

Geometria e Algebra Lineare

Rafael Mosca

Scritto con L^AT_EX

Quest'opera stata rilasciata con licenza Creative Commons Attribuzione: - Non commerciale - Non opere derivate 4.0 Internazionale.

This work is licensed under a Creative Commons “Attribution-NonCommercial-NoDerivatives 4.0 International” license.



Per leggere una copia della licenza visita il sito web <http://creativecommons.org/licenses/by-sa/4.0/deed.it> o spedisci una lettera a Creative Commons, 171 Second Street, Suite 300, San Francisco, California, 94105, USA.

Ringraziamenti

Si ringrazia il Prof. Giovanni Vannozzi per la minuziosa revisione, critiche e correzioni.

Insiemistica

Indico con A, B, \dots degli **insiemi** che possiamo definire informalmente come un raggruppamento di elementi di qualsiasi tipo appartenenti all'universo che indichiamo con U . Anche se non possiamo dare a un insieme una definizione precisa, possiamo chiederci se qualcosa, chiamiamolo x , appartiene o meno ad un insieme, e si scriverà:

$$x \in A \quad x \text{ appartiene ad } A$$

$$x \notin A \quad x \text{ non appartiene ad } A$$

Se A e B sono insiemi e succede che $\forall x \in A, x \in B$ allora si dice che A è un **sottoinsieme** di B e si scrive:

$$A \subseteq B \quad A \text{ è sottoinsieme di } B$$

in particolare diciamo che $A = B$ se $A \subseteq B$ e $B \subseteq A$ e $A \subset B$ se $A \subseteq B$ e $A \neq B$

0.1 Operazioni sugli insiemi

Su questi insiemi appena possiamo definire delle operazioni che sono facilmente rappresentabili con i diagrammi di Eulero-Venn:

UNIONE e si scrive $A \cup B$ se:

$$\{x \in U \mid x \in A \vee x \in B\}$$

INTERSEZIONE e si scrive $A \cap B$ se:

$$\{x \in U \mid x \in A \wedge x \in B\}$$

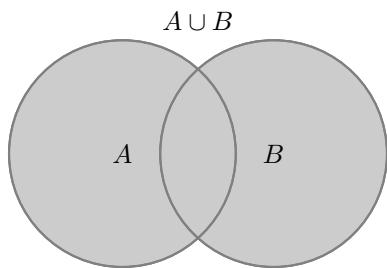


Figura 1: Unione

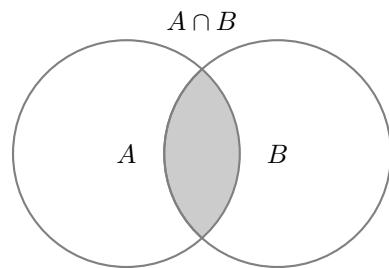


Figura 2: Intersezione

DIFFERENZA e si scrive $A \setminus B$ oppure $A - B$
se:

$$\{x \in A \mid x \notin B\}$$

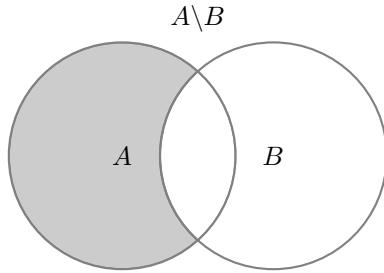


Figura 3: Differenza

DIFFERENZA SIMMETRICA e si scrive $A \Delta B$
e corrisponde a :

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

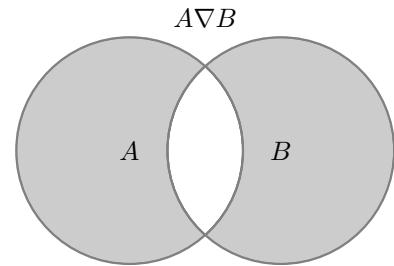


Figura 4: Differenza simmetrica

Per ultimo abbiamo l'operazione COMPLEMENTO che di solito viene scritta come A^c oppure \bar{A} ed è definita come: $\{x \in U \mid x \notin A\}$, in parole poche il complementare di A è tutto quello che non sta in A cioè non è soltanto B ma $B \cup U$. Le operazioni appena introdotte godono di certe proprietà, in particolare:

(1) *Associatività*

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

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

(2) *Distributività*

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

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

(3) *Prima Legge di De Morgan*

$$\overline{A \cap B} = \overline{A} \cup \overline{B}$$

(4) *Seconda Legge di De Morgan*

$$\overline{A \cup B} = \overline{A} \cap \overline{B}$$

0.2 Prodotto Cartesiano

Definiamo come **prodotto cartesiano** di A e B , l'insieme che ha per elementi tutte le possibili coppie ordinate di elementi dei due insiemi.

$$A \times B = \{(x, y) \mid x \in A \wedge y \in B\}$$

0.3 Relazioni

Siano A_1, A_2, \dots, A_n insiemi, chiamiamo una relazione tra A_1, A_2, \dots, A_n un sottoinsieme di $A_1 \times A_2, \dots \times A_n$, più nello specifico viene dato il nome di relazione $n - \text{aria}$ dove con aria ci riferiamo all'arietà, che non è altro che il numero degli argomenti o operandi che richiede la relazione. cioè

$$R \subseteq A \times B$$

In particolare chiamiamo:

$$\text{relazione vuota} \quad \emptyset \subseteq A \times B$$

$$\text{relazione totale} \quad A \times B \subseteq A \times B$$

Supponiamo ora di avere una relazione $R \subseteq A \times B$ e $a \in A$ e $b \in B$ per dire che a è in relazione con b rispetto a R invece di scrivere $(a, b) \in R$ possiamo scrivere

$$a R b$$

0.3.1 Relazione di equivalenza

Una relazione di equivalenza formalizza il concetto di un insieme di oggetti che hanno una stessa proprietà. Una relazione d'equivalenza soddisfa le seguenti proprietà:

- ① è RIFLESSIVA $\forall x \quad x \sim x$
- ② è SIMMETRICA $\forall x, y \quad x \sim y \Rightarrow y \sim x$
- ③ è TRANSITIVA $\forall x, y, z \quad x \sim y \text{ e } y \sim z \Rightarrow x \sim z$

0.3.2 Relazione d'ordine

Una relazione d'ordine (POSET in inglese) formalizza il concetto d'ordine o posizione degli elementi di un insieme. Una relazione d'ordine soddisfa le seguenti proprietà:

- ① è RIFLESSIVA $\forall x \quad x \sim x$
- ② è ANTISIMMETRICA $\forall x, y \quad x \sim y \text{ e } y \sim x \Rightarrow x = y$
- ③ è TRANSITIVA $\forall x, y, z \quad x \sim y \text{ e } y \sim z \Rightarrow x \sim z$

Inoltre si dice che si tratta di un ordine totale se:

- ④ è TOTALE $\forall x, y \quad x \sim y \text{ o } y \sim x$

Spazi Vettoriali

Definizione 1.1. Chiamiamo **operazione binaria interna** una funzione che associa ad una coppia di elementi di un insieme X un elemento nello stesso insieme d'arrivo. $X \times X \rightarrow X$

La struttura $(V, +)$ si dice *gruppo commutativo* o **gruppo abeliano**, rispetto a una operazione interna $+ : V \times V \rightarrow V$ se sono soddisfatte le seguenti condizioni:

1. l'operazione binaria $+$ è *associativa*:

$$(\mathbf{v}_1 + \mathbf{v}_2) + \mathbf{v}_3 = \mathbf{v}_1 + (\mathbf{v}_2 + \mathbf{v}_3)$$

2. Esiste in V un *elemento neutro* rispetto a $+$, cioè un elemento $\underline{0}_V \in V$ tale che:

$$\mathbf{v} + \underline{0}_V = \underline{0}_V + \mathbf{v} = \mathbf{v}, \forall \mathbf{v} \in V$$

3. Esiste un elemento *inverso* rispetto a $+$ in V , cioè: $\underline{0}_V \neq \mathbf{v} \in V$, tale che $\forall v \in V \exists \mathbf{v}' \in V$:

$$\mathbf{v} + \mathbf{w} = \underline{0}_V \quad e \quad \mathbf{w} + \mathbf{v} = \underline{0}_V$$

Indichiamo tale elemento con $-\mathbf{v}$.

4. L'operazione di somma tra vettori è *commutativa* con $\mathbf{v}, \mathbf{w} \in V$:

$$\mathbf{v} + \mathbf{w} = \mathbf{w} + \mathbf{v}$$

Si noti che l'elemento neutro non è il numero zero ma un elemento appartenente all'insieme che chiamiamo **vettore nullo** e il sotto-indice sta a indicare l'appartenenza all'insieme V .

Lo stesso vale per una struttura (V, \cdot) dove l'operazione binaria è un prodotto esterno su V con \mathbb{R} , cioè un'operazione $\cdot : \mathbb{R} \times V \rightarrow V$ che soddisfa le proprietà seguenti:

1. È *associativo*, vale a dire:

$$\alpha \cdot (\beta \cdot \mathbf{v}) = (\alpha\beta) \cdot \mathbf{v}$$

2. Esiste l'*elemento neutro* rispetto all'operazione di prodotto di un vettore per uno scalare:

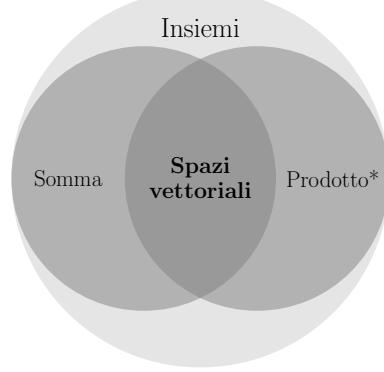
$$1 \cdot \mathbf{v} = \mathbf{v}$$

3. Il prodotto di un vettore per uno scalare è *distributivo rispetto alla somma di vettori*:

$$\alpha \cdot (\mathbf{v} + \mathbf{w}) = \alpha \cdot \mathbf{v} + \alpha \cdot \mathbf{w}$$

4. Il prodotto per uno scalare è *distributivo rispetto alla somma di elementi di \mathbb{R}* (che indichiamo con abuso di notazione come $+ : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$:

$$(\alpha + \beta) \cdot \mathbf{v} = \alpha \cdot \mathbf{v} + \beta \cdot \mathbf{v}$$



Definizione 1.2. Sia V un insieme, " + " un' operazione $+ : V \times V \rightarrow V$ e " . " un' operazione $\cdot : \mathbb{R} \times V \rightarrow V$ tali che:

1. $(V, +)$ è un gruppo abeliano.
2. (V, \cdot) è un gruppo abeliano.

Allora si dice che $(V, +, \cdot)$ è uno **spazio vettoriale reale** (poiché definito in \mathbb{R} nel prodotto per uno scalare).

Figura 1.1: Spazio vettoriale

Definizione 1.3. Se V è uno spazio vettoriale allora ogni suo elemento si chiama **vettore**.

ESEMPIO

Il piano cartesiano a cui siamo abituati non è altro che la rappresentazione di uno spazio vettoriale (\mathbb{R}^2) e alcuni suoi elementi (vettori) sono per esempio $(1,0)$ $(3,2)$ $(4,9)$, ecc. Quello che però non si deve pensare è che ne esistano soltanto spazi di forma \mathbb{R}^n , ce ne sono tanti altri, per esempio è anche uno spazio vettoriale l'insieme dei polinomi di grado al più n a coefficienti reali, che indichiamo con:

$$\mathbb{R}_n[x] := \{a_0 + a_1x + a_2x^2 + \dots + a_nx^n \mid a_i \in \mathbb{R} \ \forall i = 1, \dots, n\}$$

Definizione 1.4. Sia V uno spazio vettoriale e siano $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n \in V$ n vettori che gli appartengono e siano $\lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{R}$ coefficienti moltiplicativi, si chiama **combinazione lineare** dei vettori \mathbf{v}_i di coefficienti λ_i l'espressione:

$$\lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2 + \dots + \lambda_n\mathbf{v}_n \quad \equiv \quad \sum_{i=1}^n \lambda_i\mathbf{v}_i$$

Una combinazione lineare tra vettori prende tale nome perché è definita solamente mediante operazioni lineari: la somma tra vettori e il prodotto di un vettore per uno scalare. È importante notare che una combinazione lineare di vettori è sempre un vettore e non un numero come si potrebbe pensare.

Definizione 1.5. Sia V uno spazio vettoriale e siano $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ vettori. Si dice che i vettori sono (fra loro) **linearmente indipendenti** se l'unica loro combinazione lineare che dia il vettore nullo è quella con coefficienti tutti nulli, cioè:

$$\lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2 + \dots + \lambda_n\mathbf{v}_n = \underline{0}_V \quad e \quad \lambda_1 = \lambda_2 = \dots = \lambda_n = 0$$

Analogamente, se $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ non sono linearmente indipendenti si dice che sono **linearmente dipendenti** questo significa che esiste una loro combinazione lineare che fornisce il vettore nullo con coefficienti non tutti nulli.

