v}$).

La *componente ortogonale* di $\vec{v}$ rispetto ad un vettore non nullo $\vec{u}$ è il <u>numero reale</u>

$$v_{\vec{u}} = \|\vec{v}\|\cos\widehat{\vec{u}\vec{v}} = \vec{v}\cdot\operatorname{vers}\vec{u} \in \mathbb{R}.$$

La *proiezione ortogonale* di $\vec{v}$ su $\vec{u}$ è il <u>vettore</u>

$$\vec{v}_{\vec{u}} = v_{\vec{u}}\operatorname{vers}\vec{u}.$$

Esempi ed esercizi.

- Determinare l'angolo tra la diagonale di un cubo ed un lato.

  Si consideri il cubo individuato dai vettori $\vec{\imath}, \vec{\jmath}, \vec{k}$. Allora la diagonale è data dal vettore $\vec{OP} = (1, 1, 1)$. Poiché $\vec{\imath} = (1, 0, 0)$, segue

  $$\cos\widehat{\vec{v}\vec{\imath}} = \frac{\vec{v}\cdot\vec{\imath}}{\|\vec{v}\|\,\|\vec{\imath}\|} = \frac{1}{\sqrt{3}} \quad\Rightarrow\quad \widehat{\vec{v}\vec{\imath}} = \arccos\frac{1}{\sqrt{3}}.$$

- Determinare l'angolo tra la diagonale di un cubo e la diagonale di una faccia.

- Se $\mathcal{A}$ è l'area del parallelogramma costruito sui vettori $\vec{u}$ e $\vec{v}$, provare che

$$\mathcal{A}^2 = \begin{vmatrix} \vec{u}\cdot\vec{u} & \vec{u}\cdot\vec{v} \\ \vec{v}\cdot\vec{u} & \vec{v}\cdot\vec{v} \end{vmatrix}.$$

- Dal punto precedente dedurre che

$$\begin{aligned} |\vec{u}\cdot\vec{v}| &\leq \|\vec{u}\|\,\|\vec{v}\|, && \text{disug. di } \underline{\text{Cauchy–Schwarz}} \\ \|\vec{u}+\vec{v}\| &\leq \|\vec{u}\| + \|\vec{v}\|, && \text{disug. } \underline{\text{triangolare}}. \end{aligned}$$

- Verificare l'identità

$$\|\vec{u}+\vec{v}\|^2 + \|\vec{u}-\vec{v}\|^2 = 2(\|\vec{u}\|^2 + \|\vec{v}\|^2)$$

ed esprimerne il significato geometrico.

2.4.b Prodotto vettoriale

Il *prodotto vettoriale* tra vettori è l'applicazione

$$h\colon \mathbf{V}_3 \times \mathbf{V}_3 \to \mathbf{V}_3, \quad h(\vec{u},\vec{v}) = \vec{u}\wedge\vec{v}$$

così definita:

$$\vec{u}\wedge\vec{v} = \begin{cases} \vec{0} & \text{se } \vec{u} \parallel \vec{v} \\ \vec{w} & \end{cases}$$

dove $\vec{w}$ ha direzione perpendicolare a $\vec{u}$ e $\vec{v}$, verso tale che la terna $(\vec{u},\vec{v},\vec{w})$ sia equiversa a $(\vec{\imath},\vec{\jmath},\vec{k})$ e modulo $\|\vec{w}\| = \|\vec{u}\|\,\|\vec{v}\| \sin\widehat{\vec{u}\vec{v}}$.

Il prodotto vettoriale verifica le seguenti proprietà (la cui dimostrazione è lasciata al lettore):

1. $\vec{u}\wedge\vec{v} = -\vec{v}\wedge\vec{u}$, <u>anticommutatività</u>,
2. $(\lambda\vec{u})\wedge\vec{v} = \vec{u}\wedge(\lambda\vec{v}) = \lambda(\vec{u}\wedge\vec{v})$ $\forall\lambda\in\mathbb{R}$, <u>omogeneità</u>,
3. $\vec{u}\wedge(\vec{v}+\vec{w}) = \vec{u}\wedge\vec{v} + \vec{u}\wedge\vec{w}$, <u>distributività</u>.,

Se $\mathcal{B} = \{\vec{\imath},\vec{\jmath},\vec{k}\}$ è una base ortonormale, allora

$$\vec{u}\wedge\vec{v} = \begin{vmatrix} \vec{\imath} & \vec{\jmath} & \vec{k} \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix}$$

Per dimostrare la precedente espressione, si scrivano $\vec{u} = u_1\vec{\imath}+u_2\vec{\jmath}+u_3\vec{k}$ e $\vec{v} = v_1\vec{\imath}+v_2\vec{\jmath}+v_3\vec{k}$ e si tenga conto delle identità $\vec{\jmath}\wedge\vec{k} = \vec{\imath}$, $\vec{\imath}\wedge\vec{\jmath} = \vec{k}$, $\vec{\imath}\wedge\vec{k} = -\vec{\jmath}$.

2.4.c Prodotto misto

Il *prodotto misto* di tre vettori $\vec{u}, \vec{v}, \vec{w} \in \mathbf{V}_3$ è dato dal numero reale $(\vec{u} \wedge \vec{v}) \cdot \vec{w} \in \mathbb{R}$. Considerata una base ortonormale $\mathcal{B} = \{\vec{\imath}, \vec{\jmath}, \vec{k}\}$ si ha

$$(\vec{u} \wedge \vec{v}) \cdot \vec{w} = \begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix}$$

Infatti, basta calcolare le coordinate di $\vec{u} \wedge \vec{v}$ come detto nel paragrafo precedente e tener conto dell'espressione analitica del prodotto scalare.

Esercizi.

- Confrontare le definizioni di prodotto scalare e vettoriale con quelle impartite nei corsi di Fisica.
- Provare che $\|\vec{u} \wedge \vec{v}\| = \mathcal{A}$, area del parallelogramma costruito sui vettori $\vec{u}$ e $\vec{v}$.
- Provare che $|(\vec{u} \wedge \vec{v}) \cdot \vec{w}| = \mathcal{V}$, volume del parallelepipedo costruito sui vettori $\vec{u}$, $\vec{v}$ e $\vec{w}$.

  Naturalmente, il volume del tetraedro costruito sui vettori $\vec{u}$, $\vec{v}$, $\vec{w}$ è $1/6\,\mathcal{V}$.
- Dati i vettori $\vec{v}_1(1, 2, 0)$, $\vec{v}_2(1, -1, 2)$, $\vec{v}_3(3, 0, 2)$, determinare il vettore $\vec{w} = (\vec{v}_1 \wedge \vec{v}_2) \wedge \vec{v}_3$. Verificare che $\{\vec{v}_1, \vec{v}_2, \vec{w}\}$ sono linearmente dipendenti ed esprimere $\vec{w}$ come combinazione lineare di $\vec{v}_1$ e $\vec{v}_2$.
- Provare che il prodotto vettoriale non soddisfa la proprietà associativa, ossia che

  $$\vec{v} \wedge (\vec{w} \wedge \vec{z}) \neq (\vec{v} \wedge \vec{w}) \wedge \vec{z}.$$

  Interpretare geometricamente il risultato.

CAPITOLO 3

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. In esso considereremo un riferimento ortonormale costituito da un punto O e da una base ortormale $\mathcal{B} = \{\vec{\imath}, \vec{\jmath}, \vec{k}\}$ di vettori dello spazio. Le rette per O individuate dai vettori $\vec{\imath}$, $\vec{\jmath}$, $\vec{k}$, si dicono *assi coordinati* e nell'ordine asse x, asse y, asse z. Diremo perciò che si è scelto un riferimento cartesiano $RC(Oxyz)$.

La geometria analitica 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.

3.1 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.

3.1.a Piano: equazione cartesiana

$$\begin{vmatrix} x - x_1 & y - y_1 & z - z_1 \\ x_2 - x_1 & y_2 - y_1 & z_2 - z_1 \\ x_3 - x_1 & y_3 - y_1 & z_3 - z_1 \end{vmatrix} = 0 \quad \Leftrightarrow \quad \begin{vmatrix} x & y & z & 1 \\ x_1 & y_1 & z_1 & 1 \\ x_2 & y_2 & z_2 & 1 \\ x_3 & y_3 & z_3 & 1 \end{vmatrix} = 0. \tag{3.1.1}$$

Sviluppando il determinante, si ha un'*equazione cartesiana del piano*

$$ax + by + cz + d = 0, \qquad (a, b, c) \neq (0, 0, 0), \tag{3.1.2}$$

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

$$\vec{n} \cdot \overrightarrow{P_0 P} = 0 \qquad \forall P \in \alpha \tag{3.1.3}$$

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.

3.1.b Piano: equazioni parametriche

$$\overrightarrow{P_1 P} = u\overrightarrow{P_1 P_2} + v\overrightarrow{P_1 P_3}, \qquad u, v \in \mathbb{R}, \tag{3.1.4}$$

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 $\overrightarrow{P_1 P_2} = (0, 1, 1)$, $\overrightarrow{P_1 P_3} = (0, 0, 1)$, dunque

$$\begin{cases} x = 1 \\ y = u \\ z = u + v \end{cases},$$

da cui l'equazione cartesiana $x = 1$.

3.1.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

$$\begin{array}{rcccl} \text{sistema incompatibile} & \Leftrightarrow & \mathrm{rg}(A) \neq \mathrm{rg}(\tilde{A}) & \Leftrightarrow & \alpha \cap \alpha' = \varnothing, \\ \text{sistema compatibile} & \Leftrightarrow & \mathrm{rg}(A) = \mathrm{rg}(\tilde{A}) & \Leftrightarrow & \alpha \cap \alpha' \neq \varnothing. \end{array}$$

Inoltre

$$\begin{array}{lcccl} \mathrm{rg}(A) = \mathrm{rg}(\tilde{A}) = 2 & \Leftrightarrow & \infty^1 \text{ soluzioni} & \Leftrightarrow & \alpha \cap \alpha' = r, \\ \mathrm{rg}(A) = \mathrm{rg}(\tilde{A}) = 1 & \Leftrightarrow & \infty^2 \text{ soluzioni} & \Leftrightarrow & \alpha \equiv \alpha', \end{array}$$

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 passante per il punto $(3, -1, 5)$ è

$$1(x - 3) + (-1)(y + 1) + 2(z - 5) = 0.$$

3.2 Retta

Due punti distinti P_1 e P_2 (quindi $\vec{P_1P_2} \neq \vec{0}$)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.

3.2.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, solo se ha senso la scrittura seguente, 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 3.1.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 $\vec{v}$ parallelo ad r. Ovviamente, se $P_1, P_2 \in r$ e $P_1 \neq P_2$, allora $\vec{v} = \overrightarrow{P_1P_2}$ 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 $\vec{v} = (l, m, n)$ 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 $\vec{v}$ 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.

3.2.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.

3.2.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

$$\begin{cases} x = x_1 + t(x_2 - x_1) \\ y = y_1 + t(y_2 - y_1) \end{cases} \qquad t \in \mathbb{R}.$$

3.2.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.

$$
\begin{array}{rclcl}
\alpha \parallel \alpha' & \Rightarrow & \vec{n} \parallel \vec{n'} & \Rightarrow & (a,b,c) \sim (a',b',c') \\
\alpha \perp \alpha' & \Rightarrow & \vec{n} \perp \vec{n'} & \Rightarrow & aa' + bb' + cc' = 0 \\
\alpha \perp r & \Rightarrow & \vec{n} \parallel \vec{r} & \Rightarrow & (a,b,c) \sim (l,m,n) \\
\alpha \parallel r & \Rightarrow & \vec{n} \perp \vec{r} & \Rightarrow & al + bm + cn = 0 \\
r \parallel r' & \Rightarrow & \vec{r} \parallel \vec{r'} & \Rightarrow & (l,m,n) \sim (l',m',n') \\
r \perp r' & \Rightarrow & \vec{r} \perp \vec{r'} & \Rightarrow & ll' + mm' + nn' = 0
\end{array}
$$

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

$$\cos \widehat{\alpha\alpha'} = \cos \widehat{\vec{n}\vec{n'}} = \frac{\vec{n} \cdot \vec{n'}}{\|\vec{n}\|\,\|\vec{n'}\|} = \pm \frac{aa' + bb' + cc'}{\sqrt{a^2+b^2+c^2}\,\sqrt{a'^2+b'^2+c'^2}}$$

$$\sin \widehat{\alpha r} = |\cos \widehat{\vec{n}\vec{r}}| = \frac{|\vec{n} \cdot \vec{r}|}{\|\vec{n}\|\,\|\vec{r}\|} = \frac{|al + bm + cn|}{\sqrt{a^2+b^2+c^2}\,\sqrt{l^2+m^2+n^2}}$$

3.2.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.

3.3 Fasci e stelle

Siano α ed α' due piani. Se $\alpha \cap \alpha' = r$, si chiama *fascio di piani proprio* di asse r la totalità dei piani dello spazio passanti per r, che si dice *asse* del fascio proprio. Se $\alpha \parallel \alpha'$, la totalità dei piani dello spazio paralleli ad α (e ad α') costituisce il *fascio di piani improprio* individuato dalla giacitura di α (e di α').

Se $\alpha\colon ax + by + cz + d = 0$ e $\alpha'\colon a'x + b'y + c'z + d' = 0$ il fascio è rappresentato da

$$\lambda(ax + by + cz + d) + \mu(a'x + b'y + c'z + d') = 0$$

al variare dei parametri omogenei λ e μ, con $(\lambda, \mu) \neq (0, 0)$. Se $\lambda \neq 0$, ponendo $k = \mu/\lambda$, il fascio è rappresentato dall'equazione

$$ax + by + cz + d + k(a'x + b'y + c'z + d') = 0,$$

che esplicita il fatto che i piani di un fascio sono ∞^1.

Si osservi che nell'equazione precedente, al variare di k in $\mathbb{R}$, il piano α' non è rappresentato; esso però si può pensare ottenuto per $k = \pm\infty$. Ciò porta ad ampliare $\mathbb{R}$ in modo spontaneo, aggiungendo <u>un solo</u> punto improprio (mentre in Analisi l'ampliamento è fatto con i due punti impropri $\pm\infty$):

$$\bar{\mathbb{R}} = \mathbb{R} \cup \{\infty\}.$$

La totalità dei piani dello spazio passanti per un fissato punto $P_0(x_0, y_0, z_0)$ costituisce la *stella di piani* di centro P_0. I piani di una stella sono ∞^2, come si vede dall'equazione

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0,$$

dove a, b, c sono parametri omogenei tali che $(a, b, c) \neq 0$.

La totalità delle rette dello spazio passanti per un punto P_0 costituisce la *stella propria* di centro P_0; la totalità delle rette parallele ad una data retta r costituisce la *stella impropria* individuata dalla direzione di r.

Nota. Un fascio è sinonimo di ente con ∞^1 elementi, una stella di ente con ∞^2 elementi.

Nel piano i concetti espressi precedentemente danno luogo soltanto a quello di *fascio di rette proprio* e *fascio di rette improprio* come la totalità delle rette del piano passanti per un punto e la totalità delle rette del piano aventi una data direzione.

Esempi ed esercizi.

- Date le rette $r = \alpha \cap \beta$ ed $r' = \alpha' \cap \beta'$, studiare $r \cap r'$ discutendo il sistema lineare costituito dalle equazioni di α, β, α', β'. *Suggerimento: si proceda come nel paragrafo 3.1.c.*
- Data la retta $r = \alpha \cap \beta$ ed il piano γ, studiare $r \cap \gamma$ come nel punto precedente.
- Trovare il piano passante per $A(0, 2, -1)$ e per la retta

  $$r\colon \begin{cases} x + 2y + z = 0 \\ x - z = 0 \end{cases}$$

  Poiché $A \notin r$, il piano è univocamente individuato. Allora si considera il fascio di piani di asse r e si impone il passaggio per A del generico piano.

  Il piano generico $x + 2y + z + k(x - z) = 0$ passa per A se $k = -3$, quindi il piano cercato è $x - y - 2z = 0$. Si verifichi che questo piano passa per A e contiene r.
- Si risolva l'esercizio precedente considerando il piano passante per A e per due punti scelti di r.
- Trovare il piano comune al fascio di asse r e alla stella di centro $A \notin r$ (vedere gli esercizi precedenti).
- Nel piano siano date le rette $r\colon 3x - y - 4 = 0$, $s\colon x + 2y + 1 = 0$. Il fascio di rette individuato da r e da s ha equazione

  $$3x - y - 4 + k(x + 2y + 1) = 0, \qquad k \in \bar{\mathbb{R}}.$$

  e rappresenta la totalità delle rette passanti per $P_0 = r \cap s$. Se k varia in $\mathbb{Z}$, <u>non</u> si ha un fascio di rette, poiché non sono rappresentate <u>tutte</u> le rette passanti per P_0.
- Scrivere il fascio di rette del piano $\alpha\colon 3x - y + 5z + 1 = 0$ di centro $P_0(0, 1, 0) \in \alpha$.