ESEMPIO

Consideriamo i vettori $(0,3) (2,1) (0,2) \in \mathbb{R}^2$, vogliamo verificare se i vettori sono tra loro linearmente indipendenti o dipendenti. Dobbiamo quindi valutare l'espressione della combinazione lineare

$$\alpha \begin{bmatrix} 0 \\ 3 \end{bmatrix} + \beta \begin{bmatrix} 2 \\ 1 \end{bmatrix} + \gamma \begin{bmatrix} 0 \\ 2 \end{bmatrix} = \underline{0}$$

e trovare, se esistono, i valori di α, β, γ che sostituiti nell'espressione ci diano il vettore nullo. Possiamo dunque scrivere il seguente sistema di equazioni:

$$\begin{cases} 0\alpha + 2\beta + 0\gamma = 0 \\ 3\alpha + 1\beta + 2\gamma = 0 \end{cases} \Rightarrow \begin{cases} 2\beta = 0 \\ 3\alpha + 1\beta + 2\gamma = 0 \end{cases} \Rightarrow \begin{cases} \beta = 0 \\ 3\alpha + 2\gamma = 0 \end{cases} \Rightarrow \begin{cases} \beta = 0 \\ \alpha = -\frac{2}{3}\gamma \end{cases}$$

Questo sistema ci ha portato a una soluzione non nulla, infatti quello che ci dice la soluzione e che l'espressione

$$\frac{-2}{3}\gamma \begin{bmatrix} 0 \\ 3 \end{bmatrix} + 0\beta \begin{bmatrix} 2 \\ 1 \end{bmatrix} + \gamma \begin{bmatrix} 0 \\ 2 \end{bmatrix} = \underline{0} \quad \xrightarrow{\text{tutto per } 3} \quad -2\gamma \begin{bmatrix} 0 \\ 3 \end{bmatrix} + 0\beta \begin{bmatrix} 2 \\ 1 \end{bmatrix} + 3\gamma \begin{bmatrix} 0 \\ 2 \end{bmatrix} = \underline{0}$$

da sempre il vettore nullo indipendentemente dal valore scelto per γ , assumiamo per esempio $\gamma = 2$ allora si ha:

$$-4 \begin{bmatrix} 0 \\ 3 \end{bmatrix} + 6 \begin{bmatrix} 0 \\ 2 \end{bmatrix} = \underline{0} \quad \Rightarrow \quad \begin{bmatrix} 0 \\ -12 \end{bmatrix} + \begin{bmatrix} 0 \\ 12 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

quindi possiamo dire che i vettori sono tra loro linearmente dipendenti poiché abbiamo trovato dei coefficienti diversi da zero che danno il vettore nullo nell'espressione della combinazione lineare . Analizziamo ora la lineare dipendenza/indipendenza dei vettori $(2,4) (9,1)$ seguendo lo stesso ragionamento di prima, valutiamo l'espressione:

$$\alpha \begin{bmatrix} 2 \\ 4 \end{bmatrix} + \beta \begin{bmatrix} 9 \\ 1 \end{bmatrix} = \underline{0}$$

che ci da il sistema

$$\begin{cases} 2\alpha + 9\beta = 0 \\ 4\alpha + 1\beta = 0 \end{cases} \Rightarrow \begin{cases} 2\alpha + 9\beta = 0 \\ \beta = -4\alpha \end{cases} \Rightarrow \begin{cases} 2\alpha - 36\alpha = 0 \\ \beta = -4\alpha \end{cases} \Rightarrow \begin{cases} -34\alpha = 0 \\ \beta = -4\alpha \end{cases} \Rightarrow \begin{cases} \alpha = 0 \\ \beta = -4 \cdot 0 = 0 \end{cases}$$

Poiché l'unica loro combinazione lineare che da il vettore nullo è quella a coefficienti tutti nulli, i vettori sono linearmente indipendenti.

La lineare dipendenza vuol dire grossolanamente che un vettore può essere ottenuto moltiplicando e sommando opportunamente gli altri vettori. Con questo ragionamento risulta facile distinguere per insiemi di vettori relativamente piccoli se i vettori sono linearmente dipendenti e in questi casi risulta inutile calcolare tutta l'espressione, ad esempio nel primo caso bastava vedere che il primo e

il terzo erano proporzionali e potevamo concludere senza fare tutti quei conti che erano linearmente dipendenti. Se volessimo analizzare ad esempio la lineare dipendenza/indipendenza dei vettori $(1,2,1);(1,0,1);(3,4,3)$ basterebbe vedere che il terzo vettore può essere ottenuto come due volte il primo più il secondo vettore per cui sono tra loro linearmente dipendenti.

Il concetto di lineare dipendenza/indipendenza è di fondamentale importanza nell'algebra lineare poiché rappresenta in qualche modo la quantità d'informazione utile, infatti supponiamo di avere un sistema di equazioni

$$\begin{cases} 2x + 3y = 2 \\ 4x + 6y = 4 \end{cases}$$

Anche se è un sistema di due equazioni in due incognite non possiamo trovare un'unica soluzione poiché le equazioni risultano essere tra loro linearmente dipendenti, infatti, la seconda può essere ottenuta dalla prima moltiplicandola per 2. Riprenderemo questo concetto successivamente.

Teorema 1.1. Se $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ sono linearmente dipendenti, allora esiste $1 \leq i \leq n$ tale che \mathbf{v}_i è combinazione lineare di $\mathbf{v}_1, \dots, \mathbf{v}_{i-1}, \mathbf{v}_{i+1}, \dots, \mathbf{v}_n$

Questo teorema non è altro che una conseguenza della definizione di lineare dipendenza, cioè esistono $\lambda_1, \dots, \lambda_n$ non tutti nulli tali che

$$\lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 + \dots + \lambda_i \mathbf{v}_i + \dots + \lambda_n \mathbf{v}_n = \underline{0}$$

Supponiamo che λ_i sia non nullo allora portando il termine λ_i a destra del segno di uguaglianza ci rimane l'espressione.

$$\lambda_1 \mathbf{v}_1 + \dots + \lambda_{i-1} \mathbf{v}_{i-1} + \lambda_{i+1} \mathbf{v}_{i+1} + \dots + \lambda_n \mathbf{v}_n = -\lambda_i \mathbf{v}_i$$

Per cui \mathbf{v}_i può essere scritto come:

$$\mathbf{v}_i = \frac{1}{-\lambda_i} (\lambda_1 \mathbf{v}_1 + \dots + \lambda_{i-1} \mathbf{v}_{i-1} + \lambda_{i+1} \mathbf{v}_{i+1} + \dots + \lambda_n \mathbf{v}_n)$$

Definizione 1.6. Sia V uno spazio vettoriale e siano vettori $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n \in V$ si chiama **insieme generato** da i vettori \mathbf{v} l'insieme di tutte le combinazioni lineari dei vettori \mathbf{v} :

$$\langle \mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n \rangle = \{ \lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 + \dots + \lambda_n \mathbf{v}_n \mid \lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{R} \}$$

Definizione 1.7. Sia V uno spazio vettoriale e sia $G = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ un insieme di vettori. Si dice che G è un **sistema di generatori** di V se ogni vettore $\mathbf{v} \in V$ può essere scritto come combinazione lineare di vettori di G :

$$\mathbf{v} = \lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 + \dots + \lambda_n \mathbf{v}_n$$

che può essere anche espressa come:

$$\mathbf{v} \in \langle \mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n \rangle$$

Inoltre se G è un insieme finito di vettori si dice che \mathbf{v} è **finitamente generabile**.

Proprietà 1.1. Sia V uno spazio vettoriale, sia $G = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ un sistema di generatori e $\mathbf{u} \in V$ allora anche $G \cup \{\mathbf{u}\}$ è un sistema di generatori.

Definizione 1.8. Sia V uno spazio vettoriale, sia $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ si dice che \mathcal{B} è una **base** di V se:

1. \mathcal{B} è un insieme di generatori.
2. \mathcal{B} è un insieme di vettori linearmente indipendenti.

Definizione 1.9. Sia V uno spazio vettoriale, si chiama **dimensione** di V il numero di vettori costituenti una sua base. Si scrive:

$$\dim(V)$$

In casi molto rari, viene chiamata dimensione di Hamel o dimensione algebrica, per distinguerla da altri tipi di dimensione.

Teorema 1.2 (UNICITÀ DELLA COMBINAZIONE LINEARE). Sia V uno spazio vettoriale e $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ una base di V allora ogni vettore di V si scrive in modo unico come combinazione lineare di elementi di \mathcal{B} .

Definizione 1.10. Sia V uno spazio vettoriale e $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ una base di V si chiamano **coordinate** di u rispetto a \mathcal{B} gli unici coefficienti $\lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{R}$ tale che $u = \lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 + \dots + \lambda_n \mathbf{v}_n$ e si scrive:

$$u = (\lambda_1, \lambda_2, \dots, \lambda_n)_{\mathcal{B}}$$

Si osservi che se $\mathcal{B} \neq \mathcal{B}'$ le coordinate sono diverse

ESEMPIO

I concetti introdotti possono sembrare complessi dalla definizione ma in realtà sono abbastanza facili. Facciamo un esempio concreto, consideriamo tutti i colori visibili anche se non sono un vero e proprio spazio vettoriale. Come tutti sappiamo, uno schermo elettronico funziona accendendo dei pixel, questi pixel assumono tutti i colori possibili facendo variare l'intensità di tre colori: rosso, verde e blu (RGB) quindi possiamo dire che questi tre colori sono un insieme di generatori per i colori visibili (poiché possono generare tutti gli altri) e in particolare una base (poiché i tre colori sono linearmente indipendenti), questo vuol dire che basta combinare opportunamente l'intensità di questi tre colori per generare tutti gli altri. Nell'ambito dei colori succede una cosa affascinante, le stampanti a iniezione riescono a generare lo stesso spazio di colori facendo una combinazione lineare di colori diversi: ciano, magenta e giallo (CMYK) (in realtà usano anche il nero perché si otterrebbe combinando tutti al 100% e risulterebbe poco efficiente), questo esempio è un esempio concreto che ci aiuta a capire che anche se il sistema di generatori (ovvero i colori con cui si generano tutti gli altri) è diverso, ciò non vuol dire che lo spazio generato sia diverso, in realtà lo spazio generato è lo stesso!

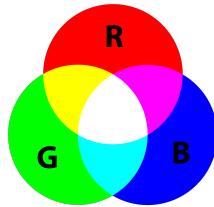


Figura 1.2: RGB

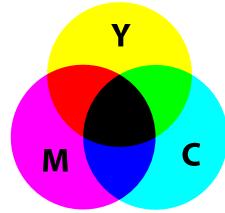
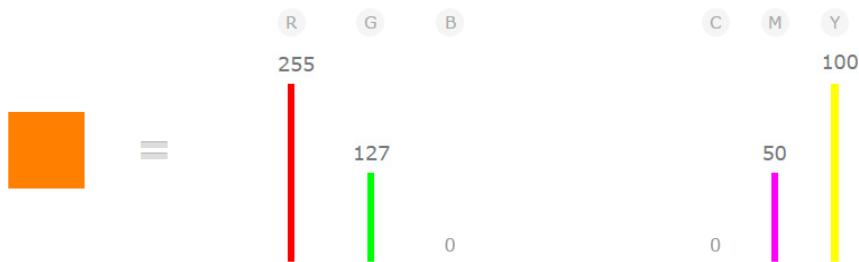
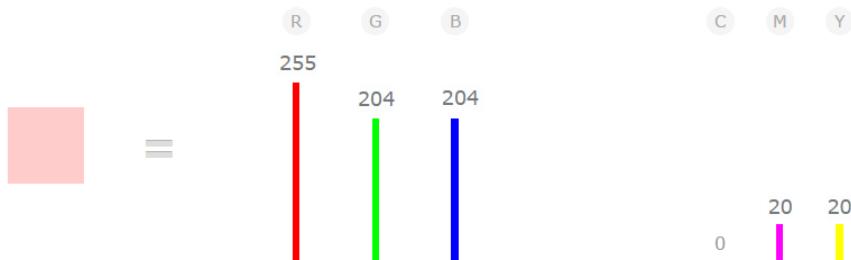


Figura 1.3: CYMK

Questi però, non sono gli unici insiemi di generatori, infatti se considerassimo i colori rosso, verde, blu, arancione e rosa, questi formano anche un insieme di generatori capaci di generare tutto lo spazio ma non sono una base giacché non sono tra loro linearmente indipendenti, ad esempio nella maggior parte dei computer i colori sono rappresentati come combinazione lineare del rosso, verde e blu, ove ogni colore può assumere 256 intensità diverse (8bit per colore) che ci danno $256^3 = 16\,777\,216$ combinazioni diverse, l'arancione in questo caso può scriversi come:



e il rosa come:



per cui questi due colori, così come tutto il resto, possono essere generati da solo tre di essi. In questo caso i coefficienti d'intensità assumono il ruolo di quello che chiamiamo coordinata ovvero i coefficienti unici che moltiplicano gli elementi della base e che lo contraddistinguono da tutti gli altri colori. Possiamo vedere che c'è un unico modo per generare il colore rispetto a una base, cioè non esiste un'altra combinazione rispetto alla base che ci dia lo stesso colore, può esistere però una combinazione diversa che ci dia lo stesso colore solo se assumiamo una base diversa!