  Sia r una retta per P_0 <u>non</u> contenuta in α; ad esempio $r\colon x = 0$, $z = 0$. L'equazione $x + kz = 0$, con $k \in \mathbb{R}$, rappresenta il fascio di piani di asse r e

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

  rappresenta il fascio di rette richiesto.

- Nella stella di piani di centro $P_0(2, -3, 1)$, trovare il piano che passa per l'asse z. (Suggerimento: se A e B sono due punti dell'asse z, non allineati con P_0, il piano ABP_0 è quello richiesto.)

3.4 Superficie e curve

Abbiamo visto che (nello spazio) un piano si rappresenta con un'equazione, mentre una retta con due equazioni. Ogni equazione, ponendo un vincolo tra le incognite, riduce di uno il grado di libertà delle incognite. Ciò traduce il fatto che il piano è un ente di dimensione 2, mentre la retta è un ente di dimensione 1.

3.4.a Definizioni

Chiamiamo *superficie* Σ il luogo dei punti $P(x, y, z)$ dello spazio le cui coordinate verificano un'equazione del tipo

$$f(x, y, z) = 0,$$

che è detta *equazione cartesiana* di Σ.

Se f è un polinomio, la superficie si dirà *algebrica*: le superfici algebriche di grado 1 sono i piani, quelle di grado 2 si chiamano *quadriche*.

Una superficie si può rappresentare *parametricamente* tramite equazioni del tipo

$$x = x(u, v), \quad y = y(u, v), \quad z = z(u, v), \qquad (u, v) \in A \subset \mathbb{R}^2$$

dove A è un aperto del piano; quindi $(x(u, v), y(u, v), z(u, v)) \in \Sigma$ dipende da <u>due</u> parametri.

Un punto P descrive una *curva* $\mathcal{C}$ dello spazio se esso dipende da <u>un solo</u> parametro:

$$x = x(t), \quad y = y(t), \quad z = z(t), \qquad t \in I \subset \mathbb{R},$$

che rappresentano *equazioni parametriche* di $\mathcal{C}$. Eliminando il parametro si perviene (spesso con difficoltà) ad *equazioni cartesiane* di $\mathcal{C} = \Sigma_1 \cap \Sigma_2$, dove

$$\Sigma_1\colon f_1(x, y, z) = 0, \qquad \Sigma_2\colon f_2(x, y, z) = 0.$$

Le curve $\mathcal{C}_u$ e $\mathcal{C}_v$, dette *curve coordinate*, (ottenute rispettivamente per $v = \text{cost}$ e $u = \text{cost}$) al variare di u e v costituiscono su Σ un reticolato, usato spesso per visualizzare il grafico di Σ.

Esempio. Se $\Sigma\colon f(x, y, z) = 0$ e $\Sigma'\colon g(x, y, z) = 0$ sono equazioni algebriche di primo grado, ognuna di esse rappresenta un piano ed il loro sistema rappresenta la retta $r = \Sigma \cap \Sigma'$, che è dunque una particolare curva.

Si osservi che, date due superficie $\Sigma_1\colon f_1(x,y,z)=0$ e $\Sigma_2\colon f_2(x,y,z)=0$, allora

$$\Sigma_1\cap\Sigma_2\colon \begin{cases} f_1(x,y,z)=0,\\ f_2(x,y,z)=0,\end{cases} \qquad \Sigma_1\cup\Sigma_2\colon f_1(x,y,z)\,f_2(x,y,z)=0.$$

Evidentemente, $\Sigma_1\cap\Sigma_2$ è una curva, mentre $\Sigma_1\cup\Sigma_2$ è una superficie. L'intersezione di una curva $\mathcal{C}$ e di una superficie Σ è, generalmente, un insieme finito di punti, eventualmente vuoto. Naturalmente, può succedere $\mathcal{C}\subset\Sigma$ in casi particolari.

Esempi ed esercizi.

- Le equazioni

  $$\Sigma\colon \begin{cases} x=3u^2-uv\\ y=e^{-u}+v\\ z=v^2\end{cases}$$

  rappresentano una superficie. Se fissiamo uno dei parametri, otteniamo una curva. Ponendo, ad esempio, $u=1$, si ha

  $$\mathcal{C}_v\colon \quad x=3-v,\quad y=\frac{1}{e}+v,\quad z=v^2.$$

  Analogamente se $v=0$ si ha la curva

  $$\mathcal{C}_u\colon \quad x=3u^2,\quad y=e^{-u},\quad z=0,$$

  che giace tutta nel piano $z=0$.

- Provare che

  $$\Sigma\colon x=u,\quad y=2u^2v,\quad z=v$$

  è una superficie algebrica che contiene la curva

  $$\mathcal{D}\colon x=2\sin t\cos t,\quad y=2t(1-\cos 4t),\quad z=2t.$$

- Siano

  $$\Sigma_1\colon x^2+y^2+z^2=1,\qquad \Sigma_2\colon x^2+y^2-x=0$$

  due superficie. Allora il sistema

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

  rappresenta la curva $\mathcal{C}=\Sigma_1\cap\Sigma_2$. Ma il sistema precedente è equivalente al seguente

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

cioè $\mathcal{C}$ si può scrivere anche come intersezione delle superficie $\Sigma_1'\colon x = 1 - z^2$ e Σ_2.

Si osservi che una stessa curva può essere rappresentata in più modi come intersezione di superfici (si pensi anche ad una retta).

Equazioni parametriche di $\mathcal{C}$ sono

$$x = \sin^2\phi, \quad y = \sin\phi\cos\phi, \quad z = \cos\phi,$$

come si verifica facilmente.[1]

3.4.b Curve piane

Una curva $\mathcal{C}$ dello spazio si dice *piana* se esiste un piano che la contiene. Naturalmente una sezione di una superficie con un piano dà una curva piana.

Esempio. Data la curva

$$\mathcal{C}\colon x = t^2 - 1, \quad y = t^2 + 1, \quad z = 2t,$$

dimostriamo che è piana. Bisogna vedere se esiste un piano

$$ax + by + cz + d = 0, \qquad (a, b, c) \neq (0, 0, 0)$$

tale che $ax(t) + by(t) + cz(t) + d = 0$ per ogni t. Ora

$$a(t^2 - 1) + b(t^2 + 1) + 2tc + d = 0 \quad \Rightarrow \quad (a + b)t^2 + 2tc + d - a + b = 0,$$

che porta (per il principio di identità dei polinomi) al sistema omogeneo

$$a + b = 0, \quad 2c = 0, \quad d - a + b = 0,$$

che ha soluzioni $c = 0$, $a = -b$, $d = -2b$. Quindi $\mathcal{C}$ è piana ed è contenuta nel piano $\alpha\colon x - y + 2 = 0$. Si noti che si ha $t = z/2$, da cui le equazioni cartesiane di $\mathcal{C}$:

$$x = \frac{z^2}{4} - 1, \qquad y = \frac{z^2}{4} + 1,$$

oppure

$$4x - z^2 + 4 = 0, \qquad x - y + 2 = 0.$$

Se una curva $\mathcal{C}$ giace nel piano $z = 0$, la sua equazione cartesiana in questo piano è del tipo $f(x, y) = 0$, ed una rappresentazione parametrica è

$$x = x(t), \quad y = y(t); \qquad t \in \mathbb{R}.$$

Si noti che $\mathcal{C}$ può essere ben lontana dall'idea di curva che abbiamo.

[1] Questa curva è detta *finestra di Viviani*.

Osservazioni. Il simmetrico rispetto all'asse x del punto $P(x, y)$ è $P'(x, -y)$; analogamente, il simmetrico rispetto all'asse y è $P''(-x, y)$ ed il simmetrico rispetto ad O è $P^*(-x, -y)$. Quindi, una curva $\mathcal{C}$ è

$$\begin{array}{lcl} \text{simmetrica rispetto all'asse } x & \Leftrightarrow & P, P' \in \mathcal{C}, \\ \text{simmetrica rispetto all'asse } y & \Leftrightarrow & P, P'' \in \mathcal{C}, \\ \text{simmetrica rispetto all'origine} & \Leftrightarrow & P, P^* \in \mathcal{C}. \end{array}$$

Esempi.

- $f(x, y) = x^2 + y^2 + 1 = 0$ non ha punti reali.
- $f(x, y) = x^2 + y^2 = 0$ ha il solo punto reale $(0, 0) \in \mathbb{R}^2$, mentre in $\mathbb{C}^2$ si scrive come unione delle due rette $\{y = ix\}$ e $\{y = -ix\}$.
- $f(x, y) = (x^2 + y^2)[(x - 1)^2 + (y - 1)^2] = 0$ ha solo i punti reali $\{(0, 0), (1, 1)\} \subset \mathbb{R}^2$.
- Sia $\mathcal{C}$ una curva del piano $z = 0$ di equazione $f(x, y) = 0$. Provare che se $f(x, y) = 0 \Rightarrow f(x, -y) = 0$ allora $\mathcal{C}$ è simmetrica rispetto all'asse x; se $f(x, y) = 0 \Rightarrow f(-x, y) = 0$ allora $\mathcal{C}$ è simmetrica rispetto all'asse y; se $f(x, y) = 0 \Rightarrow f(-x, -y) = 0$ allora $\mathcal{C}$ è simmetrica rispetto all'origine.

Una curva algebrica di grado 1 è una retta, una curva algebrica di grado 2 è una *conica*. Nella scuola secondaria sono già state incontrate le seguenti coniche:

$$\begin{array}{lr} x^2 + y^2 + 2ax + 2by + c = 0 & \underline{\text{circonferenza}}, \\ \dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 1 & \underline{\text{ellisse}}, \\ \dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} = 1 & \underline{\text{iperbole}}, \\ y = ax^2 + bx + c & \underline{\text{parabola}}. \end{array}$$

Vedremo più avanti che le ellissi, iperboli e parabole esauriscono tutti i possibili tipi di coniche. Il nome "coniche" (anticamente "sezioni coniche") deriva dal fatto che esse sono le sezioni piane di un cono circolare (retto o obliquo). Come inventore delle coniche si ritiene generalmente Menecmo (circa 350 a.C.), mentre chi le ha studiate in profondità è Apollonio di Perga (II sec. a.C.). Le coniche sono molto importanti nelle applicazioni. SI pensi che noi vediamo attraverso sezioni di coni visivi.

Osservazione. La distinzione tra curve algebriche e curve non algebriche (o trascendenti) riguarda le coordinate cartesiane.

Esempio. L'equazione $y = 3x + 1$ è algebrica, mentre in coordinate polari (ρ, θ) la stessa retta si rappresenta con una equazione non algebrica

$$\rho \sin\theta = 3\rho\cos\theta + 1.$$

Le coordinate polari risultano molto convenienti per rappresentare curve che si avvolgono intorno ad un punto.

$\rho = d$	circonferenza di centro l'origine e raggio d,
$\rho = c\theta, \quad c > 0, 0 \le \theta \le +\infty$	spirale di Archimede,
$\rho = ae^{\theta}, \quad a > 0$	spirale logaritmica,
$\rho = \dfrac{c}{\theta}, \quad c > 0, 0 \le \theta \le +\infty$	spirale iperbolica.

Sia $\mathcal{C}^h$ una curva algebrica di grado h, allora in generale $\mathcal{C}^m \cap \mathcal{C}^n$ è costituito da mn punti. In particolare $\mathcal{C}^n$ incontra una retta in posizione generica in n punti contati ciascuno con la sua molteplicità.

Esempi ed esercizi.

- Si consideri la curva $\mathcal{C}$ di equazioni parametriche

$$x = \frac{2t}{t^2+1}, \qquad y = \frac{1}{t^2+1}, \qquad t \in \mathbb{R}.$$

  Eliminando il parametro $(x/y = 2t)$ si perviene all'equazione cartesiana

$$4y^3 + x^2y - 4y^2 = 0,$$

  che si scompone nella retta $y = 0$ e nella conica $4y^2 + x^2 - 4y = 0$. Si osservi che nel passaggio dalle equazioni parametriche a quella cartesiana sono stati introdotti punti che non fanno parte di $\mathcal{C}$.

- Si consideri la curva
$$\begin{cases} x = \sin t \\ y = \sin t \end{cases}, \quad t \in \mathbb{R}.$$

  L'equazione cartesiana è $y = x$ con $|x| \le 1$ e quindi $|x| \le 1$.

- Tentare di rappresentare le seguenti curve (anche con l'aiuto di un elaboratore)

$$\rho = \cos\theta, \quad \rho = \cos(2\theta), \quad \rho = \cos(3\theta)$$

  Che cosa si nota?

- Che cosa rappresenta nel piano un'equazione del tipo $f(x) = 0$?

- Provare che la curva
$$\mathcal{C}\colon x^4 + 3x^2y^2 - 4y^4 = 0$$
è costituita da 4 rette per l'origine (2 reali e 2 complesse coniugate).

- Dire che cosa rappresenta $f(x, y) = 0$ dove
$$f(x, y) = \sin \frac{x + y}{x}.$$
La curva rappresentata è algebrica?

3.4.c 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:
$$\begin{array}{lr} a^2 + b^2 + c^2 - d > 0 & \text{sfera ordinaria,} \\ a^2 + b^2 + c^2 - d = 0 & \text{sfera di raggio nullo,} \\ a^2 + b^2 + c^2 - d < 0 & \text{sfera immaginaria.} \end{array}$$

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.

3.4.d Coni e cilindri

Sia P un punto dello spazio ed α un piano. Proiettare P su α *da un fissato punto* V vuol dire considerare il punto $P'=VP\cap\alpha$; proiettare P su α *secondo una direzione data* $\vec{w}$ vuol dire considerare il punto $P'=s\cap\alpha$, dove s è la retta per P parallela a $\vec{w}$.

Se P descrive una curva $\mathcal{C}$, il punto P' descrive una curva $\mathcal{C}'\subset\alpha$, che è la *proiezione* di $\mathcal{C}$.

Si chiama *cono*, di *vertice* V e di *direttrice* una curva $\mathcal{C}$, la superficie $\mathcal{K}$ luogo delle rette (dette *generatrici* di $\mathcal{K}$) uscenti da V e che si appoggiano a $\mathcal{C}$. Si osservi che se $\mathcal{C}$

è una circonferenza, allora il cono di vertice V e di direttrice $\mathcal{C}$ è proprio il cono a due falde della geometria elementare.

Naturalmente per la curva $\mathcal{C}'$, proiezione di $\mathcal{C}$ su α da V, si ha $\mathcal{C}' = \mathcal{K} \cap \alpha$. Quindi, $\mathcal{C}$ e $\mathcal{C}'$ sono entrambe direttrici del cono, che proietta da V sia $\mathcal{C}$ sia $\mathcal{C}'$.

Si chiama *cilindro* la superficie Γ luogo delle rette (dette *generatrici* di Γ) incidenti una curva $\mathcal{C}$ ed aventi la stessa direzione individuata da un vettore $\vec{w}$.

Naturalmente per la curva $\mathcal{C}'$, proiezione di $\mathcal{C}$ su α parallelamente a $\vec{w}$, si ha $\mathcal{C}' = \Gamma \cap \alpha$.

I coni ed i cilindri sono esempi di superfici *rigate*, cioè superfici costituite da ∞^1 rette.

Troviamo ora equazioni parametriche di un cono e di un cilindro. Sia

$$\mathcal{C}\colon \quad x = x(u), \quad y = y(u), \quad z = z(u).$$

Se $V(x_0, y_0, z_0)$ e $\vec{w}(l, m, n)$ allora, posto $P(x, y, z) \in \Sigma$ e $Q \in \mathcal{C}$ si ha

$$\mathcal{K}\colon \begin{cases} x = x_0 + v(x(u) - x_0) \\ y = y_0 + v(y(u) - y_0) \\ z = z_0 + v(z(u) - z_0) \end{cases} \qquad \Gamma\colon \begin{cases} x = x(u) + lv \\ y = y(u) + mv \\ z = z(u) + nv \end{cases}$$

basta considerare, nel caso del cono, $\vec{PV} = v\vec{VQ}$, nel caso del cilindro $\vec{PQ} = v\vec{w}$.

Esempi ed esercizi.