In sintesi, un sistema di generatori è un insieme di elementi (vettori, colori) che genera tutto lo spazio mentre una base è un insieme minimo di vettori che genera tutto lo spazio. Possiamo inoltre dire che la dimensione dei colori visibili è $\dim = 3$ poiché bastano tre colori linearmente indipen-

denti per generare tutti gli altri.

Questo esempio non è a caso, infatti un grande studioso dell'algebra lineare, Hermann Günther Grassman (che rincontreremo a fine capitolo), riuscì a capire che i colori possono essere generati come combinazione lineare di tre colori linearmente indipendenti, questa sua scoperta e ora quella che viene chiamata la prima legge della mescolanza additiva di stimoli di colore.

Grassman's first law of colour mixture:

"Any colour can be matched by a linear combination of three other colours, provided that none of those three can be matched by a combination of the other two"

Infatti ci basta prendere tre colori linearmente indipendenti e sommandoli opportunamente possiamo ottenere tutti i colori possibili.

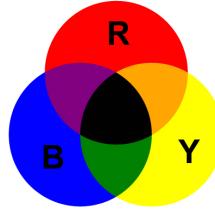


Figura 1.4: Un'altra possibile base per i colori (RYB)

Lemma 1.1. Sia V uno spazio vettoriale e siano:

$G = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ un sistema di generatori di V .

$I = \{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_j\}$ un insieme di vettori linearmente indipendenti.

Allora:

$$j \leq n$$

Quello che ci dice questo lemma è che il numero di vettori linearmente indipendenti è sempre minore o in certi casi uguale al numero di vettori di un sistema di generatori.

Teorema 1.3 (CARDINALITÀ DELLE BASI). Sia V uno spazio vettoriale e siano \mathcal{B} e \mathcal{B}' due basi finite di V ; allora

$$|\mathcal{B}| = |\mathcal{B}'|$$

Questo teorema ci dice che due basi hanno la stessa **cardinalità** ovvero lo stesso numero di elementi, riprendendo l'esempio dei colori possiamo vedere che anche se i colori possono essere generati da basi diverse (RGB), (CMYK) e (RYB) il numero di elementi (colori) è sempre lo stesso (3).

Ritornando all'ambito degli spazi vettoriali, $V = \{\underline{0}_V\}$ (il solo vettore nullo) è uno spazio vettoriale ma non ha una base ovvero $\dim(V) = 0$; si deve notare anche che il vettore nullo $\underline{0}_V$ non può mai far parte di una base giacché è linearmente dipendente.

Teorema 1.4 (TEOREMA DI ESISTENZA DELLE BASI). Sia V uno spazio vettoriale non nullo ($V \neq \underline{0}_V$) se V ha un sistema di generatori, allora ha anche una base.

In realtà ci basta notare che ogni spazio vettoriale ha una base poiché se prendiamo tutti i vettori appartenenti a V , questi sono un sistema di generatori di se stessi. Ora per renderlo una base quello che ci basta fare è togliere tutti i vettori linearmente dipendenti finché non si ha un insieme linearmente indipendente (che è una base per definizione).

ESEMPIO

Ci viene detto che $\mathbf{v}_1 = (2, 1, 3)$, $\mathbf{v}_2 = (3, 1, 0)$, $\mathbf{v}_3 = (0, 0, 4)$, $\mathbf{v}_4 = (1, 1, 2)$ è un insieme di generatori per \mathbb{R}^3 di cui vogliamo trovarne una base. Se tutti i vettori sono linearmente indipendenti allora l'insieme di vettori è già una base, altrimenti toglieremo all'insieme i vettori linearmente dipendenti finché non ci rimane un insieme di vettori linearmente indipendenti. Consideriamo la combinazione lineare

$$\alpha \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} + \beta \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix} + \gamma \begin{bmatrix} 0 \\ 0 \\ 4 \end{bmatrix} + \delta \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} = \underline{0}$$

$$\begin{cases} 2\alpha + 3\beta + 0\gamma + 1\delta = 0 \\ 1\alpha + 1\beta + 0\gamma + 1\delta = 0 \\ 3\alpha + 0\beta + 4\gamma + 2\delta = 0 \end{cases} \Rightarrow \begin{cases} 2\alpha + 3\beta + \delta = 0 \\ \alpha + \beta + \delta = 0 \\ 3\alpha + 4\gamma + 2\delta = 0 \end{cases} \Rightarrow \begin{cases} \delta = -2\alpha - 3\beta \\ \alpha + \beta - 2\alpha - 3\beta = 0 \\ 3\alpha + 4\gamma + 2(-2\alpha - 3\beta) = 0 \end{cases}$$

$$\begin{cases} \delta = -2\alpha - 3\beta \\ \alpha = -2\beta \\ -\alpha + 4\gamma - 6\beta = 0 \end{cases} \Rightarrow \begin{cases} \delta = \beta \\ \alpha = -2\beta \\ 4\gamma = 4\beta \end{cases} \Rightarrow \begin{cases} \delta = \beta \\ \alpha = -2\beta \\ \gamma = \beta \end{cases}$$

Poiché α, β e δ sono non nulli, i vettori sono tra loro linearmente dipendenti

$$-2\beta \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} + \beta \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix} + \beta \begin{bmatrix} 0 \\ 0 \\ 4 \end{bmatrix} + \beta \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} = \underline{0}$$

Per il teorema 1.1 possiamo affermare che il vettore $\mathbf{v}_4 = (1, 1, 2)$ può essere scritto come:

$$\beta \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} = 2\beta \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} - \beta \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix} - \beta \begin{bmatrix} 0 \\ 0 \\ 4 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} = 2 \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} - \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ 4 \end{bmatrix}$$

$\mathbf{v}_3 = (0, 0, 4)$ come:

$$\beta \begin{bmatrix} 0 \\ 0 \\ 4 \end{bmatrix} = 2\beta \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} - \beta \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix} - \beta \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} \Rightarrow \begin{bmatrix} 0 \\ 0 \\ 4 \end{bmatrix} = 2 \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} - \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$$

$\mathbf{v}_2 = (3, 1, 0)$ come:

$$\beta \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix} = 2\beta \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} - \beta \begin{bmatrix} 0 \\ 0 \\ 4 \end{bmatrix} - \beta \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} \Rightarrow \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix} = 2 \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ 4 \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$$

$\mathbf{v}_1 = (2, 1, 3)$ come:

$$-2\beta \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} = -\beta \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix} - \beta \begin{bmatrix} 0 \\ 0 \\ 4 \end{bmatrix} - \beta \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} \quad \Rightarrow \quad \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix} + \frac{1}{2} \begin{bmatrix} 0 \\ 0 \\ 4 \end{bmatrix} + \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$$

Quindi le possibili basi che possiamo ottenere da questo insieme di generatori sono $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$, $\mathcal{B}' = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_4\}$, $\mathcal{B}'' = \{\mathbf{v}_1, \mathbf{v}_3, \mathbf{v}_4\}$, $\mathcal{B}''' = \{\mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$

Teorema 1.5 (TEOREMA DI COMPLETAMENTO DELLA BASE). Sia V uno spazio vettoriale e $I = \{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_j\}$ un insieme di vettori linearmente indipendenti, allora è sempre possibile completare I a una base di V , detta in altre parole esiste una base \mathcal{B} di V con $I \subseteq \mathcal{B}$

Definizione 1.11. Possiamo dire che \mathcal{B} è **base canonica** per V se le coordinate di ogni vettore di V coincidono con le proprie componenti.

Nello spazio vettoriale \mathbb{R}^n i vettori:

$$\mathbf{e}_1 = (1, 0, 0, \dots, 0); \mathbf{e}_2 = (0, 1, 0, \dots, 0); \dots; \mathbf{e}_n = (0, 0, 0, \dots, 1)$$

formano una base che chiamiamo base canonica e viene indicata con \mathcal{E} .

ESEMPIO

Consideriamo i vettori $\mathbf{v}_1 = (2, 1, 0)$, $\mathbf{v}_2 = (1, 4, 0)$ vogliamo completare questo insieme di vettori linearmente indipendenti a una base di \mathbb{R}^3 , non verifichiamo che sono linearmente indipendenti poiché essendo soltanto due vettori, questi possono essere linearmente dipendenti soltanto se sono tra loro proporzionali o se uno di essi è il vettore nullo e si vede facilmente che non rientrano in uno di questi due casi. Si può apprezzare facilmente che ogni vettore nella forma $(k, k, 0)$ può essere generato da \mathbf{v}_1 e \mathbf{v}_2 ma non si è in grado di generare un solo vettore con una terza componente non nulla $(0, 0, k)$ per cui per rendere l'insieme di vettori una base quello che potremmo fare è aggiungere il vettore $\mathbf{e}_3 = (0, 0, 1)$ o un qualunque vettore con la terza componente non nulla, infatti sia un vettore generico (λ, μ, k) con $\lambda, \mu, k \in \mathbb{R}$ e $k \neq 0$ se consideriamo la combinazione lineare:

$$\alpha \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} + \beta \begin{bmatrix} 1 \\ 4 \\ 0 \end{bmatrix} + \gamma \begin{bmatrix} \lambda \\ \mu \\ k \end{bmatrix} = \underline{0}$$

che ci da il sistema di equazioni:

$$\begin{cases} 2\alpha + \beta + \lambda\gamma = 0 \\ \alpha + 4\beta + \mu\gamma = 0 \\ k\gamma = 0 \end{cases} \Rightarrow \begin{cases} 2\alpha + \beta = 0 \\ \alpha + 4\beta = 0 \\ \gamma = 0 \end{cases} \Rightarrow \begin{cases} 2(-4\beta) + \beta = 0 \\ \alpha = -4\beta \\ \gamma = 0 \end{cases} \Rightarrow \begin{cases} \beta = 0 \\ \alpha = 0 \\ \gamma = 0 \end{cases}$$

per cui abbiamo dimostrato che se $k \neq 0$ allora i tre vettori risulteranno linearmente indipendenti per qualunque valore di λ e μ e formeranno una base di \mathbb{R}^3 .

In altri casi completare la base non è così immediato ma risulta comunque facile, basta ricordarsi il significato della lineare dipendenza e prendere un qualunque vettore che non sia proporzionale a un vettore assegnato e che non si possa ottenere come somma degli altri vettori appartenenti all'insieme quindi si può sempre completare l'insieme di vettori

Definizione 1.12. Sia V uno spazio vettoriale e $U \subseteq V$. Se U è uno spazio vettoriale si dice che U è un **sottospazio vettoriale** di V . Se $U \subseteq V$ allora U è un sottospazio vettoriale di V se e solo se:

1. $\underline{0}_V \in U$
2. $\forall \mathbf{u}_1, \mathbf{u}_2 \in U, \forall \alpha, \beta \in \mathbb{R} \quad \text{si ha} \quad \alpha\mathbf{u}_1 + \beta\mathbf{u}_2 \in U$
(chiuso rispetto alla combinazione lineare)

Finora abbiamo detto che uno spazio vettoriale è un insieme ove è definita la somma e il prodotto per uno scalare. Pensiamo ora a un sottoinsieme di uno spazio vettoriale; per essere anch'esso uno spazio vettoriale deve essere definita anche su esso la somma e il prodotto per uno scalare. Molte delle proprietà necessarie sono automatiche proprio per il fatto di essere un sottoinsieme di uno spazio vettoriale, uno dei problemi che potrebbe incontrare è che il vettore nullo non sia nel sottoinsieme e di conseguenza la somma non sarebbe definita poiché non esisterebbe l'elemento neutro o potrebbe non essere chiusa rispetto alla combinazione lineare.

Un sottospazio è l'insieme di tutti i vettori che soddisfano la condizione di appartenenza ad esso, ad esempio un possibile sottospazio sarebbe $U = \{(x, y, z) \in \mathbb{R}^3 \mid x + y - z = 0\}$ che è formato da tutti i vettori appartenenti a \mathbb{R}^3 che soddisfano l'equazione data.

ESEMPIO

Per verificare che un sottoinsieme è realmente un sottospazio vettoriale dobbiamo verificare che valgono le due proprietà fondamentali, ossia che il vettore nullo appartiene al sottoinsieme e che è chiuso rispetto alla combinazione lineare, consideriamo il sottoinsieme $U = \{(x, y, z) \in \mathbb{R}^3 \mid x - 2y + 3z = 2\}$ vediamo immediatamente che il vettore nullo non appartiene al sottoinsieme $(0, 0, 0) \notin U$ poiché non soddisfa l'equazione di appartenenza $0 - 2 \cdot 0 + 3 \cdot 0 \neq 2$.