- Riconoscere che la superficie seguente è costituita da rette

$$x = u + u^2 v, \quad y = (u^2 + 1)v, \quad z = \frac{1}{u} + v.$$

- Sia data la sfera

$$\Sigma\colon x^2 + y^2 + z^2 - 3x - y + 2 = 0.$$

  1. Determinare il cono $\mathcal{K}$ con vertice nell'origine e circoscritto a Σ.
  2. Scrivere l'equazione del piano contenente la circonferenza $\mathcal{C} = \mathcal{K} \cap \Sigma$.
  3. Trovare il centro ed il raggio di $\mathcal{C}$.

  (L'equazione del cono richiesto è $x^2 - 7y^2 - 8z^2 + 6xy = 0$.)

- Proiettare sul piano $x + y = 1$ la curva

$$\mathcal{C}\colon x = t, \quad y = t^2, \quad z = t^3.$$

- Si provi che la superficie

$$\Gamma\colon x^2 + y^2 - x = 0$$

  è un cilindro, con generatrici parallele all'asse z e con direttrice la curva

$$\mathcal{C}\colon \quad z = 0, \quad x^2 + y^2 - x = 0.$$

3.4.e Superficie di rotazione

Si chiama *superficie di rotazione* la superficie generata dalla rotazione di una curva $\mathcal{C}$ (piana o sghemba) intorno ad una retta a, che si chiama *asse* della superficie.

L'asse a può essere dato mediante un suo punto $A(x_0, y_0, z_0)$ e i suoi parametri direttori (l, m, n), la curva $\mathcal{C}$ mediante equazioni parametriche

$$\mathcal{C}\colon \quad x = x(u), \quad y = y(u), \quad z = z(u).$$

Il generico punto $P \in \mathcal{C}$, quando $\mathcal{C}$ ruota, descrive una circonferenza, detta *parallelo*,

$$\mathcal{P}_u = \tau_u \cap S_u,$$

dove τ_u è il piano per P e perpendicolare ad a ed S_u la sfera di centro A e raggio $\|\vec{AP}\|$:

$$\begin{aligned}
&\tau_u\colon\ l(x - x(u)) + m(y - y(u)) + n(z - z(u)) = 0,\\
&S_u\colon\ (x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2 = (x(u) - x_0)^2 + (y(u) - y_0)^2 + (z(u) - z_0)^2.
\end{aligned}$$

Se a coincide con l'asse z, le precedenti equazioni si semplificano notevolmente perché $(l, m, n) \sim (0, 0, 1)$ e si può prendere $A(0, 0, 0)$.

Naturalmente la superficie di rotazione Σ sarà l'unione di tutti i paralleli, al variare di $P \in \mathcal{C}$. In simboli

$$\Sigma = \bigcup_u \mathcal{P}_u.$$

Ne segue che l'equazione cartesiana di Σ si ottiene eliminando il parametro u dalle equazioni di τ_u e $\mathcal{P}_u$.

Esempio. Determinare la superficie di rotazione ottenuta ruotando attorno all'asse z la retta $r\colon x = 1,\ y = 2z$.

Il parallelo descritto dal punto $P(1, 2t, t) \in r$ è

$$\begin{cases} z = t \\ x^2 + y^2 + z^2 = (1 - 0)^2 + (2t - 0)^2 + (t - 0)^2 \end{cases}$$

Ne segue che la superficie di rotazione è $x^2 + y^2 - 4z^2 = 1$.

3.4.f Retta tangente ad una curva

Sia $\mathcal{C}$ una curva dello spazio e P_0 un fissato punto di $\mathcal{C}$. Si chiama *retta tangente* in P_0 a $\mathcal{C}$ la retta r, se esiste, posizione limite della secante congiungente i punti P_0 e P di $\mathcal{C}$, al tendere di P a P_0.

Se

$$\mathcal{C}\colon \quad x = x(t), \quad y = y(t), \quad z = z(t) \tag{3.4.5}$$

allora r sarà la retta per $P_0 = P(t_0)$ di parametri direttori (x'_0, y'_0, z'_0). Quindi in P_0 esiste la retta tangente se le funzioni (3.4.5) sono derivabili e $\vec{P'(t_0)} = (x'_0, y'_0, z'_0) \neq \vec{0}$. In tal caso P_0 si dice *regolare* o *non singolare*.

Esempio. Trovare la retta tangente in $P_0(0,0,1)$ alla curva

$$x = \sin^2\phi, \quad y = \sin\phi\cos\phi, \quad z = \cos\phi.$$

Poiché

$$x' = 2\sin\phi\cos\phi, \quad y' = \cos^2\phi - \sin^2\phi, z' = -\sin\phi,$$

e P_0 si ottiene per $\phi = 0$, si ha $\vec{P'_0} = (0,1,0)$, quindi la retta tangente è $x = 0, z = 1$.

3.4.g Piano tangente ad una superficie

Sia Σ una superficie e $P_0 \in \Sigma$. Consideriamo l'insieme $\{\mathcal{C}\}$ delle curve passanti per P_0 ed appartenenti a Σ. Per ognuna di queste curve si consideri (se esiste) la retta tangente in P_0. Se le rette tangenti a tutte le curve per P_0 appartengono ad uno stesso piano, allora questo sarà chiamato *piano tangente* in P_0 a Σ.

Sia $\Sigma\colon f(x,y,z) = 0$ una superficie. Il punto $P_0 \in \Sigma$ è detto *regolare* o *semplice* se

$$\left(\vec{\mathrm{grad}} f\right)_0 = (\nabla f)_0 = (f^0_x, f^0_y, f^0_z) \neq (0,0,0),$$

dove f^0_x, f^0_y, f^0_z sono le derivate parziali di f calcolate in P_0. In tal caso esiste il piano tangente ed ha equazione

$$f^0_x(x - x_0) + f^0_y(y - y_0) + f^0_z(z - z_0) = 0.$$

È intuitivo che nel vertice di un cono non esiste il piano tangente. Se Σ è una sfera, il piano tangente in P_0 a Σ è il piano per P_0 ed ortogonale a $\vec{P_0C}$, dove C è il centro di Σ.

Esempi ed esercizi.

- Trovare il piano tangente in $P_0(0,0,1)$ alla sfera $\Sigma\colon x^2 + y^2 + z^2 = 1$. Si ha $f_x = 2x$, $f_y = 2y$, $f_z = 2z$, quindi il piano tangente è $z = 1$.
- Trovare il piano tangente in $P_0(0,0,1)$ al cilindro $\Gamma\colon x^2 + y^2 - x = 0$.

Se

$$\Sigma\colon \quad x = x(u,v), \quad y = y(u,v), \quad z = z(u,v),$$

si prova che l'equazione del piano tangente è

$$\begin{vmatrix} x - x_0 & y - y_0 & z - z_0 \\ x^0_u & y^0_u & z^0_u \\ x^0_v & y^0_v & z^0_v \end{vmatrix} = 0.$$

Si tratta del piano per

$$P_0(x_0, y_0, z_0) = (x(u_0, v_0), y(u_0, v_0), z(u_0, v_0))$$

contenente i vettori $\vec{P_u^0} = (x_u^0, y_u^0, z_u^0)$ e $\vec{P_v^0} = (x_v^0, y_v^0, z_v^0)$. Quindi il piano esiste ed è individuato se $\vec{P_u^0}$ e $\vec{P_v^0}$ sono *indipendenti.*

Il vettore $\vec{P_u^0}$ è tangente alla curva coordinata

$$\mathcal{C}_u\colon \quad x = x(u, v_0), \quad y = y(u, v_0), \quad z = z(u, v_0),$$

giacente su Σ, che si ottiene per $v = \text{cost} = v_0$. Analogamente $\vec{P_v^0}$ è il vettore tangente alla curva coordinata

$$\mathcal{C}_v\colon \quad x = x(u_0, v), \quad y = y(u_0, v), \quad z = z(u_0, v).$$

Nota. Se $\mathcal{C} = \Sigma_1 \cap \Sigma_2$ e $P_0 \in \mathcal{C}$, allora la retta tangente a $\mathcal{C}$ in P_0 è $r = \tau_1 \cap \tau_2$, dove τ_i è il piano tangente in P_0 a Σ_i. Quindi se $\mathcal{C}$ è piana e $\mathcal{C} = \Sigma \cap \alpha$, la retta tangente a $\mathcal{C}$ in P_0 appartiene al piano α.

Se $\mathcal{C}$ è una curva del piano xy di equazioni

$$\mathcal{C}\colon \quad f(x, y) = 0, \quad z = 0,$$

allora $f_z^0 = 0$ e la retta tangente in P_0 a $\mathcal{C}$ è

$$f_z^0(x - x_0) + f_y^0(y - y_0) = 0, \quad z = 0.$$

Dunque nel piano xy la retta tangente a $\mathcal{C}$ in P_0 esiste se P_0 è semplice, cioè $(f_x^0, f_y^0) \neq (0, 0)$.

Esempio. Abbiamo già visto che la curva

$$\mathcal{C}\colon \quad x = \sin^2\phi, \quad y = \sin\phi\cos\phi, \quad z = \cos\phi,$$

è intersezione delle superficie

$$\Sigma_1\colon x^2 + y^2 + z^2 = 1, \qquad \Sigma_2\colon x^2 + y^2 - x = 0.$$

Quindi la retta r, tangente a $\mathcal{C}$ in $P_0(0, 0, 1)$ è $r = \tau_1 \cap \tau_2$. Risulta, infatti,

$$\tau_1\colon z - 1 = 0, \qquad \tau_2\colon x = 0.$$

3.4.h Coordinate cilindriche

Sia α un piano ed r una retta perpendicolare ad α (detta *asse delle quote*). Posto $O = \alpha \cap r$, consideriamo nel piano α un riferimento polare (ρ, φ) e nella retta r un riferimento cartesiano.

Se P è un punto dello spazio, consideriamo P', la sua proiezione ortogonale su α, e P'', proiezione ortogonale di P su r. Denotiamo (ρ, φ) le coordinate polari di P' in α ed h la coordinata di P'' su r. I tre numeri (ρ, φ, h), associati a P, si chiamano *coordinate cilindriche* di P. Fuori dall'asse r, la corrispondenza è biunivoca. Le coordinate si chiamano *cilindriche* poiché per $\rho = c \in \mathbb{R}$ si ottiene un cilindro rotondo intorno all'asse r di raggio c.

Spesso ad un riferimento cilindrico si associa un riferimento cartesiano $RC(Oxyz)$ tale che r coincida con l'asse z, il semiasse positivo delle x con l'asse polare nel piano α. Allora

$$\begin{cases} x = \rho\cos\varphi & 0 \leq \varphi < 2\pi \\ y = \rho\sin\varphi & \rho \in \mathbb{R}^+ \\ z = h & h \in \mathbb{R} \end{cases}$$

3.4.i Coordinate sferiche

Fissiamo nello spazio un riferimento polare costituito da

- un punto O detto *polo*;
- una retta orientata r per O detta *asse polare*;
- un semipiano α di origine r detto *semipiano polare*;
- un'unità di misura per le lunghezze ed un verso positivo per le rotazioni intorno all'asse polare.

Poniamo

$\rho = \|\vec{OP}\|$ <u>raggio vettore</u>,

$\varphi = \widehat{\alpha\beta}$ <u>longitudine</u>, dove β è il piano per r e P, $0 \leq \varphi < 2\pi$,

$\theta = \widehat{\vec{OP}\vec{r}}$ <u>colatitudine</u>, $0 \leq \theta \leq \pi$ ($\psi = \pi/2 - \theta$ <u>latitudine</u>).

I tre numeri (ρ, φ, θ) sono detti *coordinate sferiche*. Al riferimento polare si può associare un riferimento $RC(Oxyz)$ tale che O coincida con il polo, z coincida con l'asse polare, il semiasse positivo delle x appartenga al semipiano polare e coincidano le unità di misura per i segmenti. Allora

$$\begin{cases} x = \rho\sin\theta\cos\varphi & \rho \in \mathbb{R}^+ \\ y = \rho\sin\theta\sin\varphi & 0 \leq \varphi < 2\pi \\ z = \rho\cos\theta & 0 \leq \theta \leq \pi \end{cases}$$

Le coordinate si dicono *sferiche* poiché, per $\rho = \text{cost}$, si ottengono sfere concentriche. Pertanto, per $\rho = R$, le equazioni precedenti sono equazioni parametriche della sfera di centro O e raggio R; le coordinate (φ, θ) sono *coordinate geografiche* sulla sfera.

3.4.j Cambiamenti di riferimento

Se $RC(O'x'y'z')$ è un altro riferimento cartesiano, allora

$$X' = AX + X'_0,$$

dove A è una matrice ortogonale (${}^tA = A^{-1}$) e X'_0 è il vettore colonna le cui componenti sono le coordinate di O in RC'.

Se $\det A = +1$, i due riferimenti sono *equiversi*; se $\det A = -1$, sono *contraversi.*

Se consideriamo la geometria piana, la matrice A avrà una delle forme seguenti

$$\begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}, \quad \begin{pmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{pmatrix}.$$

La prima matrice rappresenta una rotazione (intorno ad O), mentre la seconda un ribaltamento rispetto alla retta $y = \text{tg}(\varphi/2)\, x$.

Esercizio 3.1. *Verificare che la curva di equazione*

$$\mathcal{C} \colon x^4 - 4x^3 + 5x^2 - 2x + y = 0$$

è simmetrica rispetto alla retta $x = 1$.

PARTE II

ALGEBRA LINEARE

CAPITOLO 4

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.

Il concetto base è quello di spazio vettoriale, generalizzazione dello spazio dei vettori ordinari, incontrato al capitolo 2.

4.1 Spazi vettoriali

Questa sezione è dedicata alla definizione di spazio vettoriale.

4.1.a Definizione

Definizione 4.1. *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* $\vec{v}, \vec{w}, \vec{z} \in V$
   (a) $(\vec{v} + \vec{w}) + \vec{z} = \vec{v} + (\vec{w} + \vec{z})$
   (b) $\exists\, \vec{0}$ *tale che* $\vec{v} + \vec{0} = \vec{v}$
   (c) $\exists\, (-\vec{v})$ *tale che* $\vec{v} + (-\vec{v}) = \vec{0}$
   (d) $\vec{v} + \vec{w} = \vec{w} + \vec{v}$
2. $\lambda(\mu\vec{v}) = (\lambda\mu)\vec{v}$ *per ogni* $\vec{v} \in V,\ \lambda, \mu \in \mathbb{K}$.
3. $(\lambda + \mu)\vec{v} = \lambda\vec{v} + \mu\vec{v}$ *per ogni* $\vec{v} \in V,\ \lambda, \mu \in \mathbb{K}$.
4. $\lambda(\vec{v} + \vec{w}) = \lambda\vec{v} + \lambda\vec{w}$ *per ogni* $\vec{v}, \vec{w} \in V,\ \lambda \in \mathbb{K}$.

5. $1\vec{v} = \vec{v}$ *per ogni* $\vec{v} \in V$.

Gli elementi di V sono detti *vettori* e quelli di $\mathbb{K}$ *scalari*. Con abuso di notazione si scriverà V al posto di $(V, +, \cdot)$. L'elemento $\vec{0}_V \stackrel{\text{def}}{=} \vec{0}$, elemento neutro di $(V, +)$, si chiama *vettore nullo* e si vede facilmente che $0\vec{v} = \vec{0}$ per ogni $\vec{v} \in V$.

Esempi ed esercizi.

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

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

  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 $\mathbb{R} (= \mathbb{R}^1)$, $\mathbb{R}^2$, $\mathbb{R}^3$.

- L'insieme $\mathbb{K}^{m,n}$ delle matrici di tipo m, n a coefficienti in $\mathbb{K}$ è uno spazio vettoriale rispetto alle operazioni di somma tra matrici e di moltiplicazione di una matrice per uno scalare (vedi 1.2.b).

- 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.

- Sia X un insieme non vuoto e V uno spazio vettoriale su $\mathbb{K}$. L'insieme

$$\mathcal{M}(X, V) \stackrel{\text{def}}{=} \{f \colon X \to V\} \tag{4.1.1}$$

  ha una struttura di spazio vettoriale rispetto alle operazioni di seguito definite. Siano $f, g \in \mathcal{M}(X, V)$ e $\lambda \in \mathbb{K}$. Si definiscono le funzioni $f + g \in \mathcal{M}(X, V)$ e $\lambda f \in \mathcal{M}(X, V)$ in questo modo

$$(f + g)(x) \stackrel{\text{def}}{=} f(x) + g(x), \qquad (\lambda f)(x) \stackrel{\text{def}}{=} \lambda f(x),$$

  per ogni $x \in X$. Si verificano facilmente le proprietà richieste.

4.1.b Sottospazi vettoriali

Definizione 4.2. *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 4.1. *Sia* $(V, +, \cdot)$ *uno spazio vettoriale, e* $W \subset V$. *Allora si equivalgono*

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

Ne segue che *tutti i sottospazi di* V *contengono il vettore nullo* $\vec{0} \in V$. 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 $W \stackrel{\text{def}}{=} \{\vec{x} \in \mathbb{R}^n \mid x_1 = 0\}$. Dimostriamo che W è un sottospazio vettoriale di $\mathbb{R}^n$. Siano $\vec{x}, \vec{y} \in W$ e $\lambda \in \mathbb{R}$. Allora

$$\vec{x} + \vec{y} = (x_1 + y_1, \dots, x_n + y_n), \quad \lambda\vec{x} = (\lambda x_1, \dots, \lambda x_n),$$

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

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

- Sia $U \stackrel{\text{def}}{=} \{\vec{x} \in \mathbb{R}^3 \mid x_1^2 + x_2^2 = 0\}$; U è un sottospazio vettoriale di $\mathbb{R}^3$?
- Sia $K \stackrel{\text{def}}{=} \{A \in \mathbb{K}^{n,n} \mid \det A = 0\}$; K è un sottospazio vettoriale?

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*.

4.1.c Somma e somma diretta

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

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

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

$S + T$ 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.

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

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

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

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

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

$$S_1 + \cdots + S_h \stackrel{\text{def}}{=} \{\vec{v}_1 + \cdots + \vec{v}_h \mid v_i \in S_i, i = 1, \ldots, h\}$$

$$S_1 \oplus \cdots \oplus S_h \stackrel{\text{def}}{=} \{\vec{v} = \vec{v}_1 + \cdots + \vec{v}_h \mid v_i \in S_i, i = 1, \ldots, h, \text{in modo unico}\}$$

Si osservi che

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

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 $\vec{x}_0$ è un fissato vettore di V ed S un suo sottospazio vettoriale, poniamo

$$\vec{x}_0 + S \stackrel{\text{def}}{=} \{\, \vec{x} = \vec{x}_0 + \vec{s} \mid \vec{s} \in S \,\}.$$

Si osservi che, per $\vec{x}_0 \neq \vec{0}$, $\vec{x}_0 + S$ non è un sottospazio vettoriale di V se $\vec{x}_0 \notin S$; se $\vec{x}_o \in S$ allora $\vec{x}_0 + S = S$.

4.1.d Dipendenza ed indipendenza lineare

Sia V uno spazio vettoriale su $\mathbb{K}$ e $\{\vec{v}_i\}_{1 \leq i \leq n}$ un suo sottoinsieme finito. Si dice che un vettore $\vec{v} \in V$ è *combinazione lineare* di $\{\vec{v}_i\}$ se esistono $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$ tali che

$$\vec{v} = \lambda_1 \vec{v}_1 + \cdots + \lambda_n \vec{v}_n.$$

Sia I un insieme non vuoto, anche infinito, e $\{\vec{v}_i\}_{i \in I}$ un sottoinsieme di vettori di V. Si dice che $\vec{v} \in V$ è *combinazione lineare* di $\{\vec{v}_i\}$ se esiste un sottoinsieme *finito* $J \subset I$ ed un insieme $\{\lambda_j\}_{j \in J}$ di scalari di $\mathbb{K}$ tali che

$$\vec{v} = \sum_{j \in J} \lambda_j \vec{v}_j.$$

Se I è un insieme finito, si può scegliere $J = I$ e la definizione attuale si riduce alla precedente.

Definizione 4.3. *Nelle ipotesi precedenti, l'insieme $\{\vec{v}_i\}_{i \in I}$ si dice*

1. indipendente *se il vettore $\vec{0}$ è combinazione lineare di $\{\vec{v}_i\}_{i \in I}$ in modo* unico, *e quindi con coefficienti tutti nulli, ossia se per ogni sottoinsieme finito $J \subset I$ vale*

$$\sum_{j \in J} \lambda_j \vec{v}_j = \vec{0} \quad \Rightarrow \quad \lambda_j = 0 \quad \forall j \in J;$$

2. dipendente *se non è indipendente, ovvero se almeno uno dei vettori $\{\vec{v}_i\}_{i \in I}$ è combinazione lineare dei restanti.*

4.1.e Sottospazi generati

Sia $X \subset V$ un sottoinsieme non vuoto. Si dice *sottospazio vettoriale generato da X* l'insieme

$$\mathcal{L}(X) \stackrel{\text{def}}{=} \{\vec{x} \in V \mid \vec{x} \text{ è comb. lin. di elementi di } X\}.$$

Si può dimostrare che $\mathcal{L}(X)$ è il più piccolo sottospazio vettoriale (rispetto all'inclusione) contenente X. È altresì ovvio che, se S, T sono sottospazi di V, allora

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

Si noti che:

- se σ è una permutazione di $\{1, \dots, h\}$, allora
$$\mathcal{L}(\vec{u}_1, \dots, \vec{u}_h) = \mathcal{L}(\vec{u}_{\sigma(1)}, \dots, \vec{u}_{\sigma(h)});$$
- se $u \in \mathcal{L}(\vec{u}_1, \dots, \vec{u}_h)$ allora
$$\mathcal{L}(\vec{u}_1, \dots, \vec{u}_h, \vec{u}) = \mathcal{L}(\vec{u}_1, \dots, \vec{u}_h).$$

Quindi, procedendo per scarti successivi si perviene a

$$\mathcal{L}(\vec{u}_1, \dots, \vec{u}_h) = \mathcal{L}(\vec{u}_{r_1}, \dots, \vec{u}_{r_p}),$$

dove i p vettori $\vec{u}_{r_1}, \dots, \vec{u}_{r_p}$ sono indipendenti e $p \leq h$. Il numero p risulta così essere il massimo numero di vettori indipendenti in $\mathcal{L}(X)$. Trovare questo numero è vantaggioso perché $p \leq h$ ed ogni vettore di $\mathcal{L}(X)$ si esprime in modo *unico* come combinazione lineare di quei vettori.

Si può dimostrare che

Teorema 4.1. *Sia $X = \{\vec{u}_i\}_{1 \leq i \leq n}$ e $\{\vec{v}_j\}_{1 \leq j \leq m} \subset \mathcal{L}(X)$ un sottoinsieme indipendente di $\mathcal{L}(X)$. Allora $m \leq n$.*

4.1.f Basi e dimensione

Sia V uno spazio vettoriale. Un sottoinsieme non vuoto $G \subset V$ è detto un insieme di *generatori* di V se $V = \mathcal{L}(G)$. Se esiste un insieme *finito* $F \subset V$ di generatori di V allora V si dice *finitamente generato*.

Un sottoinsieme $\mathcal{B} \subset V$ si dice *base* di V se $\mathcal{B}$ è un insieme di *generatori indipendenti*.

Si può dimostrare che ogni spazio vettoriale ammette una base [8]. Più precisamente:

Teorema 4.2. *Sia V uno spazio vettoriale, $G \subset V$ un insieme di generatori e $I \subset G$ un insieme indipendente, anche vuoto. Allora esiste una base $\mathcal{B}$ tale che*

$$I \subset \mathcal{B} \subset G.$$

Se $I \neq \varnothing$ il teorema è detto *teorema del completamento di una base.*

Teorema 4.3. *Sia V uno spazio vettoriale. Allora tutte le sue basi hanno la stessa cardinalità.*

Definizione 4.4. *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 $\dim\{\vec{0}\} = 0$.

In particolare, uno spazio vettoriale ha

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

Si osservi 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.

Nel seguito considereremo per lo più spazi vettoriali di dimensione finita. In tal caso possiamo dimostrare facilmente il teorema 4.3. Infatti, se $\mathcal{B}$ e $\mathcal{B}'$ sono due basi di V di n ed n' elementi, rispettivamente, allora il teorema 4.1 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 $n = n' = \dim V$.

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.

Se $\mathcal{B} = \{\vec{e}_i\}_{1 \leq i \leq n}$, allora per $\vec{x} \in V$

$$\vec{x} = x_1\vec{e}_1 + \cdots + x_n\vec{e}_n,$$

con $(x_1, \ldots, x_n)$ n-pla univocamente individuata. Gli scalari $x_1, \ldots, x_n$ sono chiamati *coordinate* di $\vec{x}$ rispetto alla base $\mathcal{B}$.

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

  $$\vec{e}_1 \stackrel{\text{def}}{=} (1, 0, \ldots, 0), \ldots, \vec{e}_n \stackrel{\text{def}}{=} (0, \ldots, 0, 1).$$

  Si osservi che $\mathbb{K}^0 \stackrel{\text{def}}{=} \{\vec{0}\}$.

- 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 $\vec{e}_1 \stackrel{\text{def}}{=} 1$).

  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 $\vec{e}_1 \stackrel{\text{def}}{=} 1$ e $\vec{e^*}_1 \stackrel{\text{def}}{=} i$. Ricordando l'identificazione $1 \mapsto (1,0)$, $i \mapsto (0,1)$ si ottiene per ogni $z \in \mathbb{C}$
  $$z = (a,b) \mapsto a\vec{e}_1 + b\vec{e^*}_1$$
  in modo unico.

- Analogamente $\dim_{\mathbb{C}} \mathbb{C}^n = n$, $\dim_{\mathbb{R}} \mathbb{C}^n = 2n$. Se $\mathcal{B} = \{\vec{e}_1, \dots, \vec{e}_n\}$ è una base di $\mathbb{C}^n$ su $\mathbb{C}$, allora $\mathcal{B}' = \{\vec{e}_1, \dots, \vec{e}_n, \vec{e}^*_1, \dots, \vec{e}^*_n\}$ è una base di $\mathbb{C}^n$ su $\mathbb{R}$, con $\vec{e}^*_h \stackrel{\text{def}}{=} i\vec{e}_h$, $h = 1, \dots, n$.

- Si ha $\dim_{\mathbb{K}} \mathbb{K}^{m,n} = m \cdot n$. Una base è data dall'insieme $\mathcal{C} \stackrel{\text{def}}{=} \{E_{ij} \mid 1 \le i \le m, 1 \le j \le n\}$ 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 $\{t^h\}_{0 \le h \le n}$.

- 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 $\mathbb{R}[t]$.

4.1.g Relazione di Grassmann

Teorema 4.4. *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 $\{\vec{u}_1, \dots, \vec{u}_h\}$ una base di $S \cap T$. Essa si può completare ad una base $\{\vec{u}_1, \dots, \vec{u}_h, \vec{v}_1, \dots, \vec{v}_p\}$ di S e ad una base $\{\vec{u}_1, \dots, \vec{u}_h, \vec{w}_1, \dots, \vec{w}_q\}$ di T. Se si dimostra che
$$\mathcal{B} \stackrel{\text{def}}{=} \{\vec{u}_1, \dots, \vec{u}_h, \vec{v}_1, \dots, \vec{v}_p, \vec{w}_1, \dots, \vec{w}_q\}$$
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

$$a_1\vec{u}_1 + \cdots + a_h\vec{u}_h + b_1\vec{v}_1 + \cdots + b_p\vec{v}_p + c_1\vec{w}_1 + \cdots + c_q\vec{w}_q = \vec{0} \tag{4.1.2}$$

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

$$a_1\vec{u}_1 + \cdots + a_h\vec{u}_h + b_1\vec{v}_1 + \cdots + b_p\vec{v}_p = -c_1\vec{w}_1 - \cdots - c_q\vec{w}_q$$

si vede che $\vec{w} \stackrel{\text{def}}{=} -c_1\vec{w}_1 - \cdots - c_q\vec{w}_q \in T \cap S$, quindi

$$\vec{w} = \lambda_1\vec{u}_1 + \cdots + \lambda_h\vec{u}_h,$$

per opportuni λ_i. Segue

$$\lambda_1\vec{u}_1 + \cdots + \lambda_h\vec{u}_h + c_1\vec{w}_1 + \cdots + c_q\vec{w}_q = \vec{0},$$

che implica $\lambda_i = c_j = 0$ poiché $\{\vec{u}_1, \ldots, \vec{u}_h, \vec{w}_1, \ldots, \vec{w}_q\}$ è una base. Quindi la (4.1.2) diventa

$$a_1\vec{u}_1 + \cdots + a_h\vec{u}_h + b_1\vec{v}_1 + \cdots + b_p\vec{v}_p = \vec{0},$$

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

Nel caso di somma diretta, poiché $S \cap T = \{\vec{0}\}$, 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 $S + T = S \oplus T$.

4.1.h Rango di un insieme di vettori

Sia V uno spazio vettoriale. Si definisce *rango* di un insieme $X \subset V$ di vettori il massimo numero di vettori indipendenti che è possibile estrarre da X. Ovviamente, se $\vec{v}_1, \ldots, \vec{v}_h \in V$ e $W \stackrel{\text{def}}{=} \mathcal{L}(\vec{v}_1, \ldots, \vec{v}_h)$, allora

$$\dim W = \operatorname{rg}\{\vec{v}_1, \ldots, \vec{v}_h\}.$$

Sia $A = (a_{ij}) \in \mathbb{K}^{m,n}$. Indichiamo con

$$\vec{r}_1 = (a_{11}, \ldots, a_{1n}), \ldots, \vec{r}_m = (a_{m1}, \ldots, a_{mn})$$

le righe di A e con

$$\vec{c}_1 = \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix}, \ldots, \vec{c}_n = \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix}$$

le colonne di A.

Si può provare che

$$\mathrm{rg}\{\vec{c_1}, \dots, \vec{c_n}\} \leq \mathrm{rg}\{\vec{r_1}, \dots, \vec{r_m}\}, \qquad \mathrm{rg}\{\vec{r_1}, \dots, \vec{r_m}\} \leq \mathrm{rg}\{\vec{c_1}, \dots, \vec{c_n}\},$$

da cui

$$\mathrm{rg}\{\vec{r_1}, \dots, \vec{r_m}\} = \mathrm{rg}\{\vec{c_1}, \dots, \vec{c_n}\} = \mathrm{rg}\, A,$$

dove $\mathrm{rg}\, A$ è il rango della matrice A precedentemente incontrato in 1.3.d. Quindi, se $X = \{\vec{v}_1, \dots, \vec{v}_m\}$ è un insieme di vettori tale che $\vec{v}_i \in V$ e $\dim V = n$, allora

$$\mathrm{rg}(X) = \mathrm{rg}(A),$$

dove A è la matrice le cui colonne (o righe) sono costituite dalle coordinate di $\vec{v}_i$ rispetto ad una base di V.

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

$$\vec{u}_1 = (1,1,1,1), \quad \vec{u}_2 = (1,-1,1,-1), \quad \vec{u}_3 = (1,3,1,3),$$

e W il sottospazio generato da

$$\vec{w}_1 = (1,2,0,2), \quad \vec{w}_2 = (1,2,1,2), \quad \vec{w}_3 = (3,1,3,1).$$

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

4.2 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.

4.2.a 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

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

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

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

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

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

Esempio. L'applicazione $f\colon V \to W$ tale che $f(\vec{x}) = \vec{y}_0$ per ogni $\vec{x} \in V$ è detta *applicazione costante*, e

$$f(V) = \{\vec{y}_0\} \quad \Rightarrow \quad f^{-1}(\vec{y}_0) = V, \quad f^{-1}(\vec{y}) = \varnothing \quad \forall \vec{y} \neq \vec{y}_0.$$

Sia $f : V \to W$ un'applicazione. Essa è iniettiva se

$$\vec{x} \neq \vec{x}' \Rightarrow f(\vec{x}) \neq f(\vec{x}')$$

oppure, che è lo stesso,

$$f(\vec{x}) = f(\vec{x}') \Rightarrow \vec{x} = \vec{x}'.$$

L'applicazione f è suriettiva se $f(V) = W$.

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

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

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

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

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

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

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

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

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

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

o, in forma matriciale,

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

In tal caso,
$$f(\vec{x}+\vec{x}') = f(\vec{x}) + f(\vec{x}'), \quad f(\lambda\vec{x}) = \lambda f(\vec{x}).$$
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.b Applicazioni lineari

Definizione 4.5. *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*

$$\begin{gathered} f(\vec{x}+\vec{x'}) = f(\vec{x}) + f(\vec{x'}) \quad \forall \vec{x}, \vec{x'} \in V \\ f(\lambda\vec{x}) = \lambda f(\vec{x}) \quad \forall \vec{x} \in V, \lambda \in \mathbb{K}, \end{gathered}$$

o, equivalentemente,

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

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

Esempi.