Consideriamo ora il sottoinsieme $U = \{(x, y, z) \in \mathbb{R}^3 \mid x + z \geq 0\}$ si vede facilmente che il sottoinsieme non è uno spazio vettoriale poiché non è chiuso rispetto alla combinazione lineare infatti consideriamo la combinazione lineare:

$$\alpha \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix} + \beta \begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix}$$

per essere chiusa deve verificarsi che per qualunque valore di α e β la combinazione lineare di due qualsivoglia vettori appartenenti al sottoinsieme dia un vettore appartenente al sottoinsieme, consideriamo due vettori a caso che soddisfano l'equazione, ad esempio $(1, 0, 0)$ e $(1, 0, 1)$ se consid-

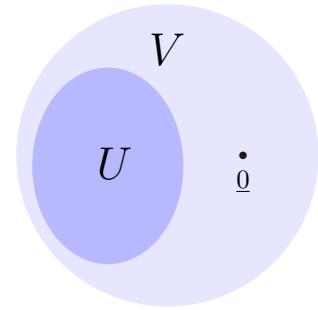


Figura 1.5: In questo caso U non è uno spazio vettoriale poiché $\underline{0}_V \notin U$

eriamo la combinazione lineare con $\alpha = 1$ e $\beta = -4$ si ha:

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} - 4 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix} \quad \text{che non soddisfa l'equazione } -3 + 1 \not\geq 0$$

Proprietà 1.2. Se V è uno spazio vettoriale, con $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n \in V$ allora $\langle \mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n \rangle$ (l'insieme generato dai vettori) è un sottospazio vettoriale di V .

ESEMPIO

Consideriamo i vettori $\mathbf{u}_1 = (2, 1, 3)$ e $\mathbf{u}_2 = (1, 1, 2)$ vogliamo trovare qual'è lo spazio vettoriale generato da questi due vettori, cioè vogliamo trovare l'equazione del sottospazio. Per trovarla ci basta considerare la combinazione lineare e ricavare le equazioni o equazioni.

$$\alpha \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} + \beta \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} \Rightarrow \begin{bmatrix} 2\alpha + \beta \\ \alpha + \beta \\ 3\alpha + 2\beta \end{bmatrix}$$

Se un vettore generico $\in \mathbb{R}^3$ lo pensiamo con componenti (x, y, z) allora possiamo scrivere il seguente sistema di equazioni:

$$\begin{cases} x = 2\alpha + \beta \\ y = \alpha + \beta \\ z = 3\alpha + 2\beta \end{cases} \Rightarrow \begin{cases} x - 2\alpha = \beta \\ y = \alpha + x - 2\alpha \\ z = 3\alpha + 2(x - 2\alpha) \end{cases} \Rightarrow \begin{cases} x - 2\alpha = \beta \\ \alpha = x - y \\ z = -\alpha + 2x \end{cases} \Rightarrow \begin{cases} x - 2\alpha = \beta \\ \alpha = x - y \\ z = -x + y + 2x \end{cases}$$

Quindi l'equazione del sottospazio è proprio $x + y - z = 0$

ESEMPIO

Vogliamo vedere ora come si può ricavare una base del sottospazio data l'equazione, consideriamo sempre il sottospazio $U = \{(x, y, z) \in \mathbb{R}^3 \mid x + y - z = 0\}$ da questa equazione possiamo scrivere: $z = x + y$, usando questa equazione possiamo scrivere il generico vettore appartenente al sottospazio come:

$$\begin{bmatrix} x \\ y \\ x + y \end{bmatrix} \quad \text{che può essere visto come:} \quad \begin{bmatrix} x \\ 0 \\ x \end{bmatrix} + \begin{bmatrix} 0 \\ y \\ y \end{bmatrix} \Rightarrow x \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + y \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$

Quindi una base del sottospazio vettoriale è $\{(1, 0, 1), (0, 1, 1)\}$ Non ci deve sorprendere il fatto che gli elementi della base siano diversi dai vettori considerati nell'esempio precedente, poiché come abbiamo visto in precedenza, possono esserci vettori diversi che generano lo stesso spazio.

Proprietà 1.3. Sia V uno spazio vettoriale e siano $\mathbf{v}_1, \mathbf{v}_2$ sottospazi vettoriali di V allora $\mathbf{v}_1 + \mathbf{v}_2$ e $\mathbf{v}_1 \cap \mathbf{v}_2$ sono sottospazi vettoriali di V mentre $\mathbf{v}_1 \cup \mathbf{v}_2$ in generale non lo è.

Se consideriamo V_1 un sottospazio vettoriale generato da $\mathbf{e}_1 \in \mathbb{R}^2$ ossia $\langle (1, 0) \rangle$ e V_2 un altro sottospazio vettoriale generato da $\mathbf{e}_2 \in \mathbb{R}^2$ cioè $\langle (0, 1) \rangle$ verifichiamo facilmente che $V_1 \cup V_2$ non è uno spazio vettoriale poiché non è chiuso rispetto alla combinazione lineare infatti se consideriamo la combinazione lineare con $\alpha = 1$ e $\beta = 1$ si ha $(1, 0) + (0, 1) = (1, 1) \notin V_1 \cup V_2$

Definizione 1.13. A questo punto ci occorre definire quello che chiamiamo spazio somma

$$V_1 + V_2 = \{\mathbf{v} \in V \text{ t.c. } \mathbf{v} = \mathbf{s} + \mathbf{t} \text{ dove } \mathbf{s} \in V_1, \mathbf{t} \in V_2\}$$

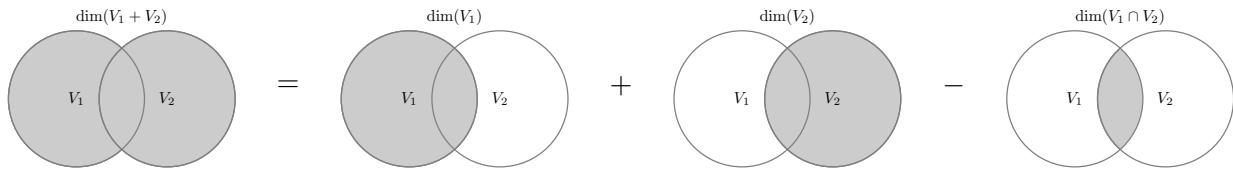
Definiamo poi l'intersezione dei due sottospazi vettoriali, e la indichiamo con $V_1 \cap V_2$.

$$V_1 \cap V_2 = \{\mathbf{v} \in V \text{ t.c. } \mathbf{v} \in V_1, \mathbf{v} \in V_2\}$$

Teorema 1.6 (FORMULA DI GRASSMANN). Siano V_1 e V_2 sottospazi vettoriali di V di dimensioni finite. Allora:

$$\dim(V_1 + V_2) = \dim(V_1) + \dim(V_2) - \dim(V_1 \cap V_2)$$

Giacché la base di $V_1 + V_2$ è l'unione delle singole basi risulta più facile memorizzare la formula ragionando con i diagrammi di Eulero-Venn.



Vale la pena notare che l'intersezione fra due sottospazi non può essere mai nulla poiché entrambi insiemi avranno sempre il vettore nullo in comune (altrimenti non sarebbero sottospazi vettoriali) però la dimensione del solo vettore nullo è nulla poiché, come abbiamo detto non ha una base. Se $V_1 \cap V_2 = \{\underline{0}_V\}$ l'intersezione è dunque minima e lo spazio somma si dice che è in **somma diretta** e si indica con $V_1 \oplus V_2$ e vale :

$$\dim(V_1 \oplus V_2) = \dim(V_1) + \dim(V_2)$$

Abbiamo già visto come si fa a trovare una base di un sottospazio vettoriale e di conseguenza come si trova la dimensione del sottospazio poiché ricordiamo che la dimensione non è altro che il numero di vettori che formano una sua base. Trovare però la base dell'intersezione fra due sottospazi richiede qualche considerazione in più.

ESEMPIO

Consideriamo $U = \{(x, y, z) \in \mathbb{R}^3 \mid x + y - z = 0\}$ e $W = \{(x, y, z) \in \mathbb{R}^3 \mid 2x - y = 0\}$ Vogliamo trovare base e dimensione di ogni singolo sottospazio così come per l'intersezione e lo spazio somma. Quello che ci conviene fare per prima è trovare le basi dei singoli sottospazi.

Dagli esempi precedenti sappiamo che una base di U è $\{(1, 0, 1), (0, 1, 1)\}$ per cui $\dim(U) = 2$, ora per trovare una base di W sfruttiamo il fatto che il generico vettore appartenente a W si scrive come:

$$\begin{bmatrix} x \\ 2x \\ z \end{bmatrix} \quad \text{che può essere visto come:} \quad \begin{bmatrix} x \\ 2x \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ z \end{bmatrix} \Rightarrow x \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} + z \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

quindi una base è $\{(1, 2, 0) (0, 0, 1)\}$ e $\dim(W) = 2$.

Per trovare una base dell'intersezione quello che dobbiamo fare è sostituire le componenti del generico vettore di un sottospazio vettoriale nell'equazione dell'altro, ad esempio possiamo sostituire il generico vettore di W nell'equazione di U , cioè, sostituendo nell'equazione $x + y - z = 0$ la componente $x = x, y = 2x, z = z$ che da luogo all'equazione $x + 2x - z = 0$ ciò si verifica solo se $z = 3x$, sostituendo nel generico vettore di cui abbiamo preso le componenti si ottiene che il vettore appartenente all'intersezione ha componenti $(x, 2x, 3x)$ quindi una base dell'intersezione è $(1, 2, 3)$ e $\dim(U \cap W) = 1$.

un metodo alternativo è sostituire le componenti $(x, y, x+y)$ del generico vettore di U nell'equazione $2x - y = 0$ di W , così facendo si otteneva l'equazione $2x - y = 0$ che ha come soluzione $y = 2x$ imponendo la soluzione trovata nel generico vettore di U si ha un vettore di componenti $(x, 2x, 3x)$ che è proprio la stessa soluzione che avevamo trovato precedentemente, per fare un ultimo controllo quello che facciamo è controllare che la base dell'intersezione soddisfi entrambe le equazioni, infatti tutti i vettori $(k, 2k, 3k)$ verificano entrambe le equazioni $2(k) - 2k = 0$ e $k + 2k - 3k = 0$.

A questo punto, per trovare una base dello spazio somma $U + W$ consideriamo l'unione delle due basi cioè $B_U \cup B_W$ poiché questo insieme è un sistema di generatori per lo spazio somma, grazie al teorema di esistenza delle basi possiamo trovare una base dello spazio somma togliendo al insieme tutti quei vettori che sono linearmente dipendenti. Cioè basta trovare tutti i vettori linearmente indipendenti dell'insieme $B_U \cup B_W = \{(1, 0, 1) (0, 1, 1) (1, 2, 0) (0, 0, 1)\}$ a questo punto dovremmo considerare la combinazione lineare ma si vede facilmente che tre qualsiasi vettori di quel insieme formano una base per lo spazio somma per cui una base di $B_{U+W} = \{(1, 0, 1) (0, 1, 1) (1, 2, 0)\}$ e $\dim(V + W) = 3$, risultato coerente con la formula di Grassmann.

Definizione 1.14. Sia V uno spazio vettoriale con $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n \in V$ si chiama **rango** di $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ la dimensione di $\langle \mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n \rangle$ cioè, chiamiamo rango il numero massimo di vettori linearmente indipendenti.

Indicato anche come (rk è dovuto al termine inglese rank):

$$\text{rk } \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$$

Il rango gode di certe proprietà:

1. Si possono riordinare i vettori senza cambiare il risultato

$$\text{rk } \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\} = \text{rk } \{\mathbf{v}_2, \mathbf{v}_1, \dots, \mathbf{v}_n\}$$

2. Si può inserire un multiplo di un vettore senza cambiare il risultato

$$\text{rk } \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\} = \text{rk } \{\lambda \mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$$

In molti casi risulta importante conoscere soltanto la dimensione di ognuno dei sottospazi, ed è molto pratico invece di calcolare la base dello spazio somma calcolarne soltanto il rango che in questi casi è proprio la dimensione dello spazio somma. Per fare ciò dovremmo considerare la combinazione lineare di tutti i vettori e togliere progressivamente i vettori linearmente dipendenti; questo algoritmo è senz'altro efficace ma non molto efficiente poiché avremmo un sistema di tante equazioni quanti sono i vettori generatori.

Nel prossimo capitolo vedremo come l'uso opportuno di strumenti come le matrici ci aiutano a semplificare sostanzialmente la risoluzione di questi sistemi di equazioni e ci aiutano a calcolare il rango in un modo molto più efficiente.

Matrici

Una matrice $m \times n$ a coefficienti in \mathbb{R} è una tabella con m righe ed n colonne che contiene come elementi dei numeri reali e si indicano con la notazione $\mathbb{R}^{m,n}$. L'insieme di tutte le matrici $m \times n$ si indica con la seguente notazione: $\mathcal{M}(m, n, \mathbb{R})$

Generalmente una generica matrice si indica con una lettera maiuscola e viene scritta in questo modo:

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

I pedici di ogni lettera non sono stati scelti a caso ma hanno un significato ben preciso, il primo numeretto indica la riga in cui si trova l'elemento, il secondo ne indica invece la colonna.