- $0\colon V \to W$, $\vec{v} \mapsto \vec{0}$, 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 $\vec{x} \mapsto \vec{x} + \vec{a}$, traslazione. Allora $T_{\vec{a}}$ è lineare se e solo se $\vec{a} = \vec{0}$.
- Si prova facilmente che, se $f\colon V \to W$ e $g\colon W \to Z$ sono due applicazioni lineari, allora
$$g \circ f\colon V \to Z, \vec{x} \mapsto g \circ f(\vec{x}) \stackrel{\text{def}}{=} g(f(\vec{x}))$$
è lineare.

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

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

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

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

Rispetto a queste operazioni, l'insieme

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

è uno spazio vettoriale, sottospazio di $\mathcal{M}(V, W)$ (4.1.1). 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 4.3. *Siano V, W spazi vettoriali, con V finitamente generato. Sia $\mathcal{B} = \{\vec{v}_1, \dots, \vec{v}_n\}$ una base di V, e siano $\vec{w}_1, \dots, \vec{w}_n \in W$. Allora esiste un'unica applicazione lineare $f\colon V \to W$ tale che $f(\vec{v}_i) = \vec{w}_i$ per $i = 1, \dots, n$.*

Dimostrazione. Infatti, per $\vec{x} = x_1\vec{v}_1 + \cdots + x_n\vec{v}_n$ si ponga

$$f(\vec{x}) \stackrel{\text{def}}{=} x_1\vec{w}_1 + \cdots + x_n\vec{w}_n, \quad \Rightarrow f(\vec{v}_i) = \vec{w}_i.$$

È immediato verificare che f è lineare. Inoltre f è unica. Infatti, se $g\colon V \to W$ è tale che $g(\vec{v}_i) = \vec{w}_i$, allora per ogni $\vec{x}$

$$f(\vec{x}) - g(\vec{x}) = x_1 f(\vec{v}_1) + \cdots + x_n f(\vec{v}_n) - x_1 g(\vec{v}_1) - \cdots - x_n g(\vec{v}_n) = \vec{0}.$$

QED

Osservazione. In generale per conoscere un'applicazione $f\colon V \to W$ bisogna conoscere $f(\vec{x})$ per ogni $\vec{x} \in V$. Il vantaggio delle applicazioni lineari è che per conoscere f basta conoscere solo $f(\vec{v}_i)$ dove $\{\vec{v}_i\}$ 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 4.4. *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 $\vec{y}_1, \vec{y}_2 \in f(S)$, allora $\vec{y}_1 = f(\vec{x}_1)$, $\vec{y}_2 = f(\vec{x}_2)$ con $\vec{x}_1, \vec{x}_2 \in S$. Per la linearità

$$\vec{y}_1 + \vec{y}_2 = f(\vec{x}_1) + f(\vec{x}_2) = f(\vec{x}_1 + \vec{x}_2) \in f(S),$$

poiché $\vec{x}_1 + \vec{x}_2 \in S$. Analogamente $\lambda\vec{y}_1 = f(\lambda\vec{x}_1) \in f(S)$ per $\lambda \in \mathbb{K}$.

L'altra implicazione è lasciata al lettore. QED

Casi particolarmente importanti sono

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

che, quindi, risultano sottospazi vettoriali.

Esercizio. Si verifichi direttamente, usando la proposizione 4.1, 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 $\dim \operatorname{Im} f \leq \dim W$.

Proposizione 4.5. *Sia $f\colon V \to W$ un'applicazione lineare. Allora*

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

Dimostrazione. Infatti,

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

QED

Quindi, in generale,

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

e dunque

$$f \text{ è iniettiva} \quad \Leftrightarrow \quad \operatorname{Ker} f = \{\vec{0}_V\} \quad \Leftrightarrow \quad f^{-1}(f(\vec{x})) = \vec{x},$$
$$f \text{ è suriettiva} \quad \Leftrightarrow \quad \operatorname{Im} f = W \quad \Leftrightarrow \quad f(f^{-1}(\vec{y})) = \vec{y}.$$

La seguente proprietà ha dimostrazione immediata.

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

$$U = \mathcal{L}(\vec{v}_1, \dots, \vec{v}_h) \quad \Rightarrow \quad f(U) = \mathcal{L}(f(\vec{v}_1), \dots, f(\vec{v}_h)),$$

da cui, se $\mathcal{B} = \{\vec{v}_1, \dots, \vec{v}_n\}$ è una base di V, allora $\operatorname{Im} f$ *è finitamente generato e*

$$\operatorname{Im} f = \mathcal{L}(f(\vec{v}_1), \dots, f(\vec{v}_n)).$$

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

Si osservi che se $\vec{v}_1, ..., \vec{v}_h$ sono indipendenti, non è detto che $f(\vec{v}_1), ..., f(\vec{v}_h)$ siano indipendenti, perciò $\dim f(U) \leq \dim U$.

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

$$\begin{aligned} \lambda_1 f(\vec{v}_1) + \cdots + \lambda_n f(\vec{v}_n) = \vec{0} \quad &\Leftrightarrow \quad f(\lambda_1 \vec{v}_1 + \cdots + \lambda_n \vec{v}_n) = \vec{0} \\ &\Leftrightarrow \quad \lambda_1 \vec{v}_1 + \cdots + \lambda_n \vec{v}_n = \vec{0} \quad \Leftrightarrow \quad \lambda_i = 0. \end{aligned}$$

Teorema 4.5 (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 $\{\vec{u}_1, \dots, \vec{u}_p\}$ una base di $\operatorname{Ker} f$ e $\vec{w}_1, \dots, \vec{w}_q$ una base di $\operatorname{Im} f$. Indicati con $\vec{v}_i \in V$ i vettori tali che $f(\vec{v}_i) = \vec{w}_i$, basta far vedere che $\{\vec{u}_1, \dots, \vec{u}_p, \vec{v}_1, \dots, \vec{v}_q\}$ costituiscono una base di V.

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

$$f(\vec{x}) = \lambda_1 \vec{w}_1 + \cdots + \lambda_q \vec{w}_q = \lambda_1 f(\vec{v}_1) + \cdots + \lambda_q f(\vec{v}_q) = f(\lambda_1 \vec{v}_1 + \cdots + \lambda_q \vec{v}_q),$$

da cui $\vec{x} - \lambda_1 \vec{v}_1 - \cdots - \lambda_q \vec{v}_q \in \operatorname{Ker} f = \mathcal{L}(\vec{u}_1, \dots, \vec{u}_p)$, quindi

$$\vec{x} - \lambda_1 \vec{v}_1 - \cdots - \lambda_q \vec{v}_q = \mu_1 \vec{u}_1 + \cdots + \mu_p \vec{u}_p.$$

Essi sono *indipendenti.* Infatti se

$$a_1 \vec{u}_1 + \cdots + a_p \vec{u}_p + b_1 \vec{v}_1 + \cdots + b_p \vec{v}_p = \vec{0}_V, \tag{4.2.3}$$

applicando la f ad entrambi i membri si ha

$$a_1 f(\vec{u}_1) + \cdots + a_p f(\vec{u}_p) + b_1 f(\vec{v}_1) + \cdots + b_q f(\vec{v}_q) = \vec{0}_W,$$

dunque

$$b_1 \vec{w}_1 + \cdots + b_q \vec{w}_q = \vec{0}_W \quad \Rightarrow \quad b_1 = \cdots = b_q = 0$$

per l'indipendenza di $\{w_j\}$, dunque l'equazione (4.2.3) diventa

$$a_1 \vec{u}_1 + \cdots + a_p \vec{u}_p = \vec{0}_V,$$

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

Chiamiamo

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

4.2.c Isomorfismi

Definizione 4.6. *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 4.7. *Due spazi vettoriali V e W si dicono* isomorfi*, in simboli $V \cong W$, se esiste un isomorfismo $f\colon V \to W$.*

Teorema 4.6. *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 $\operatorname{Ker} f = \{\vec{0}\}$ 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 $\mathcal{B} = \{\vec{e}_1, \dots, \vec{e}_n\}$ una base di V e $\mathcal{B}' = \{\vec{e'}_1, \dots, \vec{e'}_n\}$ una base di W. Definiamo un'applicazione lineare $f\colon V \to W$ ponendo $f(\vec{e}_i) \stackrel{\text{def}}{=} \vec{e'}_i$. La proposizione 4.3 assicura che f è univocamente definita. Per la linearità,

$$\vec{x} = x_1\vec{e}_1 + \dots + x_n\vec{e}_n \quad \Rightarrow \quad f(\vec{x}) = x_1\vec{e'}_1 + \dots + x_n\vec{e'}_n.$$

Se $\vec{x} \neq \vec{x'}$, allora $f(\vec{x}) \neq f(\vec{x'})$, dunque f è iniettiva.

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

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 $\mathcal{B} = \{\vec{v}_i\}_{1 \leq i \leq n}$ una base di V. Allora esiste un'unica applicazione lineare

$$c_{\mathcal{B}}\colon V \to \mathbb{K}^n, \quad \vec{v}_i \mapsto \vec{e}_i, \tag{4.2.4}$$

dove $\mathcal{C} = \{\vec{e}_i\}_{1 \leq i \leq n}$ 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 $\vec{v} \in V$ 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

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

è tale che $D(t^h) = ht^{h-1}$ 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 $f : \mathbb{R}^3 \to \mathbb{R}^3$, $(x, y, z) \mapsto (z, 0, 0)$. 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

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

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

4.2.d Matrici ed applicazioni lineari

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

$$f(\vec{e}_i) = a_{1i}\vec{e'}_1 + \cdots + a_{mi}\vec{e'}_m. \tag{4.2.5}$$

Viene così introdotta una matrice $A = (a_{ij})$ di tipo $m \times n$ dove la colonna i-esima è costituita dalle coordinate di $f(\vec{e}_i)$ 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

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

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,

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

dove X è il vettore delle componenti di $\vec{v} \in V$ rispetto a $\mathcal{B}$. Ma, usando (4.2.5), risulta

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

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

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.

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

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

è lineare e per le matrici associate vale

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

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 seguente notevole lemma di frequente utilizzo.

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

Nota. La matrice associata ad un'applicazione lineare dà tutte le informazioni sull'applicazione lineare stessa. Ad esempio,

$$\begin{gathered}\dim(\mathrm{Im}\, f) = \mathrm{rg}(f) = \mathrm{rg}\, A,\\ \dim(\mathrm{Ker}\, f) = \dim V - \mathrm{rg}\, A.\end{gathered}$$

Se A è una matrice invertibile, allora f è un isomorfismo.

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

Teorema 4.7. *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 $\mathcal{B} = \{\vec{e}_i\}_{1 \le i \le n}$ una base di V e $\mathcal{B}' = \{\vec{e}\,'_j\}_{1 \le j \le m}$ una base di W. Allora l'applicazione*

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

è un isomorfismo.

Dimostrazione. È facile verificare che

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

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

4.2.e Cambiamenti di base

Sia V uno spazio vettoriale su $\mathbb{K}$ di dimensione n e siano
$$\mathcal{B} = \{\vec{e}_1, \dots, \vec{e}_n\}, \qquad \mathcal{B}' = \{\vec{e'}_1, \dots, \vec{e'}_n\}$$
due basi distinte (anche solo per l'ordine degli elementi). Allora
$$\vec{e'}_k = \sum_{j=1}^{n} b_{jk}\vec{e}_j, \qquad \vec{e}_j = \sum_{i=1}^{n} b'_{ij}\vec{e'}_i.$$
Si vede subito che per le matrici $B = (b_{jk})$ e $B' = (c_{rj})$ vale
$$B' = B^{-1}.$$
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 $\vec{x} \in V$, e sia
$$\vec{x} = \sum_i x_i\vec{e}_i, \qquad \vec{x} = \sum_i x'_i\vec{e'}_i.$$
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)?
$$\vec{x} = \sum_i x'_i\vec{e'}_i = \sum_i x'_i \left(\sum_j b_{ji}\vec{e}_j\right) = \sum_j \left(\sum_i x'_i b_{ji}\right)\vec{e}_j$$
$$\Rightarrow \quad x_j = \sum_i b_{ji}x'_i \quad \Rightarrow \quad X = BX',$$

e, naturalmente $X' = B^{-1}X$.

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;*

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 $\mathcal{B} = \{\vec{v}_i\}_{1\leq i\leq n}$ una base di V e $\mathcal{C} = \{\vec{w}_j\}_{1\leq j\leq m}$ una base di W. Allora, se $\vec{y} = f(\vec{x})$ si ha

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

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

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

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

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

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

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

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

Resta così provato il seguente notevole teorema.

Teorema 4.8. *Matrici associate ad uno stesso endomorfismo rispetto a basi diverse sono simili.*

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

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

4.2.f Sistemi ed applicazioni lineari

Consideriamo il seguente sistema *omogeneo* di m equazioni in n incognite, in forma compatta

$$AX = O. \tag{4.2.6}$$

Sia $f_A \colon \mathbb{R}^n \to \mathbb{R}^m$ l'applicazione lineare tale che $\mathcal{M}^{\mathcal{C}}_{\mathcal{C}}(f_A) = A$, cioè

$$f_A(X) = AX.$$

Si osservi che

$$\vec{x} = (x_1, \dots, x_n) \text{ è soluzione di (4.2.6)} \quad \Leftrightarrow \quad \vec{x} \in \operatorname{Ker} f_A.$$

Quindi lo spazio delle soluzioni di (4.2.6) è lo spazio vettoriale $\operatorname{Ker} f_A$, e

$$\begin{aligned} \dim(\operatorname{Ker} f_A) = 0 \quad &\Rightarrow \quad \vec{0} \text{ è l'unica soluzione,} \\ \dim(\operatorname{Ker} f_A) > 0 \quad &\Rightarrow \quad \text{esistono autosoluzioni.} \end{aligned}$$

Più precisamente, per il teorema fondamentale,

$$\dim(\operatorname{Ker} f_A) = n - \operatorname{rg}(A).$$

Consideriamo ora un sistema lineare *generale* (omogeneo o no)

$$AX = B \qquad (A\vec{x} = \vec{b}). \tag{4.2.7}$$

Come sopra consideriamo l'applicazione f_A. Allora

$$\vec{x} \text{ soluzione di (4.2.7)} \quad \Leftrightarrow \quad f_A(\vec{x}) = \vec{b} \quad \Leftrightarrow \quad \vec{x} \in f_A^{-1}(\vec{b});$$

con altre parole

$$\text{il sistema (4.2.7) è compatibile} \quad \Leftrightarrow \quad f_A^{-1}(\vec{b}) \neq \varnothing \quad \Leftrightarrow \quad \vec{b} \in \operatorname{Im} f_A.$$

Ora

$$\operatorname{rg}(f_A) = \dim(\operatorname{Im} f_A) = \operatorname{rg}(A) = p.$$

Indichiamo con $\vec{c}_{i_1}, \dots, \vec{c}_{i_p}$ p colonne linearmente indipendenti di A (che generano $\operatorname{Im} f_A$). Allora

$$\begin{aligned} \text{(4.2.7) compatibile} \quad &\Leftrightarrow \quad \vec{b} \in \operatorname{Im} f_A \quad \Leftrightarrow \quad \vec{b} \in \mathcal{L}(\vec{c}_{i_1}, \dots, \vec{c}_{i_p}) \\ &\Leftrightarrow \quad \{\vec{b}, \vec{c}_{i_1}, \dots, \vec{c}_{i_p}\} \text{ dipendenti} \quad \Leftrightarrow \quad \operatorname{rg}\{\vec{b}, \vec{c}_{i_1}, \dots, \vec{c}_{i_p}\} = p. \end{aligned}$$

Ma $\operatorname{rg}\{\vec{b}, \vec{c}_{i_1}, \dots, \vec{c}_{i_p}\} = \operatorname{rg}(\tilde{A})$, dove $\tilde{A} = (A|B)$ è la matrice completa del sistema. Abbiamo così provato il seguente teorema.

Teorema 4.9 (Rouché–Capelli). *Il sistema $AX = B$ è compatibile se e solo se $\operatorname{rg}(A) = \operatorname{rg}(\tilde{A})$.*

Descriviamo ora lo spazio S delle soluzioni di (4.2.7). Se $\vec{x}_0$ è tale che $f(\vec{x}_0) = \vec{b}$, cioè una soluzione del sistema non omogeneo, e $\vec{x}$ è tale che $f(\vec{x}) = \vec{0}$, allora

$$f(\vec{x}_0 + \vec{x}) = \vec{b} \quad \Leftrightarrow \quad S = \vec{x}_0 + \operatorname{Ker} f, \tag{4.2.8}$$

quindi se $\vec{x}_0 \neq \vec{0}$ lo spazio S *non è uno spazio vettoriale*, ma lo spazio ottenuto traslando lo spazio vettoriale $\operatorname{Ker} f$ mediante il vettore $\vec{x}_0$, in simboli

$$S = t_{\vec{x}_0}(\operatorname{Ker} f).$$

Se U è un sottospazio vettoriale di V ed $\vec{a}$ un vettore fissato di V, lo spazio

$$S = t_{\vec{a}}(U)$$

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

$$\varphi\colon V \to W, \ \varphi = t_{\vec{a}} \circ f, \qquad \vec{a} \in W,$$

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

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

4.3 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.

4.3.a Definizioni

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

$$f\colon V \to V$$

un endomorfismo. Vogliamo vedere se esistono vettori $\vec{x}_0 \neq \vec{0}$ tali che

$$f(\vec{x}_0) = \lambda \vec{x}_0, \qquad \lambda \in \mathbb{K},$$

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

Definizione 4.8. *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> *$\vec{x}$ tale che*

$$f(\vec{x}) = \lambda \vec{x}.$$

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

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

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> $f(\vec{x})$ di una applicazione sia proporzionale alla <u>causa</u> $\vec{x}$.

In Fisica gli autovalori corrispondono agli assi principali di un ellissoide costruito tramite 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 $\vec{x}$ è un autovettore per f, allora $\mathcal{L}(\vec{x})$ è invariante per f, cioè

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

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

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

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 $\vec{0}$, perciò

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

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

Teorema 4.10. *Se $\lambda_1, \dots, \lambda_p$ sono autovalori di f* distinti *e se $\vec{x}_1, \dots, \vec{x}_p$ sono autovettori di f corrispondenti agli autovalori dati, allora i vettori $\vec{x}_1, \dots, \vec{x}_p$ sono linearmente indipendenti.*

La principale conseguenza del precedente risultato è che se $\lambda_i \neq \lambda_j$ allora

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

4.3.b Polinomio caratteristico

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

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

Nota 4.1. *Altri autori pongono come polinomio caratteristico*

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

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

Sviluppando il determinante 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:

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

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

---

[1]Altri autori usano la notazione V_λ.

Dimostrazione.

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

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

Teorema 4.12. *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

Ora, ricordando che matrici simili rappresentano lo stesso endomorfismo rispetto a basi diverse, possiamo definire il polinomio caratteristico $P_f(\lambda)$ di un endomorfismo f tra spazi vettoriali di dimensione finita come il polinomio caratteristico di una qualsiasi matrice che rappresenti l'endomorfismo.

Analogamente, poiché i coefficienti c_h di un polinomio caratteristico sono invarianti per similitudine, si può definire $\mathrm{tr}(f)$, $\det(f)$ come traccia e determinante di una qualsiasi matrice che rappresenti l'endomorfismo.

Teorema 4.13. *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 $P_f(\lambda) = 0$.*

DIMOSTRAZIONE.

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

che in forma matriciale, fissata una base, è rappresentato dal sistema omogeneo

$$(A - \lambda \operatorname{Id})X = O,$$

dove $f(X) = AX$. La conclusione si ottiene ricordando precedenti teoremi, tenendo conto che un autovettore è un vettore non nullo. $\boxed{QED}$

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

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

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

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

$$\Rightarrow \quad \begin{cases} -2x - y = 0 \\ x - 2y = 0 \end{cases} \quad \Rightarrow \quad V(2) = \{t(0,0,1) \mid t \in \mathbb{R}\} = \mathcal{L}(\vec{e}_3).$$

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

$$\begin{aligned} V(\lambda_1) &= \mathcal{L}(\vec{u}_1), & \vec{u}_1 &= (i,1,0) \\ V(\lambda_2) &= \mathcal{L}(\vec{u}_2), & \vec{u}_2 &= (-i,1,0) \\ V(\lambda_3) &= \mathcal{L}(\vec{u}_3), & \vec{u}_3 &= (0,0,1). \end{aligned}$$

Si osservi che, in questo caso,

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

I vettori $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ 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 $\det \tilde{A} = \det A$, $\operatorname{tr} \tilde{A} = \operatorname{tr} A$.

Nota. Per trovare gli autovalori di un endomorfismo $f: V \to V$ non bisogna operare con una matrice $\mathcal{M}^{\mathcal{C}}_{\mathcal{B}}(f)$, dove $\mathcal{C} \neq \mathcal{B}$. In tal caso, infatti, la matrice di Id_V in $f - \lambda \mathrm{Id}_V$ non sarebbe uguale alla matrice identità!

4.3.c Endomorfismi semplici

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

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

Se $\tilde{\mathcal{B}} = \{\vec{u}_1, \dots, \vec{u}_n\}$ è una tale base, allora $f(\vec{u}_i) = \lambda_i \vec{u}_i$ 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 $\mathcal{B} = \{\vec{e}_1, \dots, \vec{e}_n\}$ 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. Allora, indicata con B la matrice del cambuamento di base da $\mathcal{B}$ a $\tilde{\mathcal{B}}$ (base di autovettori), si ha

$$D = B^{-1}AB \Rightarrow A = BDB^{-1}.$$

Vedremo subito che non tutti gli endomorfismi sono semplici e quindi che non ogni matrice è diagonalizzabile.

Teorema 4.14 (Criterio di semplicità)**.** *L'endomorfismo $f: V \to V$ è semplice se e solo se*

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

2. *$m_i = s_i$ $\forall i = 1, 2, \dots, r \leq n$, dove m_i è la molteplicità algebrica (osservazione 1.1) ed s_i quella geometrica dell'autovalore λ_i.*

Si può dimostrare che, in generale,

$$1 \leq s_i \leq m_i,$$

quindi in particolare se $P_f(\lambda)$ ha n zeri tutti <u>distinti</u>, appartenenti a $\mathbb{K}$, allora $m_i = 1$ $\forall i$, quindi f è semplice poiché $s_i = m_i = 1$ per ogni i.

D'altra parte se $\lambda_1, \dots, \lambda_n$ sono distinti, allora autovettori corrispondenti $\vec{u}_1, \dots, \vec{u}_n$ sono indipendenti e quindi costituiscono una base.

Si noti che l'avere n autovalori distinti è condizione solo **sufficiente** ma non **necessaria**. Ad esempio, I è certamente diagonalizzabile (essendo diagonale) ma gli autovalori sono tutti coincidenti.

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' <u>non</u> 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 $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ tale che $BA = A'B$. Da qui segue $a = 0 = b$, ma in tal caso $\det B = 0$, ossia B non è invertibile.

Si osservi che A è diagonalizzabile, mentre A' non lo è.

Se $P_A(\lambda) = P_{A'}(\lambda)$ e le matrici sono entrambe diagonalizzabili, allora esse sono simili. Infatti

$$D = B^{-1}AB, \quad D = C^{-1}AC \Rightarrow A' = G^{-1}AG$$

dove $G = BC^{-1}$.

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 $W = \mathcal{L}(\vec{v}_1, \vec{v}_2)$ dove $\vec{v}_1 = (0, 3, 1)$ e $\vec{v}_2 = (1, 4, 1)$.

  1. Provare che $f \colon W \to \mathbb{R}^3$ definita da

  $$f(\vec{v}_1) = (-1, 5, 2), \qquad f(\vec{v}_2) = (0, 6, 2),$$

  è un endomorfismo.

2. Trovare gli autovalori di $f \colon W \to W$.

1. Bisogna dimostrare che $f(W) \subset W$; basta che $f(\vec{v}_1)$, $f(\vec{v}_2) \in \mathcal{L}(\vec{v}_1, \vec{v}_2)$, cioè esistono $a, b, c, d \in \mathbb{R}$ tali che

$$f(\vec{v}_1) = a\vec{v}_1 + b\vec{v}_2, \qquad f(\vec{v}_2) = c\vec{v}_1 + d\vec{v}_2.$$

   Segue facilmente che $a = 3$, $b = -1$, $c = 2$, $d = 0$.

2. Considerando $\mathcal{B} = \{\vec{v}_1, \vec{v}_2\}$ 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 5

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 $\mathbb{R}$.

5.1 Forme bilineari e forme quadratiche

In questa sezione sarà introdotto materiale preliminare alle strutture metriche.

5.1.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

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

è detto *duale* di V. Per il teorema 4.7 si ha $\dim V^* = n \cdot 1 = n$, quindi $V^* \cong V$.

Definizione 5.1. *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*

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

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

Una forma bilineare β è detta *simmetrica* se

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

è detta *antisimmetrica* o *alternante* se

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

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

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

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

Se $\mathcal{B} = \{\vec{e}_1, \ldots, \vec{e}_n\}$ è una base di V, posto

$$a_{ij} \stackrel{\text{def}}{=} \beta(\vec{e}_i, \vec{e}_j) \in \mathbb{R}, \qquad (i, j = 1, \ldots, n),$$

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

$$\mathcal{M}_{\mathcal{B}}(\beta) \stackrel{\text{def}}{=} (\beta(\vec{e}_i, \vec{e}_j)).$$

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 $\beta(\vec{x}, \vec{y})$ per ogni $\vec{x}, \vec{y} \in V$. Infatti, se

$$\vec{x} = \sum_i x_i \vec{e}_i, \qquad \vec{y} = \sum_j y_j \vec{e}_j,$$

allora

$$\beta(\vec{x}, \vec{y}) = \beta(\sum_i x_i \vec{e}_i, \sum_j y_j \vec{e}_j) = \sum_i \sum_j x_i y_j \beta(\vec{e}_i, \vec{e}_j) = \sum_{ij} a_{ij} x_i y_j = {}^t\!XAY,$$

dove ${}^t\!X = (x_1 \ldots x_n)$ 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 $\mathcal{B}' = \{\vec{e'}_1, \ldots, \vec{e'}_n\}$ è una nuova base allora $\beta(\vec{x}, \vec{y}) = {}^t\!X' A' Y'$. 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

$${}^t\!XAY = {}^t(BX')A(BY') = {}^t\!X'({}^t\!BAB)Y',$$

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 $0\vec{x} = \vec{0}$, si verifica facilmente che

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

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

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

Se

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

allora β si dice *non degenere*. Ovviamente, se $\mathcal{B} = \{\vec{e}_1, \dots, \vec{e}_n\}$, allora

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

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

$$\sum_{i=1}^{n} x_i \beta(\vec{e}_i, \vec{e}_j) = 0 \quad \Rightarrow \quad \sum_{i=1}^{n} a_{ij} x_i = 0,$$

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

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

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.

5.1.b Forme quadratiche

Definizione 5.2. *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*

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

In una base si ha l'espressione

$$Q(\vec{x}) = \beta(\vec{x}, \vec{x}) = \sum_{ij} a_{ij} x_i x_j = {}^tXAX,$$

quindi la matrice di β può essere considerata anche la matrice di Q.

Q è detta *quadratica* perchè

$$Q(\lambda\vec{x}) = \beta(\lambda\vec{x}, \lambda\vec{x}) = \lambda^2\beta(\vec{x}, \vec{x}) = \lambda^2 Q(\vec{x}).$$

Viceversa, data una forma quadratica Q si può costruire subito la forma bilineare simmetrica da cui proviene:

$$\beta(\vec{x}, \vec{y}) = \frac{1}{2}(Q(\vec{x} + \vec{y}) - Q(\vec{x}) - Q(\vec{y})).$$

Infatti:

$$Q(\vec{x} + \vec{y}) = \beta(\vec{x} + \vec{y}, \vec{x} + \vec{y}) = \beta(\vec{x}, \vec{x}) + \beta(\vec{x}, \vec{y}) + \beta(\vec{y}, \vec{x}) + \beta(\vec{y}, \vec{y}).$$

L'insieme $\{\vec{x} \in V \mid Q(\vec{x}) = 0\}$ 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 5.1. *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 $\mathcal{B}' = \{\vec{e'}_1, \dots, \vec{e'}_n\}$ rispetto alla quale $\beta(\vec{e'}_i, \vec{e'}_j) = 0$ per $i \neq j$.*

Quindi, rispetto alla base $\mathcal{B}'$, la matrice $A' = \mathcal{M}_{\mathcal{B}'}(\beta)$ è diagonale del tipo

$$A' = \begin{pmatrix} \alpha'_1 & 0 & \dots\dots\dots\dots\dots\dots & & & \\ 0 & \alpha'_2 & \dots\dots\dots\dots\dots\dots & & & \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & \dots & \alpha'_p & \dots & 0 \\ 0 & \dots\dots\dots\dots\dots & & & & 0 & \dots \\ 0 & \dots\dots\dots\dots\dots\dots & & & & & 0 \end{pmatrix},$$

con $\alpha'_i = \beta(\vec{e'}_i, \vec{e'}_i) \neq 0$ per $i = 1, \dots, p$ e $\beta(\vec{e'}_i, \vec{e'}_i) = 0$ se $i > p$. Quindi $\operatorname{rg}(\beta) = p$ e

$$Q(\vec{x}) = \alpha'_1 x_1^2 + \dots + \alpha'_p x_p^2.$$

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:

$$\begin{aligned} &\vec{E}_i \stackrel{\text{def}}{=} \frac{\vec{e'}_i}{\sqrt{\alpha_i'}} && \text{se } \alpha_i' > 0 \\ &\vec{E}_i \stackrel{\text{def}}{=} \frac{\vec{e'}_i}{\sqrt{-\alpha_i'}} && \text{se } \alpha_i' < 0 \\ &\vec{E}_i \stackrel{\text{def}}{=} \vec{e'}_i && \text{se } \alpha_i' = 0 \end{aligned}$$

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

$$Q(\vec{x}) = X_1^2 + \cdots + X_s^2 - X_{s+1}^2 - \cdots - X_p^2.$$

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

$$
\begin{array}{lcl}
\textit{semidefinita positiva} & \Leftrightarrow & Q(\vec{x}) \geq 0 \ \forall \vec{x} \in V \\
\textit{definita positiva} & \Leftrightarrow & Q(\vec{x}) > 0 \ \forall \vec{x} \in V \smallsetminus \vec{0} \\
\textit{semidefinita negativa} & \Leftrightarrow & Q(\vec{x}) \leq 0 \ \forall \vec{x} \in V \\
\textit{definita negativa} & \Leftrightarrow & Q(\vec{x}) < 0 \ \forall \vec{x} \in V \smallsetminus \vec{0} \\
\textit{indefinita} & \Leftrightarrow & \exists \vec{x}, \vec{y} \in V \ : Q(\vec{x}) > 0, \ Q(\vec{y}) < 0.
\end{array}
$$

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.

- $\beta\colon \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$, $\beta(\vec{x}, \vec{y}) = x_1y_1 + x_2y_2$. Si ha $Q(\vec{x}) = x_1^2 + x_2^2$, e

  $$\mathcal{M}_{\mathcal{C}}(\beta) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

  quindi Q è definita positiva.

- $\beta\colon \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$, $\beta(\vec{x}, \vec{y}) = x_1y_1 + x_2y_2$. Si ha $Q(\vec{x}) = x_1^2 + x_2^2$, dunque Q è semidefinita positiva, poiché $Q(\vec{v}) = 0$ con $\vec{v} = (0, 0, 1) \neq \vec{0}$.

- $\beta\colon \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$, $\beta(\vec{x}, \vec{y}) = x_1y_1 - x_2y_2$. Si ha $Q(\vec{x}) = x_1^2 - x_2^2$, dunque Q è indefinita.

5.2 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.

5.2.a Definizione

Definizione 5.3. *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 $\vec{u} \cdot \vec{v} = g(\vec{u}, \vec{v})$[1], allora le proprietà di g si traducono nelle seguenti:

1. distributività:
$$\vec{u} \cdot (\vec{v} + \vec{v}') = \vec{u} \cdot \vec{v} + \vec{u} \cdot \vec{v}', \qquad (\vec{u} + \vec{u}') \cdot \vec{v} = \vec{u} \cdot \vec{v} + \vec{u}' \cdot \vec{v};$$
2. omogeneità:
$$\vec{u} \cdot (\lambda\vec{v}) = \lambda(\vec{u} \cdot \vec{v}), \qquad (\lambda\vec{u}) \cdot \vec{v} = \lambda(\vec{u} \cdot \vec{v});$$
3. commutatività: $\vec{u} \cdot \vec{v} = \vec{v} \cdot \vec{u}$;
4. essere def. positiva: $\vec{u} \cdot \vec{u} > 0$ per ogni $\vec{u} \neq \vec{0}$, da cui $\vec{u} \cdot \vec{u} = \vec{0} \Leftrightarrow \vec{u} = \vec{0}$.

Definizione 5.4. *Si dice* spazio vettoriale euclideo *uno spazio vettoriale V sul campo $\mathbb{R}$, di dimensione finita, dotato di un prodotto scalare g.*

Nota 5.1. *Dalla proprietà 4 segue anche che*

$$\forall\, \vec{v} \quad \vec{u} \cdot \vec{v} = 0 \quad \Rightarrow \quad \vec{u} = \vec{0};$$

per dimostrare questo fatto, si prenda $\vec{v} = \vec{u}$.