Definizione 2.1. Sia A una matrice, se $A \in \mathcal{M}(m, n, \mathbb{R})$ si dice che la matrice A ha **ordine** $m \times n$

Proprietà 2.1 (Insieme di matrici come spazi vettoriali). $\forall m, n \in \mathbb{N}^+$, la struttura algebrica $(\mathbb{R}^{m,n}, +, \cdot)$ è uno spazio vettoriale.

Grazie a questa proprietà possiamo associare a ciascun vettore una matrice riga o una colonna oppure possiamo radunare un insieme di vettori in una sola matrice che può essere rettangolare o quadrata. Vediamo dunque quali sono i tipi di matrici.

2.1 Tipi di matrici

2.1.1 Matrice riga

Come lo indica il suo nome, è una matrice di tipo $(1, n)$, ossia una matrice formata da una sola riga.

ESEMPIO $[0 \ 1 \ 2 \ 3] \in \mathbb{R}^{1,4}$

2.1.2 Matrice colonna

È una matrice formata da una sola colonna, ossia del tipo $(m, 1)$.

ESEMPIO $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \in \mathbb{R}^{3,1}$

2.1.3 Matrice rettangolare

È una matrice in cui il numero delle righe è diverso dal numero delle colonne, ovvero $m \neq n$ con $m, n \neq 1$.

ESEMPIO $\begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{bmatrix} \in \mathbb{R}^{2,4}$

2.1.4 Matrice quadrata

È una matrice che ha il numero di righe uguale al numero di colonne. Tali matrici rivestono un ruolo fondamentale in Algebra Lineare. Sono infatti le uniche per cui si può calcolare quello che chiameremo più avanti il determinante e cercare la matrice inversa (se esiste).

ESEMPIO $\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \in \mathbb{R}^{3,3}$

Al posto di scrivere $\mathcal{M}(n, n, \mathbb{R})$ si pone soltanto $\mathcal{M}_n(\mathbb{R})$ per le matrici quadrate.

2.1.5 Matrice triangolare alta

È una matrice quadrata avente tutti gli elementi al di sotto della diagonale principale uguali a zero. $A \in \mathbb{R}^{n,n}$ è triangolare superiore se e solo se $\forall i, j \in \{1, 2, \dots, n\}, i > j : a_{ij} = 0$

ESEMPIO $\begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{bmatrix} \in \mathbb{R}^{3,3}$

2.1.6 Matrice triangolare bassa

È una matrice quadrata avente tutti gli elementi al di sopra della diagonale principale uguali a zero. $A \in \mathbb{R}^{n,n}$ è triangolare superiore se e solo se $\forall i, j \in \{1, 2, \dots, n\}, i < j : a_{ij} = 0$

ESEMPIO $\begin{bmatrix} 1 & 0 & 0 \\ 2 & 3 & 0 \\ 4 & 5 & 6 \end{bmatrix} \in \mathbb{R}^{3,3}$

2.1.7 Matrice trasposta

Si indica generalmente con A^T pero per non confondere con l'elevamento a potenza si indica ${}^t A$ e si ottiene scambiando le righe di A con le sue colonne. Ovvero $a_{ij} = b_{ji}$

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \in \mathbb{R}^{3,3} \quad {}^t A = \begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix} \in \mathbb{R}^{3,3}$$

Teorema 2.1. *Sia $A = \mathcal{M}(m, n, \mathbb{R})$, $B = \mathcal{M}(n, l, \mathbb{R})$ allora:*

$$\underbrace{{}^t(A \cdot B)}_{l \times m} = \underbrace{{}^t(B)}_{l \times n} \cdot \underbrace{{}^t(A)}_{n \times m}$$

2.1.8 Matrice diagonale

È una matrice quadrata in cui i soli termini della diagonale principale possono essere diversi da zero. In formule possiamo dire che $A \in \mathbb{R}^{n,n}$ è diagonale se e solo se $\forall i, j \in \{1, 2, \dots, n\}, i \neq j : a_{ij} = 0$

ESEMPIO $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} \in \mathbb{R}^{3,3}$

2.1.9 Matrice simmetrica

Si dice simmetrica se $A = {}^t A$ ovvero $a_{ij} = a_{ji}$

ESEMPIO $\begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{bmatrix} \in \mathbb{R}^{3,3}$

La dimensione del insieme S_n delle matrice simmetriche $(n \times n)$ è: $\dim S_n = \frac{n(n+1)}{2}$

2.1.10 Matrice antisimmetrica

Si dice antisimmetrica o **emisimmetrica** se $A = -{}^t A$ ovvero $a_{ij} = -a_{ji}$

ESEMPIO $\begin{bmatrix} 0 & 1 & 2 \\ -1 & 0 & 3 \\ -2 & -3 & 0 \end{bmatrix} \in \mathbb{R}^{3,3}$

Si noti che la diagonale principale è sempre nulla poiché $a_{ii} = -a_{ii}$ se soltanto se $a_{ii} = 0$

La dimensione del insieme delle matrice antisimmetriche: $(n \times n)$ è: $\dim \mathcal{A}_n = \frac{n(n-1)}{2}$

Teorema 2.2. *Ogni matrice quadrata può essere scritta come la somma di una matrice simmetrica e una antisimmetrica.*

$$\mathcal{M}_n = \mathcal{S}_n \oplus \mathcal{A}_n \quad A = \underbrace{\frac{A + {}^t A}{2}}_{\text{simmetrica}} + \underbrace{\frac{A - {}^t A}{2}}_{\text{antisimmetrica}}$$

ESEMPIO

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} = \begin{bmatrix} 1 & 3 & 5 \\ 3 & 5 & 7 \\ 5 & 7 & 9 \end{bmatrix} + \begin{bmatrix} 0 & -1 & -2 \\ 1 & 0 & -1 \\ 2 & 1 & 0 \end{bmatrix}$$

2.1.11 Matrice opposta

Data una qualsiasi matrice A la sua opposta (che è sempre definita) si indica con $-A$ ed è la matrice che si ottiene cambiando di segno tutti gli elementi della matrice A . L'opposta di una matrice A si definisce come $-A = -1 \cdot A$

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \in \mathbb{R}^{3,3} \quad -A = \begin{bmatrix} -1 & -2 & -3 \\ -4 & -5 & -6 \\ -7 & -8 & -9 \end{bmatrix} \in \mathbb{R}^{3,3}$$

2.1.12 Matrice nulla

È la matrice costituita da coefficienti tutti zeri. Si indica con O .

ESEMPIO $O = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \in \mathbb{R}^{3,3}$

2.1.13 Matrice identità

Si indica generalmente con I ed è una matrice diagonale - per cui quadrata -. avente tutti uno sulla diagonale principale e tutti zero altrove. In simboli:

$$I \in \mathbb{R}^{n,n}, \forall i, j \in \{1, 2, \dots, n\}$$

Le sue componenti sono date dal **delta di Kronecker** :

$$\delta_{ij} := \begin{cases} 1 & \text{se } i = j \\ 0 & \text{se } i \neq j \end{cases}$$

ESEMPIO $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \in \mathbb{R}^{4,4}$

2.1.14 Matrice a gradini o a scala

La matrice si dice a gradini (o a scalini) *echelon* in inglese, se per ogni coppia di righe consecutive il primo coefficiente non nullo della seconda è più a destra del primo coefficiente non nullo della precedente. Ogni matrice quadrata a scala è anche triangolare alta.

ESEMPIO $A = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 5 & 6 & 7 & 8 \\ 0 & 0 & 9 & 10 & 11 \\ 0 & 0 & 0 & 0 & 12 \end{bmatrix}$

Definizione 2.2. Chiamiamo **pivot** il primo coefficiente non nullo di ogni riga.

Più precisamente una matrice si dice a gradini (o a scalini) se valgono i seguenti due fatti:

1. Il pivot di ogni riga non nulla è più a sinistra dei pivot delle righe non nulle successive.
2. Le righe nulle sono tutte alla fine.

Proprietà 2.2. Se una matrice non è a gradini si possono fare le **operazioni elementari su righe** affinché diventi a scala. Queste operazioni consistono nello scambio di righe, la somma e sottrazione fra righe e la moltiplicazione di una riga per un coefficiente. Questo argomento verrà approfondito ulteriormente.

2.2 Operazioni con matrici

2.2.1 Somma

Consideriamo due matrici A e B i cui elementi sono numeri reali. Le due matrici sono sommabili se e solo se sono dello stesso tipo, cioè se e solo se hanno lo stesso numero di righe e di colonne, quindi la stessa dimensione. Tale somma si indica con : $A+B$. Cioè se $A, B \in \mathbb{R}^{m,n}$, allora $A+B \in \mathbb{R}^{m,n}$. Siano:

$$A = \begin{bmatrix} 2 & 5 & -3 \\ 1 & -2 & 4 \end{bmatrix} \quad B = \begin{bmatrix} 7 & -5 & 2 \\ -9 & 4 & -1 \end{bmatrix}$$

$$A + B = \begin{bmatrix} 2+7 & 5+(-5) & -3+2 \\ 1+(-9) & -2+4 & 4+(-1) \end{bmatrix} = \begin{bmatrix} 9 & 0 & -1 \\ -8 & 2 & 3 \end{bmatrix}$$

La sottrazione funziona allo stesso modo.

2.2.2 Prodotto per uno scalare

Chiamiamo prodotto di una matrice per uno scalare il prodotto della matrice A per lo scalare λ , il risultato è una nuova matrice con lo stesso numero di righe e di colonne con coefficienti che si ottengono moltiplicando tutti gli elementi di A per lo scalare λ . Ad esempio: $A \in \mathcal{M}(3, 2, \mathbb{R})$.

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \quad e \quad \lambda = 2 \quad \lambda \cdot A = \lambda \cdot \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} = \begin{bmatrix} 2 \cdot 1 & 2 \cdot 2 \\ 2 \cdot 3 & 2 \cdot 4 \\ 2 \cdot 5 & 2 \cdot 6 \end{bmatrix} = \begin{bmatrix} 2 & 4 \\ 6 & 8 \\ 10 & 12 \end{bmatrix}$$

2.2.3 Prodotto righe per colonne

Sia $A \in \mathbb{R}^{m,n}$ $B \in \mathbb{R}^{n,l}$ cioè la prima matrice deve avere lo stesso numero di colonne che la seconda righe, allora si può definire $A \cdot B = C$ con $C \in \mathbb{R}^{m,l}$. Se consideriamo $1 \leq i \leq m$, $1 \leq k \leq n$, $1 \leq j \leq l$. Allora $c_{ij} = \sum_{k=1}^n a_{ik} \cdot b_{kj}$

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 5 & -1 \\ 3 & 2 & 0 \end{bmatrix} \quad B = \begin{bmatrix} 7 & 1 \\ 1 & 0 \\ 0 & 4 \end{bmatrix}$$

$$A \cdot B = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 5 & -1 \\ 3 & 2 & 0 \end{bmatrix} \cdot \begin{bmatrix} 7 & 1 \\ 1 & 0 \\ 0 & 4 \end{bmatrix} = \begin{bmatrix} 1 \cdot 7 + 0 \cdot 1 + 1 \cdot 0 & 1 \cdot 1 + 0 \cdot 0 + 1 \cdot 4 \\ 1 \cdot 7 + 5 \cdot 1 + (-1) \cdot 0 & 1 \cdot 1 + 5 \cdot 0 + (-1) \cdot 4 \\ 3 \cdot 7 + 2 \cdot 1 + 0 \cdot 0 & 3 \cdot 1 + 2 \cdot 0 + 0 \cdot 4 \end{bmatrix} = \begin{bmatrix} 7 & 5 \\ 12 & -3 \\ 23 & 3 \end{bmatrix}$$

Il prodotto righe per colonne equivale a fare il prodotto di ogni elemento della riga della prima matrice per il rispettivo elemento della colonna nella seconda matrice iterando le colonne e successivamente la riga.

Date tre matrici, nell'ipotesi in cui il prodotto sia eseguibile, ovvero siano A, B, C così definite: $A = \mathcal{M}(m, n, \mathbb{R})$, $B = \mathcal{M}(n, l, \mathbb{R})$, $C = \mathcal{M}(l, p, \mathbb{R})$ avremo le seguenti proprietà:

1. Il prodotto tra due matrici è associativo

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

2. Distributività del prodotto tra matrici rispetto alla somma

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

3. Elemento neutro rispetto al prodotto tra matrici

In questo caso l'elemento neutro non è un numero ma una matrice, e questa matrice è la matrice identità I

$$A \cdot I = A$$

4. Prodotto con la matrice nulla

Moltiplicando qualsiasi matrice per la matrice nulla, costituita da soli zeri, si ottiene sempre la matrice nulla.

$$A \cdot O = O$$

È importante notare che il prodotto non è commutativo ovvero può essere definito $(A \cdot B)$ ma no $(B \cdot A)$, nel caso siano entrambe definite si ha in generale $A \cdot B \neq B \cdot A$.