Esempi ed esercizi.

- $V = \mathbf{V}_3$, $\vec{u} \cdot \vec{v} \overset{\text{def}}{=} \|\vec{u}\|\,\|\vec{v}\| \cos\widehat{uv}$ è un prodotto scalare.
- $V = \mathbb{R}^n$, $\vec{u} = (u_1, \dots, u_n)$, $\vec{v} = (v_1, \dots, v_n)$, allora è un prodotto scalare
$$\vec{u} \cdot \vec{v} \overset{\text{def}}{=} u_1v_1 + \cdots + u_nv_n$$
- $V = \mathbb{R}^2$, $g(\vec{u}, \vec{v}) \overset{\text{def}}{=} u_1v_1 + 2u_1v_2 + 5u_2v_2$ non è un prodotto scalare.
- $V = \mathbb{R}^2$, $g(\vec{u}, \vec{v}) \overset{\text{def}}{=} u_1v_1 - u_2v_2$ non è un prodotto scalare.
- $V = \mathbb{R}^{n,n}$, $g(A, B) \overset{\text{def}}{=} \sum_{i,j=1}^{n} a_{ij}b_{ij}$ è un prodotto scalare.

[1] Altri autori usano la notazione $g(\vec{u}, \vec{v}) = \langle \vec{u}, \vec{v} \rangle$

- $V = \mathbb{R}^{n,n}$, $g'(A,B) \stackrel{\text{def}}{=} \operatorname{tr}(A^t B)$ è un prodotto scalare. Infatti g' è bilineare e

$$g'(B,A) = \operatorname{tr}(B\,{}^tA) = \operatorname{tr}({}^t(B\,{}^tA)) = \operatorname{tr}(A\,{}^tB) = g'(A,B).$$

Inoltre, $g'(A,A) = \sum_{ij} a_{ij}^2 \geq 0$ se $A \neq O$. Dimostrare che, se $n = 2$, $g = g'$.

Il numero

$$\|\vec{u}\|_g \stackrel{\text{def}}{=} \sqrt{g(\vec{u},\vec{u})} \geq 0$$

si dice *norma*, o *lunghezza*, o *modulo* di $\vec{u}$ rispetto a g. Si ha

$$\|\lambda\vec{u}\|_g = \sqrt{g(\lambda\vec{u},\lambda\vec{u})} = |\lambda|\|\vec{u}\|_g.$$

In poche parole,

$$\|\vec{u}\|_g^2 = Q(\vec{u}),$$

dove Q è la forma quadratica associata a g. Inoltre (paragrafo 5.1.b)

$$g(\vec{u},\vec{v}) = \frac{1}{2}(\|\vec{u}+\vec{v}\|_g^2 - \|\vec{u}\|_g^2 - \|\vec{v}\|_g^2).$$

Si dice che due vettori $\vec{u},\vec{v} \in V$ sono *ortogonali* (rispetto a g) se $g(\vec{u},\vec{v}) = 0$, e si scrive $\vec{u} \perp_g \vec{v}$, o anche $\vec{u} \perp \vec{v}$. In tal caso, vale il seguente *teorema di Pitagora* per spazi vettoriali euclidei

$$\|\vec{u}+\vec{v}\|_g^2 = \|\vec{u}\|_g^2 + \|\vec{v}\|_g^2.$$

Se $\vec{u},\vec{v} \in V$, poniamo

$$\operatorname{dist}_g(\vec{u},\vec{v}) \stackrel{\text{def}}{=} \|\vec{u}-\vec{v}\|_g,$$

distanza di $\vec{u}$ da $\vec{v}$ (rispetto a g). Infatti, se $\vec{u} = \vec{OP}$ e $\vec{v} = \vec{OQ}$, allora $\operatorname{dist}_g(\vec{u},\vec{v}) = \operatorname{dist}_g(P,Q) = \|P-Q\|_g$.

Teorema 5.2 (Disuguaglianza di Schwarz)**.** *Se V è uno spazio vettoriale euclideo con prodotto scalare g e $\vec{u},\vec{v} \in V$, allora*

$$|g(\vec{u},\vec{v})| \leq \|\vec{u}\|_g\,\|\vec{v}\|_g,$$
$$|g(\vec{u},\vec{v})| = \|\vec{u}\|_g\,\|\vec{v}\|_g \quad \Leftrightarrow \quad \vec{u} \parallel \vec{v}.$$

Dimostrazione. Se $\vec{v} = \vec{0}$ o $\vec{u} = \vec{0}$, la disuguaglianza è banale. Se $\vec{u} \neq \vec{0}$ e $\vec{v} \neq \vec{0}$ si consideri $\vec{u} + \lambda\vec{v}$, per $\lambda \in \mathbb{R}$. Allora

$$\varphi(\lambda) = \|\vec{u}+\lambda\vec{v}\|_g^2 = \|\vec{u}\|_g^2 + 2\lambda g(\vec{u},\vec{v}) + \lambda^2\|\vec{v}\|_g^2 \geq 0.$$

Da proprietà delle disequazioni (o da semplici considerazioni geometriche sulle parabole) si deduce

$$g(\vec{u},\vec{v})^2 - \|\vec{u}\|_g^2\,\|\vec{v}\|_g^2 \leq 0,$$

da cui la conclusione.

Se $g(\vec{u},\vec{v})^2 - \|\vec{u}\|_g^2\,\|\vec{v}\|_g^2 = 0$, l'equazione $\varphi(\lambda) = 0$ ha due radici concidenti,λ_0, quindi

$$\varphi(\lambda_0) = \|\vec{u} + \lambda_0\vec{v}\|^2 = 0 \Rightarrow \vec{u} = k\vec{v}.$$

Viceversa

$$|g(\vec{y},\vec{v})| = |g(k\vec{v},\vec{v})| = |k|\|\vec{v}\|^2 = |k|\|\vec{v}\|\|\vec{v}\| = \|\vec{u}\|\|\vec{v}\|.$$

QED

Dalla precedente disuguaglianza si ha il seguente teorema.

Teorema 5.3 (Disuguaglianza di Minkowski o triangolare)**.** *Se V è uno spazio vettoriale euclideo con prodotto scalare g e $\vec{u},\vec{v} \in V$, allora*

$$\|\vec{u} + \vec{v}\|_g \leq \|\vec{u}\|_g + \|\vec{v}\|_g.$$

Dimostrazione. Infatti

$$\begin{aligned}\|\vec{u} + \vec{v}\|_g^2 = g(\vec{u}+\vec{v},\vec{u}+\vec{v}) = \|\vec{u}\|_g^2 + 2g(\vec{u},\vec{v}) + \|\vec{v}\|_g^2 &\leq \|\vec{u}\|_g^2 + 2|g(\vec{u},\vec{v})| + \|\vec{v}\|_g^2 \\ &\leq \|\vec{u}\|_g^2 + 2\|\vec{u}\|_g\,\|\vec{v}\|_g + \|\vec{v}\|_g^2 = (\|\vec{u}\|_g + \|\vec{v}\|_g)^2.\end{aligned}$$

QED

Come sopra,

$$\|\vec{u} + \vec{v}\|_g = \|\vec{u}\|_g + \|\vec{v}\|_g \quad \Leftrightarrow \quad \vec{u} = k\vec{v},\ k > 0.$$

Dalla disuguaglianza di Schwarz segue anche

$$\frac{g(\vec{u},\vec{v})}{\|\vec{u}\|_g\,\|\vec{v}\|_g} \leq \frac{|g(\vec{u},\vec{v})|}{\|\vec{u}\|_g\,\|\vec{v}\|_g} \leq 1,$$

quindi possiamo <u>definire</u> il *coseno dell'angolo* $\widehat{\vec{u}\vec{v}}$ (se $\vec{u},\vec{v} \neq \vec{0}$) così

$$\cos\widehat{\vec{u}\vec{v}} \overset{\text{def}}{=} \frac{g(\vec{u},\vec{v})}{\|\vec{u}\|_g\,\|\vec{v}\|_g}.$$

Naturalmente la definizione è stata suggerita dal caso dei vettori geometrici dove il prodotto scalare è quello standard (si veda il paragrafo 2.4.a).

5.2.b Ortonormalizzazione

Premettiamo il seguente lemma.

Lemma 5.1. *Se $\{\vec{u}_1,\ldots,\vec{u}_k\}$ sono vettori di uno spazio vettoriale euclideo tra loro ortogonali e non nulli, allora essi sono indipendenti.*

Dimostrazione. Sia $\lambda_1\vec{u}_1 + \cdots + \lambda_k\vec{u}_k = \vec{0}$. Moltiplicando scalarmente per $\vec{u}_i$, e tenendo conto che $g(\vec{u}_i, \vec{u}_j) = 0$ 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 $\mathcal{B} = \{\vec{u}_1, \ldots, \vec{u}_n\}$ di V è *ortonormale* se

$$g(\vec{u}_i, \vec{u}_j) = \begin{cases} 0 & \text{se } i \neq j, \\ 1 & \text{se } i = j. \end{cases}$$

Geometricamente, ciò significa che i vettori sono perpendicolari tra loro e di lunghezza unitaria. Si vede facilmente che la matrice B associata ad un cambiamento di basi ortonormali è ortogonale, cioè $B^t B = I$.

Teorema 5.4. *Ogni spazio vettoriale euclideo ammette una base ortonormale.*

Dimostrazione. La dimostrazione è costruttiva. Si fa vedere come da ogni base $\{\vec{u}_i\}_{1\leq i\leq n}$ si può costruire una base ortonormale $\{\vec{e}_i\}_{1\leq i\leq n}$ con un procedimento detto *ortonormalizzazione di Gram–Schmidt*.

Sia, dunque, $\{\vec{u}_i\}_{1\leq i\leq n}$ una base in uno spazio vettoriale euclideo con prodotto scalare g. Basta porre

$$\vec{e}_1 \stackrel{\text{def}}{=} \frac{\vec{u}_1}{\|\vec{u}_1\|_g}, \qquad \vec{v}_2 = \vec{u}_2 - g(\vec{u}_2, \vec{e}_1)\vec{e}_1, \quad \vec{e}_2 \stackrel{\text{def}}{=} \frac{\vec{v}_2}{\|\vec{v}_2\|_g},$$
$$\vec{v}_h = \vec{u}_h - \sum_{i=1}^{h-1} g(\vec{u}_h, \vec{e}_i)\vec{e}_i, \qquad \vec{e}_h \stackrel{\text{def}}{=} \frac{\vec{v}_h}{\|\vec{v}_h\|_g}.$$

QED

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

$$g(\vec{x}, \vec{y}) = \sum_{i=1}^{n} x_i y_i = {}^tXY, \qquad \|\vec{x}\|_g^2 = \sum_{i=1}^{n} x_i^2 = {}^tXX.$$

5.2.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

$$U^\perp \stackrel{\text{def}}{=} \{\vec{x} \in V \mid g(\vec{x}, \vec{u}) = 0 \ \forall \vec{u} \in U\}$$

è detto *complemento ortogonale* di U. Si lascia al lettore di provare che $U^\perp$ è un sottospazio vettoriale di V. Si dimostra anche che

$$V = U \oplus U^\perp, \tag{5.2.1}$$

che giustifica la dizione 'complemento ortogonale'.

Se $\{\vec{e}_i\}_{1\leq i\leq n}$ è una base <u>ortonormale</u> per V e $U = \mathcal{L}(\vec{e}_1, \dots, \vec{e}_p)$, allora

$$U^\perp = \mathcal{L}(\vec{e}_{p+1}, \dots, \vec{e}_n).$$

Ogni vettore $\vec{x} \in V$ si decompone in modo unico come segue

$$\vec{x} = \vec{x}_U + \vec{x}_{U^\perp}.$$

Essendo la base ortonormale

$$\vec{x} = \sum_{i=1}^{n} \lambda_i \vec{e}_i \quad \Rightarrow \quad \lambda_i = g(\vec{x}, \vec{e}_i);$$

gli scalari λ_i sono detti *coefficienti di Fourier di* $\vec{x}$. Segue che

$$\vec{x}_U = \sum_{i=1}^{p} \lambda_i \vec{e}_i, \qquad \vec{x}_{U^\perp} = \sum_{i=p+1}^{n} \lambda_i \vec{e}_i,$$

con $\lambda_i = g(\vec{x}, \vec{e}_i)$. Il vettore $\vec{x}_U$ è detto *proiezione ortogonale di* $\vec{x}$ *su* U: infatti $\vec{x}_U \in U$ e $\vec{x} - \vec{x}_U \in U^\perp$, cioè è ortogonale a U.

L'applicazione lineare

$$p_U \colon V \to V, \ \vec{x} \mapsto \vec{x}_U \tag{5.2.2}$$

è detta *proiezione ortogonale su* U e si vede facilmente che $p_U^2 = p_U$. Naturalmente se $\vec{x} \in U$ allora $p_U(\vec{x}) = \vec{x}$, mentre se $\vec{x} \in U^\perp$ allora $p_U(\vec{x}) = \vec{0}$, cioè $\operatorname{Im} p_U = U$, $\operatorname{Ker} p_U = U^\perp$.

5.2.d Applicazione aggiunta

Sia $f \colon V \to W$ un'applicazione lineare dallo spazio euclideo V (con prodotto scalare g) allo spazio vettoriale euclideo W (con prodotto scalare h). Allora si dimostra che esiste una ed una sola applicazione lineare $f^* \colon W \to V$ detta *aggiunta* di f tale che

$$h(f(\vec{x}), \vec{y}) = g(\vec{x}, f^*(\vec{y})) \qquad \forall \vec{x} \in V, \ \forall \vec{y} \in W.$$

Se $\mathcal{B}$ è una <u>base ortonormale</u> per V e $\mathcal{B}'$ è una <u>base ortonormale</u> per W, allora

$$\mathcal{M}^{\mathcal{B}}_{\mathcal{B}'}(f^*) = {}^t\!(\mathcal{M}^{\mathcal{B}'}_{\mathcal{B}}(f)).$$

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

$$(f_1 \circ f_2)^* = f_2^* \circ f_1^*.$$

5.2.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

$$g(f(\vec{x}), \vec{y}) = g(\vec{x}, f(\vec{y})) \qquad \forall \vec{x}, \vec{y} \in V.$$

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 $p_U \colon V \to V$, $\vec{x} \mapsto \vec{x}_U$ è un endomorfismo simmetrico. Infatti, se $\vec{x}, \vec{y} \in V$, allora $\vec{x} = \vec{x}_U + \vec{x}_{U^\perp}$ e $\vec{y} = \vec{y}_U + \vec{y}_{U^\perp}$.

$$\begin{aligned}
p_U(\vec{x}) \cdot \vec{y} &= p_U(\vec{x}_U + \vec{x}_{U^\perp}) \cdot (\vec{y}_U + \vec{y}_{U^\perp}) \\
&= (p_U(\vec{x}_U) + p_U(\vec{x}_{U^\perp})) \cdot (\vec{y}_U + \vec{y}_{U^\perp}) \\
&= \vec{x}_U \cdot (\vec{y}_U + \vec{y}_{U^\perp}) \\
&= \vec{x}_U \cdot \vec{y}_U, \\
\vec{x} \cdot p_U(\vec{y}) &= (\vec{x}_U + \vec{x}_{U^\perp}) \cdot \vec{y}_U \\
&= \vec{x}_U \cdot \vec{y}_U.
\end{aligned}$$

La funzione

$$Q_f \colon V \to \mathbb{R}, \quad Q_f(\vec{x}) \stackrel{\text{def}}{=} g(f(\vec{x}), \vec{x})$$

si dice *forma quadratica associata ad f*. Dall'espressione di Q_f

$$Q_f(\vec{x}) = {}^t(AX)X = {}^tX{}^tAX = {}^tXAX$$

si vede che Q_f è anche la forma quadratica associata all'applicazione bilineare simmetrica

$$\beta(\vec{x}, \vec{y}) \stackrel{\text{def}}{=} g(f(\vec{x}), \vec{y}) = {}^tXAY,$$

dove $A = (a_{ij})$ e $a_{ij} = \beta(\vec{e}_i, \vec{e}_j)$.

Gli endomorfismi simmetrici sono molto importanti per il seguente teorema.

Teorema 5.5. *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. *V ammette una base ortonormale di autovettoririspetto ad f.*

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 (${}^tB = B^{-1}$).