Definizione 2.3. Se $A \cdot B = B \cdot A$ si dice che A e B sono **permutabili**.

Inoltre non vale la proprietà di annullamento del prodotto di matrici – che dice che $A \cdot B = O$ se e solo se $A = O$, $B \neq O \vee B = O$, $A \neq O$ – dato che non vale, il prodotto $A \cdot B$ può essere uguale alla matrice nulla O anche se né A né B sono nulle.

$$\text{ESEMPIO} \quad A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \quad B = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$$

$$A \cdot B = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \quad B \cdot A = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$$

Definizione 2.4. La struttura $(\mathcal{M}_n(\mathbb{R}), +, \cdot)$ si chiama **anello** delle matrici quadrate.

2.2.4 Elevamento a potenza

Sia $A \in \mathcal{M}_n(\mathbb{R})$ e $k \in \mathbb{N}$, allora definiamo l'elevamento a potenza come segue:

$$A^k = \begin{cases} \underbrace{A \cdot A \cdots A}_{k \text{ volte}} & k \geq 2 \\ A & k = 1 \\ I & k = 0 \end{cases}$$

Diciamo che una matrice è **idempotente** se:

$$A^k = A \quad (\text{uguale a se stessa})$$

Diciamo che è invece **nilpotente** se

$$A^k = O \quad (\text{uguale alla matrice nulla})$$

dove k è rispettivamente l'indice di idempotenza o nilpotenza , diciamo inoltre che la matrice ha nilpotenza o idempotenza di grado k .

ESEMPIO

La matrice M è idempotente:

$$M = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

In realtà tutte le matrici triangolari alte con 0 come coefficienti nella diagonale principale sono sempre nilpotenti.

ESEMPIO

La matrice N è nilpotente:

$$N = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$$

Sistemi Lineari

Dati $a_{11}, \dots, a_{1n}, a_{2n}, \dots, a_{m1}, \dots, a_{mn} \in \mathbb{R}$ detti **coefficientsi** e $b_1, b_2, \dots, b_m \in \mathbb{R}$ chiamati **termini noti**, chiamiamo **sistema lineare** un sistema composto di m equazioni lineari (equazioni di primo grado) con n incognite x_1, x_2, \dots, x_n , ossia:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \vdots \quad \vdots \quad \vdots \quad = \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m \end{cases} \quad (3.1)$$

Definizione 3.1. Una **soluzione** del sistema è una n -pla di \mathbb{R} che, assegnati i valori alle incognite x_1, x_2, \dots, x_n , soddisfa simultaneamente tutte le equazioni del sistema.

Definizione 3.2. Due sistemi si dicono **equivalenti** quando ammettono le stesse soluzioni.

I sistemi lineari possono avere almeno una soluzione o nessuna soluzione, se ammette al meno una soluzione il sistema si dice **possibile** o **risolubile**, inoltre se ammette un'unica soluzione si dice **sistema determinato** e se ne ammette infinite si dice **sistema indeterminato**, nel caso non abbia soluzioni si dice **sistema impossibile**.

I sistemi lineari possono essere risolti con il metodo di sostituzione che tuttavia, risulta molto inefficiente quando si è davanti a un sistema con molte equazioni, motivo per il cui utilizzeremo le matrici che sono molto più efficienti e poiché gli si può applicare un algoritmo ben definito possono essere implementate facilmente a calcolatore. Consideriamo il sistema lineare 3.1, per risolverlo con le matrici dobbiamo prima riscrivere come una matrice per cui consideriamo A la matrice dei coefficienti delle incognite, x è la colonna delle incognite e b la colonna dei termini noti, cioè:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

Il sistema si può allora rappresentare così:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

Questo si può dunque scrivere con l'equazione matriciale :

$$A \cdot \mathbf{x} = \mathbf{b}$$

In realtà tutto il sistema può essere scritto con una sola matrice, detta **matrice associata al sistema**:

$$(A|b) = \left[\begin{array}{ccc|c} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & \ddots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{array} \right]$$

Essa è ottenuta dalla giustapposizione della matrice A dei coefficienti e della colonna b , detta colonna dei termini noti. Dunque d'ora in poi chiameremo **matrice incompleta** la matrice A e **matrice completa o matrice orlata** la matrice $[A|b]$.

Tutta questa formalizzazione ci serve per determinare se il sistema ha soluzioni, e quante ne ha, anche senza risolverlo esplicitamente ma con il solo studio della matrice $(A|b)$.

3.1 Algoritmo di eliminazione di Gauss

Per determinare se il sistema ha soluzioni, e quante ne ha ci occorre una matrice $(A|b)$ a gradini, e come era già stato accennato, questa si può ottenere da una matrice non a gradini con le operazioni elementari su righe o **mosse di Gauss**:

1. sommare a una riga un multiplo di una riga precedente
2. scambiare fra loro due righe
3. moltiplicare una riga per un numero diverso da zero

Queste mosse hanno l'importante proprietà che se sono applicate alla matrice completa di un sistema lineare $(A|b)$, queste non modificano lo spazio delle soluzioni del sistema, in altre parole, cambiano il sistema ma le soluzioni restano invariate (sistema equivalente). In generale possiamo formulare un algoritmo che ci permetta di ottenere la matrice a scalini con mosse elementari, questo algoritmo prende il nome di **riduzione gaussiana o metodo di eliminazione di Gauss** a volte abbreviato MEG e funziona come segue:

1. Se la prima riga ha il primo elemento nullo, scambiala con una riga che ha il primo elemento non nullo. Se tutte le righe hanno il primo elemento nullo, vai al punto 3.
2. Per ogni riga a_i con primo elemento non nullo, eccetto la prima ($i > 1$), moltiplica la prima riga per un coefficiente scelto in maniera tale che la somma tra la prima riga e a_i abbia il primo elemento nullo (quindi coefficiente $= -a_{i1}/a_{11}$). Sostitisci a_i con la somma appena ricavata.
3. Adesso sulla prima colonna tutte le cifre, eccetto la prima, sono nulle. A questo punto ritorna al punto 1 considerando la **sottomatrice** che ottieni cancellando la prima riga e la prima colonna.

Si mostra qui un esempio:

$$\begin{array}{c}
 \left[\begin{array}{ccc|c} 3 & 2 & -1 & 0 \\ -1 & 2 & 3 & 8 \\ 9 & -3 & -6 & 3 \end{array} \right] \xrightarrow{\frac{3R_2}{2}} \left[\begin{array}{ccc|c} 3 & 2 & -1 & 0 \\ -3 & 6 & 9 & 24 \\ 3 & -1 & -2 & 1 \end{array} \right] \xrightarrow{\frac{R_3-R_1}{2}} \left[\begin{array}{ccc|c} 3 & 2 & -1 & 0 \\ 0 & 8 & 8 & 24 \\ 0 & -3 & -1 & 1 \end{array} \right] \\
 \\
 \xrightarrow{\frac{8R_3}{3R_2}} \left[\begin{array}{ccc|c} 3 & 2 & -1 & 0 \\ 0 & 24 & 24 & 72 \\ 0 & -24 & -8 & 8 \end{array} \right] \xrightarrow{R_3+R_2} \left[\begin{array}{ccc|c} 3 & 2 & -1 & 0 \\ 0 & 24 & 24 & 72 \\ 0 & 0 & 16 & 80 \end{array} \right]
 \end{array}$$

Il risultato dell'algoritmo non è sempre lo stesso, dipende dalle scelte effettuate, ma quello che hanno tutte le matrice a gradini ottenute da una stessa matrice completa in comune sono i numeri di pivot che si trovano già, sia nella forma minima sia in un multiplo.

Supponendo che le quattro colonne si riferiscano a x, y, z e l'ultimo sia quella dei termini noti b possiamo ottenere il sistema con soluzioni ricavabili in pochi passaggi:

$$\left\{ \begin{array}{l} 3x + 2y - z = 0 \\ 24y + 24z = 72 \\ 16z = 80 \end{array} \right. \Rightarrow \left\{ \begin{array}{l} x = 3 \\ y = -2 \\ z = 5 \end{array} \right.$$

Una volta che è già stata raggiunta la forma a gradini possiamo proseguire con ulteriori mosse elementari, questa volta cercando di trasformarla in una matrice identità, questo algoritmo esteso viene detto **metodo di eliminazione di Gauss-Jordan** e dice:

1. Se i numeri di pivot sono tutti uno, vai al passo 2, altrimenti moltiplica la riga per l'inverso del pivot affinché diventi uno.
2. Posizionati sopra la riga più in basso che abbia solo il termine noto e l'uno come pivot.
3. Per eliminare i coefficienti della riga, somma un multiplo della riga sottostante affinché diventi zero, ripeti il passo 2 finché non si ha la matrice identità.

Nell'esempio precedente:

$$\begin{array}{c}
 \left[\begin{array}{ccc|c} 3 & 2 & -1 & 0 \\ 0 & 24 & 24 & 72 \\ 0 & 0 & 16 & 80 \end{array} \right] \xrightarrow{\frac{1}{24}R_2} \left[\begin{array}{ccc|c} 3 & 2 & -1 & 0 \\ 0 & 1 & 1 & 3 \\ 0 & 0 & 1 & 5 \end{array} \right] \xrightarrow{\frac{1}{3}R_1} \left[\begin{array}{ccc|c} 1 & \frac{2}{3} & \frac{-1}{3} & 0 \\ 0 & 1 & 1 & 3 \\ 0 & 0 & 1 & 5 \end{array} \right] \\
 \\
 \xrightarrow{\frac{R_2-R_3}{R_1+\frac{1}{3}R_3}} \left[\begin{array}{ccc|c} 1 & \frac{2}{3} & 0 & \frac{5}{3} \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 5 \end{array} \right] \xrightarrow{\frac{R_1-\frac{2}{3}R_2}{R_3}} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 3 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 5 \end{array} \right]
 \end{array}$$

Come si può osservare la colonna dei termini noti contiene le soluzioni del sistema.

Definizione 3.3. Sia $A \in \mathcal{M}_{m,n}(\mathbb{R})$ si definisce rango di A il massimo numero di vettori riga linearmente indipendenti tra loro, o equivalentemente il massimo numero di vettori colonna linearmente indipendenti. Chiameremo tali numeri, rispettivamente, rango per righe e rango per colonne. Il rango di una matrice è minore o uguale sia di m (il numero di righe) che di n (numero di colonne) in formule $\text{rk}(A) \leq \min\{m, n\}$. Si indica con $\text{r}(A)$ oppure $\text{rk}(A)$.

$$(A|b) = \left[\begin{array}{ccc|c} 1 & 2 & 4 & 8 \\ 0 & 7 & 2 & -4 \\ 0 & 0 & 9 & 3 \\ 0 & 0 & 0 & 0 \end{array} \right] \quad \text{rk}(A|b) = 3$$

Il rango può essere anche definito come il numero di pivot che compaiono nella matrice ridotta con l'eliminazione di Gauss. Il concetto di rango è molto importante nell'algebra lineare poiché in un certo senso rappresenta la quantità d'informazione contenuta nella matrice.

Definizione 3.4. Chiamo $\mathcal{R}(A)$, **spazio riga**, lo spazio generato dalle righe di A .

Definizione 3.5. Chiamo $\mathcal{C}(A)$, **spazio colonna**, lo spazio generato dalle colonne di A .

Teorema 3.1.

$$\dim \mathcal{R}(A) = \dim \mathcal{C}(A) = \text{rk}(A)$$

Questo teorema ci permette di chiamare indifferentemente il rango per righe e il rango per colonne, dato che coincidono, lo chiameremo d'ora in poi semplicemente rango.

È importante notare che anche se la dimensione dello spazio riga e dello spazio colonna è lo stesso, gli spazi generati sono ben diversi!

Teorema 3.2. Il sistema NON HA SOLUZIONI se e solo se la matrice a gradini ottenuta da la riduzione di Gauss ha un pivot nell'ultima colonna.

$$eg. \quad (A|b) = \left[\begin{array}{ccc|c} 1 & 2 & 4 & 8 \\ 0 & 7 & 2 & -4 \\ 0 & 0 & 9 & 3 \\ 0 & 0 & 0 & 2 \end{array} \right] \quad \text{sistema impossibile}$$

Questo teorema può essere anche espresso come:

Teorema 3.3. Il sistema NON HA SOLUZIONI se e solo se:

$$\text{rk}(A) \neq \text{rk}(A|b)$$

In effetti:

$$(A|b) = \left[\begin{array}{ccc|c} 1 & 2 & 4 & 8 \\ 0 & 7 & 2 & -4 \\ 0 & 0 & 9 & 3 \\ 0 & 0 & 0 & 2 \end{array} \right] \quad \text{rk}(A|b) = 4 \quad A = \left[\begin{array}{ccc} 1 & 2 & 4 \\ 0 & 7 & 2 \\ 0 & 0 & 9 \\ 0 & 0 & 0 \end{array} \right] \quad \text{rk}(A) = 3$$

Teorema 3.4 (Teorema di Rouché-Capelli o Rouché-Frobenius). Il sistema lineare $A \cdot x = b$ con n incognite ha soluzioni se e solo se:

$$\text{rk}(A) = \text{rk}(A|b) \quad (3.2)$$

Un sistema con t variabili libere o **variabili indipendenti** ha:

$$\infty^t \text{ soluzioni} \quad \text{ove} \quad t = n - \text{rk}$$

Corollario 3.4.1. Un sistema lineare dipende da t variabili libere quando $n - \text{rk} > 0$, che significa che il sistema ha più incognite che equazioni linearmente indipendenti.

Corollario 3.4.2. Un sistema lineare ha una sola soluzione se il numero di incognite è uguale al numero di equazioni linearmente indipendenti, ossia: $n = \text{rk}$

È importante notare che le t variabili libere saranno quelle variabili a cui non corrisponde un numero di pivot.

3.2 Sistemi Omogenei

Si chiama **sistema omogeneo** il sistema lineare che ha come termini noti tutti 0, cioè $\mathbf{b} = O$, il che provoca che l'equazione matriciale si scriva come $A \cdot \mathbf{x} = O$. Questi sistemi sono molto importanti perché sono SEMPRE risolubili, infatti essendo la matrice colonna dei termini noti tutta nulla si avrà $(A|O)$ che implica che sarà sempre soddisfatto il requisito per essere risolubile, $\text{rk}(A) = \text{rk}(A|O)$. In un sistema omogeneo, $\mathbf{x} = O$, cioè tutte le incognite uguali a 0 è sempre una soluzione dell'equazione matriciale $A \cdot \mathbf{x} = O$ ma non è detto che sia l'unica.

Teorema 3.5. L'insieme delle soluzioni del sistema omogeneo è un sottospazio vettoriale di \mathbb{R}^n e la sua dimensione è pari a t , le variabili libere del sistema, ($t = n - \text{rk}(A)$).

$$S = \{\mathbf{x} \in \mathbb{R}^n \mid A \cdot \mathbf{x} = O\}$$

ove n è il numero di incognite o equivalentemente il numero di colonne della matrice A .

Teorema 3.6. (TEOREMA DI STRUTTURA DEI SISTEMI LINEARI O DELLE FIBRE)

Il teorema dice che una volta trovata una soluzione particolare \mathbf{x}_p del sistema $A \cdot \mathbf{x} = \mathbf{b}$ tutte le altre soluzioni si ottengono come:

$$\mathbf{x} = \mathbf{x}_{om} + \mathbf{x}_p$$

ove \mathbf{x}_{om} è l'insieme delle soluzioni del sistema omogeneo.

3.3 Inversa di una matrice

Facciamo una piccola parentesi e torniamo un attimo alle matrici, si dice che $A \in \mathcal{M}_n(\mathbb{R})$ ha una inversa (indicata con A^{-1}) se:

$$A \cdot A^{-1} = I_n = A^{-1} \cdot A$$

ovvero se il prodotto della matrice per la sua inversa è la matrice identità. (Si noti che la matrice inversa A^{-1} e A sono permutabili). Di conseguenza si dice **invertibile** o non singolare se la matrice ha una inversa, le matrici che non ammettono matrice inversa si chiamano **singolari**. Tutte le matrici invertibili si caratterizzano per non avere righe nulle (tutte zero) e per avere un determinante $\neq 0$, il concetto di determinante verrà chiarito successivamente.

Chiariamo ora a cosa ci serve conoscere l'inversa di una matrice; in generale, noi abbiamo un sistema $A \cdot \mathbf{x} = \mathbf{b}$ dove si ricorda che \mathbf{x} è il vettore delle incognite, che supponiamo essere in questo caso x, y, \dots se potessimo moltiplicare entambi i lati l'equazione matriciale per A^{-1} avremo:

$$A \cdot \mathbf{x} = \mathbf{b} \quad \rightarrow \quad A^{-1} \cdot A \cdot \begin{bmatrix} x \\ y \\ \vdots \end{bmatrix} = A^{-1} \cdot \mathbf{b}$$

Per definizione di inversa $A^{-1} \cdot A$ diventa I che a sua volta è per definizione l'elemento neutro rispetto al prodotto righe per colonne, quindi qualunque cosa si moltipichi per I rimane inalterato, mettiamo assieme quanto detto:

$$\begin{bmatrix} x \\ y \\ \vdots \end{bmatrix} = A^{-1} \mathbf{b} \tag{3.3}$$

Questa equazione matriciale risulta avere proprio come soluzione la matrice colonna delle incognite che è proprio quello che si cerca con i sistemi lineari. Possiamo quindi concludere che calcolare l'inversa equivale praticamente a risolvere il sistema.

Ora vediamo come si calcola, supponiamo di avere una matrice A , una generica matrice quadrata B , e la sua corrispondente matrice identità I :

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \quad B = \begin{bmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{n1} & \cdots & b_{nn} \end{bmatrix} \quad I = \begin{bmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{bmatrix}$$

Possiamo sfruttare la definizione d'inversa e applicare il prodotto righe per colonne fra la matrice A e la generica B imponendo che il risultato sia $A \cdot B = I$, così facendo otteniamo un sistema di equazioni fatto così:

$$\begin{cases} a_{11}b_{11} + a_{12}b_{21} + \dots = 1 \\ a_{21}b_{11} + a_{22}b_{21} + \dots = 0 \\ \vdots \\ a_{n1}b_{1n} + a_{n2}b_{2n} + \dots = 1 \end{cases}$$

Le soluzioni di questo sistema ovvero b_{11}, \dots, b_{nn} sono i coefficienti della matrice B che messi tutti assieme formano la matrice inversa cercata.

Consideriamo ora, la matrice generica B con i coefficienti incogniti come tre matrici colonna cioè $B = (x_1, x_2, \dots, x_n)$ e la matrice identità come i vettori colonna di dimensione n canonici $I = (e_1, e_2, \dots, e_n)$. Possiamo dunque riscrivere $A \cdot B = I$ come $A \cdot (x_1, x_2, \dots, x_n) = (e_1, e_2, \dots, e_n)$ ottenendo dunque un sistema fatto così:

$$\begin{cases} A \cdot x_1 = e_1 \\ A \cdot x_2 = e_2 \\ \vdots \quad \vdots \quad \vdots \\ A \cdot x_n = e_n \end{cases}$$

ogni sistema si può risolvere separatamente ma visto che avremo bisogno dello stesso metodo, possiamo risolvere simultaneamente tutti gli n sistemi.

I metodi risolutivi mostrati hanno l'inconveniente di essere un poco lenti. L'algoritmo di Gauss-Jordan visto ci permette di trovare una matrice identità partendo da una generica matrice quadrata, ma se dalla matrice generica A ottengo la matrice identità I questo vuol dire che in qualche modo le operazioni fatte fanno sì che si ottenga lo stesso risultato che avremo moltiplicandola per la matrice inversa, quindi se applichiamo la stessa procedura alla matrice identità troveremo proprio la matrice inversa.

Consideriamo dunque la matrice $(A|I)$ ottenuta dalla matrice A affiancata alla matrice identità I , applichiamo a questa il procedimento di Gauss-Jordan per trovare una matrice $(I|A^{-1})$ dove A^{-1} è la matrice inversa cercata. Per chiarire il concetto facciamo un esempio, consideriamo il seguente sistema lineare:

$$\begin{cases} x + 2y = 4 \\ 2x + 5y - 3z = 0 \\ x + 4y - 8z = -18 \end{cases} \quad \text{ossia} \quad A = \begin{bmatrix} 1 & 2 & 0 \\ 2 & 5 & -3 \\ 1 & 4 & -8 \end{bmatrix}$$

Cominciamo innanzitutto riscrivendo la matrice A affiancata con la matrice identità e poi applichiamo Gauss-Jordan.

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & 0 & 0 \\ 2 & 5 & -3 & 0 & 1 & 0 \\ 1 & 4 & -8 & 0 & 0 & 1 \end{array} \right] \xrightarrow{\substack{III-I \\ II-2I}} \left[\begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & 1 & -3 & -2 & 1 & 0 \\ 0 & 2 & -8 & -1 & 0 & 1 \end{array} \right] \xrightarrow{\substack{III-2II \\ I-2II}} \left[\begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & 1 & -3 & -2 & 1 & 0 \\ 0 & 0 & -2 & 3 & -2 & 1 \end{array} \right]$$

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & 1 & -3 & -2 & 1 & 0 \\ 0 & 0 & 1 & -3/2 & 1 & -1/2 \end{array} \right] \xrightarrow{II+3III} \left[\begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & -13/2 & 4 & -3/2 \\ 0 & 0 & 1 & -3/2 & 1 & -1/2 \end{array} \right] \xrightarrow{I-2II} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 14 & -8 & 3 \\ 0 & 1 & 0 & -13/2 & 4 & -3/2 \\ 0 & 0 & 1 & -3/2 & 1 & -1/2 \end{array} \right]$$

Ora che abbiamo l'inversa possiamo riscrivere il sistema come equazione matriciale:

$$\begin{bmatrix} 14 & -8 & 3 \\ -13/2 & 4 & -3/2 \\ -3/2 & 1 & -1/2 \end{bmatrix} \cdot \begin{bmatrix} 4 \\ 0 \\ -18 \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad \text{ovvero} \quad \begin{cases} x = (14)4 + (-8)0 + 3(-18) = 2 \\ y = (-13/2)4 + (4)0 + (-3/2)(-18) = 1 \\ z = (-3/2)4 + (1)0 + (-1/2)(-18) = 3 \end{cases}$$

Teorema 3.7 (Unicità della matrice inversa). Sia $A \in \mathcal{M}_n(\mathbb{R})$ se A ha un'inversa allora è unica.

Ora che sappiamo risolvere sistemi con matrici inverse conviene vedere prima di calcolarla se la matrice di partenza è non singolare, ovvero se ammette inversa, altrimenti è inutile calcolarla!

Teorema 3.8. Sia $A \in \mathcal{M}_n(\mathbb{R})$, questa ammette inversa se

$$\text{rk}(A) = n$$

Poiché dal teorema di Rouché-Capelli sappiamo che se $\text{rk}(A) < n$ il sistema o è impossibile o ha infinite soluzioni ma non può avere infinite soluzioni poiché queste sono scartate dal teorema precedente che ci dice che l'inversa è unica, per cui ci rimane l'ultimo caso che è che ammetta una soluzione, e questo succede solo se $\text{rk}(A) = n$.

Teorema 3.9 (Prodotto di matrici invertibili). Sia $A, B \in \mathcal{M}_n(\mathbb{R})$ entrambe invertibili allora $A \cdot B$ è invertibile e vale:

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

Teorema 3.10. Sia $A, B \in \mathcal{M}_n(\mathbb{R})$ entrambe invertibili allora:

$$A \cdot B \neq O$$

Geometria in \mathbb{R}^3

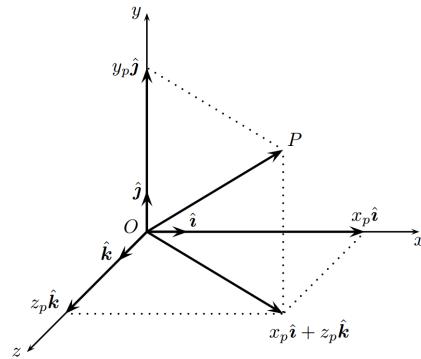
\mathbb{R}^3 così come \mathbb{R}^2 , sono spazi vettoriali che risultano fondamentali in ambiti scientifici come la fisica e la matematica, per questo motivo, analizzeremo in questo capitolo le loro principali proprietà, in modo da poterle utilizzare quando ci servono per risolvere un problema su essi. Innanzitutto, sappiamo dai capitoli precedenti che \mathbb{R}^3 è uno spazio vettoriale reale:

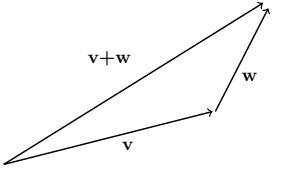
$$\mathbb{R}^3 = \{(x_1, x_2, x_3) \mid x_1, x_2, x_3 \in \mathbb{R}\}$$

questo spazio in particolare è molto importante giacché può rappresentare la posizione di un qualunque oggetto nello spazio. La posizione dell'oggetto è identificata grazie a tre assi, con uno possiamo associare la quota o altezza dell'oggetto rispetto a qualcosa, con un altro possiamo misurare quanto un oggetto è largo e con un ultimo quanto l'oggetto è "spesso". Essendo \mathbb{R}^3 uno spazio vettoriale, per quanto visto nei capitoli precedenti, deve avere una base che ci permetta di generare tutti vettori o punti appartenenti ad esso. Per comodità, useremo come base in questo capitolo, la base canonica $\mathcal{E} = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ dove si ricorda che \mathbf{e}_n è un vettore con coefficienti tutti nulli con 1 nella colonna n ovvero $\mathbf{e}_1 = (1, 0, 0)$ $\mathbf{e}_2 = (0, 1, 0)$ $\mathbf{e}_3 = (0, 0, 1)$. A questi vettori diamo un nome particolare, li chiamiamo $\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}$ rispettivamente. Dai capitoli precedenti sappiamo che la dimensione di uno spazio vettoriale è uguale al numero di vettori di una sua base, in realtà, d'ora in poi possiamo affermare quello che normalmente si intuisce senza sapere, che la dimensione di spazi vettoriali è uguale al numero di coordinate che necessitano per esprimere un punto al suo interno, è per questo che diciamo che \mathbb{R}^3 ha dimensione 3.