Se $\vec{x} = (x_1, \dots, x_n)$ rispetto alla base ortonormale di autovettori $\{\vec{e}_i\}_{1\leq i\leq n}$, allora
$$f(\vec{x}) = f(\sum_i x_i\vec{e}_i) = \sum_i x_i f(\vec{e}_i) = \sum_i x_i\lambda_i\vec{e}_i,$$
e quindi
$$Q_f(\vec{x}) = g(f(\vec{x}), \vec{x}) = \lambda_1 x_1^2 + \dots + \lambda_n x_n^2.$$
Abbiamo così un metodo standard per ridurre a forma canonica una forma quadratica: basta trovare gli autovalori della matrice simmetrica associata. Gli autovettori corrispondenti ad autovalori distinti risulteranno ortogonali, mentre per autovettori indipendenti corrispondenti ad uno stesso autovalore può rendersi necessario il procedimento di ortonormalizzazione.

5.2.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 $f = \lambda\,\mathrm{Id}_V$ 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 $Q_f(\vec{x}) = a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2$. Per quanto sopra detto, esiste una base ortonormale di autovettori $\mathcal{B} = \{\vec{e'}_1, \vec{e'}_2\}$ rispetto alla quale Q_f assume la forma canonica
$$Q_f(\vec{x}) = \lambda_1 {x'}_1^2 + \lambda_2 {x'}_2^2.$$

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 $\vec{v} = (x, y)$ e $Q(\vec{v}) = a_{11}x^2 + 2a_{12}xy + a_{22}y^2$ 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

   $$\mathcal{C}\colon \lambda_1\tilde{x}^2 + \lambda_2\tilde{y}^2 + d = 0.$$

   $$\begin{aligned} &\lambda_1\lambda_2 > 0 && \Rightarrow \quad \mathcal{C} \ \underline{\text{ellisse}}, \\ &\lambda_1\lambda_2 < 0 && \Rightarrow \quad \mathcal{C} \ \underline{\text{iperbole}}. \end{aligned}$$

   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.

5.2.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 $\{\vec{e'}_1, \vec{e'}_2, \vec{e'}_3\}$ con $\vec{e'}_i \in V(\lambda_i)$ 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 $\{\vec{e'}_1, \vec{e'}_2, \vec{e'}_3\}$ con $\vec{e'}_1, \vec{e'}_2 \in V(\lambda_1)$ ortonormali e $\vec{e'}_3 \in V(\lambda_3)$.

- Siano $\lambda_1 = \lambda_2 = \lambda_3$. Allora $f = \lambda \operatorname{Id}_V$ 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

$$Q_f(\vec{x}) = a_{11}x^2 + a_{22}y^2 + a_{33}z^2 + 2a_{12}xy + 2a_{13}xz + 2a_{23}yz.$$

Per quanto sopra detto, esiste una base ortonormale di autovettori $\mathcal{B}'$ rispetto alla quale Q_f assume la forma canonica $Q_f(\vec{x}') = \lambda_1 x'^2 + \lambda_2 y'^2 + \lambda_3 z'^2$.

Quanto detto è utile per la riduzione a forma canonica di una quadrica, che si può scrivere nella forma

$$\Sigma \colon Q_f(\vec{x}') + 2a'_{14}x' + 2a'_{24}y' + 2a'_{34}z' + a'_{44} = 0.$$

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

$$\lambda_1 \tilde{x}^2 + \lambda_2 \tilde{y}^2 + \lambda_3 \tilde{z}^2 + d = 0. \tag{5.2.3}$$

Esaminando i diversi casi, se $d \neq 0$ si perviene ad una delle seguenti forme:

$\frac{\tilde{x}^2}{a^2} + \frac{\tilde{y}^2}{b^2} + \frac{\tilde{z}^2}{c^2} = 1$ ellissoide reale

$\frac{\tilde{x}^2}{a^2} + \frac{\tilde{y}^2}{b^2} + \frac{\tilde{z}^2}{c^2} = -1$ ellissoide immaginario

$\frac{\tilde{x}^2}{a^2} + \frac{\tilde{y}^2}{b^2} - \frac{\tilde{z}^2}{c^2} = 1$ iperboloide ad una falda (o iperbolico)

$\frac{\tilde{x}^2}{a^2} - \frac{\tilde{y}^2}{b^2} - \frac{\tilde{z}^2}{c^2} = 1$ iperboloide a due falde (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

$\frac{\tilde{x}^2}{a^2} + \frac{\tilde{y}^2}{b^2} = 2z$ paraboloide ellittico

$\frac{\tilde{x}^2}{a^2} - \frac{\tilde{y}^2}{b^2} = 2z$ paraboloide iperbolico

Se nell'equazione (5.2.3) poniamo $d = 0$ si ottiene l'equazione canonica di coni quadrici (di vertice l'origine, che risulta un punto doppio). Più precisamente

$\frac{\tilde{x}^2}{a^2} + \frac{\tilde{y}^2}{b^2} + \frac{\tilde{z}^2}{c^2} = 0$ cono con generatrici immaginarie

$\frac{\tilde{x}^2}{a^2} + \frac{\tilde{y}^2}{b^2} - \frac{\tilde{z}^2}{c^2} = 0$ 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.

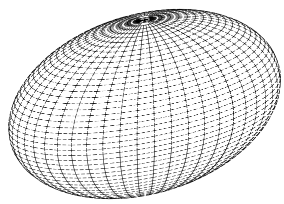

Figura II – 5.1: Ellissoide

[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.

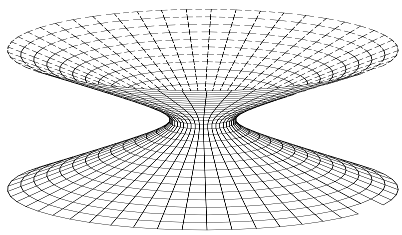

Figura II – 5.2: Iperboloide ad una falda

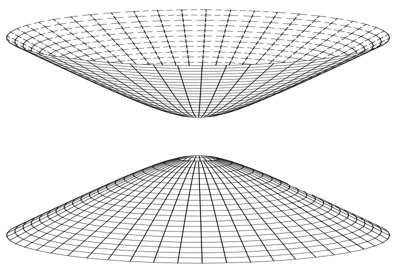

Figura II – 5.3: Iperboloide a due falde

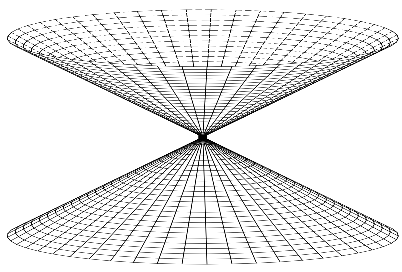

Figura II – 5.4: Cono

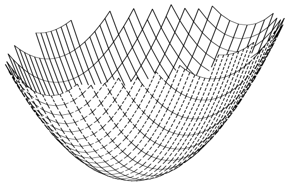

Figura II – 5.5: Paraboloide ellittico

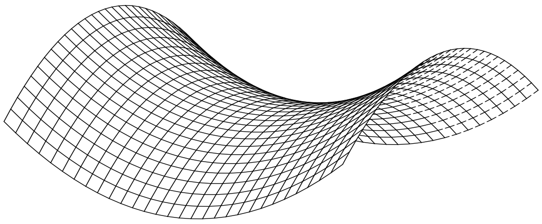

Figura II – 5.6: Paraboloide iperbolico

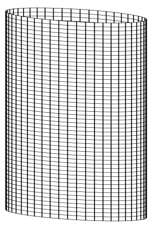

Figura II – 5.7: Cilindro ellittico

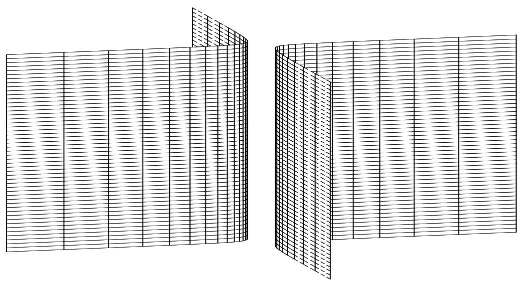

Figura II – 5.8: Cilindro iperbolico

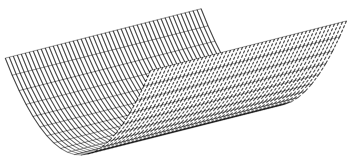

Figura II – 5.9: Cilindro parabolico

5.3 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.

5.3.a Definizione

Sia (V, g) uno spazio vettoriale euclideo. Un endomorfismo $f\colon V \to V$ si dice *trasformazione ortogonale* se

$$g(f(\vec{x}), f(\vec{y})) = g(\vec{x}, \vec{y}) \qquad \forall \vec{x}, \vec{y} \in V,$$

cioè se f è un'applicazione che conserva il prodotto scalare. In particolare, f trasforma basi ortonormali in basi ortonormali.

Teorema 5.6. *Sia f un endomorfismo di uno spazio vettoriale euclideo V con prodotto scalare g. Le seguenti condizione sono equivalenti:*

1. $g(f(\vec{x}), f(\vec{y})) = g(\vec{x}, \vec{y})$ *per ogni* $\vec{x}, \vec{y} \in V$*;*
2. $\|f(\vec{x})\|_g = \|\vec{x}\|_g$ *per ogni* $\vec{x} \in V$*;*
3. $f^* \circ f = \mathrm{Id}_V = f \circ f^*$*, ossia* $f^* = f^{-1}$*.*

Dimostrazione. (1) ⇒ (2). Si ponga $\vec{x} = \vec{y}$.
(2) ⇒ (1). Segue tenendo conto del fatto che

$$2g(\vec{x}, \vec{y}) = \|\vec{x} + \vec{y}\|_g^2 - \|\vec{x}\|_g^2 - \|\vec{y}\|_g^2$$

e analogamente per $f(\vec{x})$ e $f(\vec{y})$.
(1) ⇒ (3). Infatti $g(f(\vec{x}), f(\vec{y})) = g(\vec{x}, f^*(f(\vec{y})))$ per ogni $\vec{x}, \vec{y} \in V$. Ma $g(f(\vec{x}), f(\vec{y})) = g(\vec{x}, \vec{y})$, quindi

$$g(\vec{x}, (f^*(f(\vec{y})) - \vec{y})) = 0 \quad \forall \vec{x}, \vec{y} \in V \quad \Rightarrow \quad f^*(f(\vec{y})) = \vec{y} \ \ \forall \vec{y}.$$

(3) ⇒ (1). Per definizione $g(f(\vec{x}), \vec{z}) = g(\vec{x}, f^*(\vec{z}))$ per ogni $\vec{x}, \vec{z} \in V$. Posto $\vec{z} = f(\vec{y})$ e ricordando che $f^* = f^{-1}$ si ha la tesi.
(2) ⇔ (3). Segue tenendo conto le equivalenze sopra dimostrate. QED

Dal punto (2) segue che

$$\mathrm{dist}(f(\vec{x}), f(\vec{y})) = \|f(\vec{x}) - f(\vec{y})\|_g = \|\vec{x} - \vec{y}\|_g = \mathrm{dist}(\vec{x}, \vec{y}),$$

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*. Un cambiamento di base ortonormale è rappresentato da una matrice ortogonale.

Dal punto (3) segue anche che f è un isomorfismo.

Teorema 5.7. *Sia V uno spazio vettoriale euclideo con prodotto scalare g, e sia $f\colon V \to V$ una trasformazione ortogonale. Gli eventuali autovalori (reali) di f sono ± 1.*

Dimostrazione. Sia λ un autovalore di f. Allora

$$f(\vec{x}) = \lambda\vec{x} \quad \Rightarrow \quad \|f(\vec{x})\|_g = \|\lambda\vec{x}\|_g = |\lambda|\|\vec{x}\|_g.$$

Ma $\|f(\vec{x})\|_g = \|\vec{x}\|_g$ da cui la tesi, essendo $\|\vec{x}\|_g \neq 0$. 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 $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. 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.

5.3.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

$$\mathbb{O}^+(n;\mathbb{R}) \stackrel{\text{def}}{=} \{A \in \mathbb{O}(n;\mathbb{R}) \mid \det A = 1\},$$
$$\mathbb{O}^-(n;\mathbb{R}) \stackrel{\text{def}}{=} \{A \in \mathbb{O}(n;\mathbb{R}) \mid \det A = -1\}.$$

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

$$U = \{\vec{x} \in \mathbb{R}^3 \mid f(\vec{x}) = \vec{x}\}.$$

Si può dimostrare che

$$\begin{aligned}
\dim U = 3 \quad &\Rightarrow \quad f \text{ identità}\\
\dim U = 2 \quad &\Rightarrow \quad f \text{ simm. ort. risp. } U\\
\dim U = 1 \quad &\Rightarrow \quad f \text{ rotaz. intorno } U\\
\dim U = 0 \quad &\Rightarrow \quad f \text{ rotosimmetria}
\end{aligned}$$

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 $U \neq \{\vec{0}\}$,

$$\begin{aligned}
\dim U = 3 \quad &\Rightarrow \quad A = \mathrm{Id}\\
\dim U = 2 \quad &\Rightarrow \quad A \sim \begin{pmatrix} -1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}\\
\dim U = 1 \quad &\Rightarrow \quad A \sim \begin{pmatrix} 1 & 0 & 0\\ 0 & \cos\varphi & -\sin\varphi\\ 0 & \sin\varphi & \cos\varphi\end{pmatrix} = \begin{pmatrix} 1 & 0\\ 0 & R(\varphi)\end{pmatrix} = \tilde{R}(\varphi)\\
\dim U = 0 \quad &\Rightarrow \quad A \sim \begin{pmatrix} -1 & 0\\ 0 & R(\varphi)\end{pmatrix} = \tilde{S}(\varphi)
\end{aligned}$$

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.

5.3.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

$$m = t_{\vec{b}} \circ f \colon V \to V,\ \vec{x} \mapsto f(\vec{x}) + \vec{b},$$

oppure $m = f \circ t_{\vec{c}}$ con $\vec{c} = f^{-1}(\vec{b})$.

Si osservi che, se $\vec{b} \neq \vec{0}$, 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

$$\operatorname{dist}(m(\vec{x}), m(\vec{y})) = \|m(\vec{x}) - m(\vec{y})\|_g == \|f(\vec{x} - \vec{y})\|_g = \|\vec{x} - \vec{y}\|_g = \operatorname{dist}(\vec{x}, \vec{y}).$$

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 5.8. *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 $\mathcal{B} = \{\vec{e}_1, \ldots, \vec{e}_n\}$ una base ortonormale di V e $\mathcal{M}^{\mathcal{B}}_{\mathcal{B}}(f) = A \in \mathbb{O}(n; \mathbb{R})$. Posto $\vec{x} = (x_1, \ldots, x_n)$, $m(\vec{x}) = (x'_1, \ldots, x'_n)$, $\vec{b} = (b_1, \ldots, b_n)$, 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

$$\begin{aligned} A &\in \mathbb{O}^+(n, \mathbb{R}) && \Rightarrow && m \text{ rototraslazione}, \\ A &\in \mathbb{O}^-(n, \mathbb{R}) && \Rightarrow && m \text{ glissosimmetria}, \end{aligned}$$

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 $\mathbb{R}^n$.

BIBLIOGRAFIA

[1] E. Abbena, E. Bargero, S. Garbiero: Esercizi di Algebra Lineare e Geometria Analitica, Levrotto e Bella, Torino 1990.

[2] S. Abeasis: Algebra lineare e geometria, Zanichelli 1990.

[3] I. Cattaneo Gasparini: Strutture algebriche, operatori lineari, Roma, 1989.

[4] G. Calvaruso, R. Vitolo: Esercizi di Geometria ed Algebra, Università di Lecce, 2001.

[5] P. de Bartolomeis: Algebra Lineare, La Nuova Italia, Firenze 1993.

[6] F.Fava, F. Tricerri: Geometria ed Algebra Lineare, Levrotto e Bella, Torino 1987.

[7] G. Gambarelli: Benvenuto alle matricole (e qualche consiglio), in AA. VV., L'endecasillabo matematico, Mathesis 1999.

[8] W. Greub: Linear Algebra, IV ed., Springer, 1975.

[9] P. Maroscia: Problemi di Geometria, Masson, 1994.

[10] P. Maroscia: Geometria ed Algebra Lineare, Zanichelli, 2002.

[11] V. K. Nicholson: Algebra lineare, dalle applicazioni alla teoria, McGraw–Hill, 2002.

[12] A. Sanini: Lezioni di Geometria, ed. Levrotto-Bella, Torino 1993; Esercizi di Geometria, ed. Levrotto-Bella, Torino 1993.

[13] E. Sernesi: Geometria 1, Boringhieri, Bologna 1989.

[14] G. Vaccaro, A. Carfagna, L. Piccolella: Lezioni di geometria ed algebra lineare, Masson, 1995.

INDICE ANALITICO