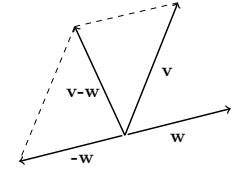
Definizione 4.1. Chiamiamo **origine** il punto $O = (0, \dots, 0_n)$ in questo caso il punto $O = (0, 0, 0)$

Un generico vettore v applicato nell'origine ci individua un punto $P \in \mathbb{R}^3$ cioè $\overline{OP} = v$ e può essere rappresentato come segue:





(a) Somma di vettori



(b) Differenza di vettori

Consideriamo le operazioni che si possono fare con i vettori, ad esempio li possiamo sommare, sottrarre, e moltiplicare per uno scalare. Sommare due vettori non è altro che sommare i coefficienti di ogni componente dei vettori, dati $\mathbf{v} = (v_x, v_y, v_z)$ e $\mathbf{u} = (u_x, u_y, u_z)$, $\mathbf{u} + \mathbf{v}$ non è altro che:

$$\mathbf{u} + \mathbf{v} = (u_x + v_x, u_y + v_y, u_z + v_z)$$

e sottrarre è uguale a sommare $-\mathbf{v}$

$$\mathbf{u} + (-\mathbf{v}) = \mathbf{u} - \mathbf{v} = (u_x - v_x, u_y - v_y, u_z - v_z)$$

Come si può apprezzare dalle figure la somma due vettori è uguale alla diagonale maggiore del parallelogramma mentre la sottrazione è la diagonale minore. La moltiplicazione per uno scalare non è altro allungare o ridurre il vettore per un coefficiente λ ove se $\lambda > 0$ si dilata e $\lambda < 0$ si contrae.

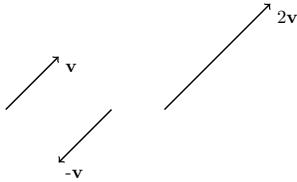


Figura 4.2: Rappresentazione del prodotto scalare per vettore o per uno scalare.

Definizione 4.2. Chiamiamo **prodotto scalare** il prodotto righe per colonne di due vettori $\mathbf{v} = (x_1, y_1, z_1)$ e $\mathbf{w} = (x_2, y_2, z_2)$ ossia:

$$\mathbf{v} \cdot \mathbf{w} = \mathbf{v} \cdot {}^t \mathbf{w} = [x_1, y_1, z_1] \begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix} = x_1 x_2 + y_1 y_2 + z_1 z_2 \quad (4.1)$$

Si noti che il prodotto scalare fra vettori non è uguale al prodotto di uno scalare per un vettore, il primo prende infatti due vettori e ne da uno scalare, mentre il prodotto per uno scalare prende uno scalare (numero) e un vettore e ne da un vettore. Sfortunatamente si usa lo stesso simbolo per denotare entrambi i casi, però si può dedurre facilmente di quale si parla d'accordo col contesto in cui è messo.

prodotto per uno scalare $\mathbf{v} \in \mathbb{R}^3 \cdot k \in \mathbb{R} \rightarrow \mathbf{w} \in \mathbb{R}^3$

prodotto scalare $\mathbf{v}_1 \in \mathbb{R}^3 \cdot \mathbf{v}_2 \in \mathbb{R}^3 \rightarrow \mathbf{w} \in \mathbb{R}$

Teorema 4.1 (Teorema di Carnot). Siano $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{R}^3$ con $\mathbf{v}_1 = (x_1, y_1, z_1)$, $\mathbf{v}_2 = (x_2, y_2, z_2)$ allora:

$$\|\mathbf{v}_2 - \mathbf{v}_1\| = \|\mathbf{v}_1\|^2 + \|\mathbf{v}_2\|^2 - 2\|\mathbf{v}_1\| \cdot \|\mathbf{v}_2\| \cos \alpha \quad (4.2)$$

Teorema 4.2. Siano $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{R}^3$ e α l'angolo tra \mathbf{v}_1 e \mathbf{v}_2 allora:

$$\textcircled{1} \quad \cos \alpha = \frac{\mathbf{v}_1 \cdot \mathbf{v}_2}{\|\mathbf{v}_1\| \cdot \|\mathbf{v}_2\|}$$

(L'angolo α formato tra due vettori è uguale all'arccos del prodotto scalare dei due vettori fratto la norma degli stessi)

$$\textcircled{2} \quad \mathbf{v}_1 \perp \mathbf{v}_2 \text{ sse } \mathbf{v}_1 \cdot \mathbf{v}_2 = 0$$

(Un vettore è perpendicolare a un altro se il loro prodotto scalare è nullo)

$$\textcircled{3} \quad \mathbf{v}_1 \parallel \mathbf{v}_2 \text{ se } \exists \lambda \in \mathbb{R} \text{ tale che: } \mathbf{v}_1 = \lambda \mathbf{v}_2$$

(Un vettore è parallelo a un altro se uno è una combinazione lineare dell'altro)

$$\textcircled{4} \quad \|\mathbf{v}_1\|^2 = \mathbf{v}_1 \cdot \mathbf{v}_1$$

(La lunghezza ovvero la norma di un vettore è uguale al prodotto scalare di se stesso)

Queste relazioni sono molto importanti poiché legano il mondo geometrico con in mondo algebrico in generale.

Introduciamo ora il concetto di lunghezza di un vettore.

Definizione 4.3. Chiamiamo **norma euclidea** la lunghezza di un vettore \mathbf{v} e la indico con $\|\mathbf{v}\|$. Per un vettore con componenti $\mathbf{v} = (x, y, z) \in \mathbb{R}^3$ questa vale:

$$\|\mathbf{v}\| = \sqrt{x^2 + y^2 + z^2} = \sqrt{\mathbf{v} \cdot \mathbf{v}} \quad (4.3)$$

Definizione 4.4. Si definisce **versore** un vettore normalizzato, ossia un vettore moltiplicato per il reciproco della propria norma cioè:

$$\hat{\mathbf{v}} = \mathbf{v} \cdot \frac{1}{\|\mathbf{v}\|} \quad (4.4)$$

$$\|\hat{\mathbf{v}}\| = 1 \quad \text{i versori sono tutti di modulo unitario}$$

I vettori $\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}$ definiti previamente sono versori.

Avendo definito lo spazio, e la lunghezza, viene la necessità di introdurre il concetto di distanza tra due punti in uno spazio reale di dimensione n e in particolare in \mathbb{R}^3 .

Definizione 4.5. Chiamiamo **distanza euclidea tra due punti** $P = (p_1, p_2, \dots, p_n)$ e $Q = (q_1, q_2, \dots, q_n)$ il numero reale non negativo:

$$d(P, Q) = \overline{PQ} = \sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2 + \dots + (q_n - p_n)^2} = \sqrt{\sum_{k=1}^n (q_k - p_k)^2} \quad (4.5)$$

In particolare si può notare che la distanza gode delle seguenti proprietà:

1. La distanza euclidea tra due punti è nulla se e solo se i due punti hanno le stesse coordinate ovvero coincidono.

$$d(P, Q) = 0 \iff P = Q$$

2. La distanza euclidea tra P e Q coincide con la distanza euclidea tra Q ovvero è *simmetrica*

$$d(P, Q) = d(Q, P)$$

3. Dati due punti P, Q , comunque scelto un terzo punto R risulta che la somma della distanza dei due punti al punto R è maggiore o al limite uguale alla distanza fra i punti, questa non è altro che la disegualanza triangolare.

$$d(P, Q) \leq d(P, R) + d(Q, R)$$

4.1 Piani

Possiamo esprimere l'equazione di un piano in relazione con un punto $P = (x_p, y_p, z_p)$ e un vettore perpendicolare al piano $n_\pi = (a, b, c)$ che messi assieme formano quella che chiamiamo **equazione cartesiana del piano**.

$$ax + by + cz = d \quad (4.6)$$

La terna (a, b, c) dei coefficienti è detta terna dei **parametri direttori del piano**, detti altresì **coefficienti direttori del piano**; questi individuano la direzione ortogonale a tutti i vettori del piano; il termine noto d si ottiene con la relazione

$$ax_p + by_p + cz_p = d$$

Quindi possiamo dedurre che qualsiasi piano che passa per l'origine allora deve avere necessariamente il termine d uguale a 0. Inoltre, se l'equazione ha nullo un parametro, ovvero nell'equazione non compare una tra le variabili x, y, z allora il piano è parallelo all'asse delle variabili che non compare, cioè all'asse il cui coefficiente direttore è 0.

Infine, possiamo scrivere l'equazione di un piano π come la superficie che passa per un punto $P = (x_p, y_p, z_p)$ e parallela a due vettori $\mathbf{u} = (u_1, u_2, u_3)$ e $\mathbf{v} = (v_1, v_2, v_3)$; questa equazione la chiamiamo **equazione parametrica del piano** e la si scrive così:

$$\pi : \begin{cases} x = x_p + tu_1 + sv_1 \\ y = y_p + tu_2 + sv_2 \\ z = z_p + tu_3 + sv_3 \end{cases} \quad \forall s, t \in \mathbb{R} \quad (4.7)$$

t, s sono dei **parametri reali indipendenti**. Al variare di t, s le precedenti equazioni individuano tutti i punti del piano π considerato.

4.2 Rette

Una retta non è altro che l'intersezione fra due piani, π e π' . Comunque non tutti i piani si intersecano, come ad esempio nel caso di due piani paralleli tra loro, il che si traduce come la lineare dipendenza dei coefficienti a', b', c' della equazione cartesiana di uno dei piani rispetto ai coefficienti a, b, c dell'altro, per cui, solo nel caso in cui i $????$ si intersechino sarà definita la retta. Per rappresentare una retta usiamo l'**equazione cartesiana della retta** che come conseguenza della definizione, risulta essere il sistema delle due equazione cartesiane dei piani.

$$r : \begin{cases} ax + by + cz = d \\ a'x + b'y + c'z = d' \end{cases} \quad (4.8)$$

Risulta anche per come è stata definita la retta che r sarà \perp a le direzioni (a, b, c) e a (a', b', c') dove si ricorda che:

(a, b, c) direzione ortogonale al piano π : $ax + by + cz = d$

(a', b', c') direzione ortogonale al piano π' : $a'x + b'y + c'z = d'$

La retta passante per i punti $A = (x_A, y_A, z_A)$ e $B = (x_B, y_B, z_B)$ si può individuare con l'**equazione parametrica della retta**

$$\begin{cases} x(t) = x_A + t(x_B - x_A) \\ y(t) = y_A + t(y_B - y_A) \\ z(t) = z_A + t(z_B - z_A) \end{cases} \quad (4.9)$$

È importante notare che l'equazione parametrica della retta non è unica, anzi, ce ne sono ∞^2 per il teorema di Rouché-Capelli poiché se si considera la retta come un asse fisso, i piani possono essere ruotati in infinite direzioni e hanno tutti la stessa retta.

Alternativamente, si può individuare con un punto $A = (x_A, y_A, z_A)$ e \parallel a $d = (l, m, n)$ dove d è il **vettore direzione** della retta che è ricavabile come nella formula precedente.

$$r : \begin{cases} x(t) = x_A + lt \\ y(t) = y_A + mt \\ z(t) = z_A + nt \end{cases} \quad \forall t \in \mathbb{R} \quad (4.10)$$

La retta può essere trovata anche con gli angoli che essa ha con gli assi del sistema cartesiano, questi angoli li chiamiamo **coseni direttori** e possono essere calcolati come segue:

$$\begin{aligned} \cos \alpha &= \frac{l}{\sqrt{l^2 + m^2 + n^2}} \\ \cos \beta &= \frac{m}{\sqrt{l^2 + m^2 + n^2}} \\ \cos \gamma &= \frac{n}{\sqrt{l^2 + m^2 + n^2}} \end{aligned} \quad (4.11)$$

Inoltre questi aiutano ad individuare un versore, cioè un vettore direzione di modulo unitario della retta :

$$\hat{\mathbf{v}} = \left(\frac{l}{\sqrt{l^2 + m^2 + n^2}}, \frac{m}{\sqrt{l^2 + m^2 + n^2}}, \frac{n}{\sqrt{l^2 + m^2 + n^2}} \right) \quad (4.12)$$

4.3 Mutue posizioni

4.3.1 Fra due piani

Supponiamo di avere l'equazione di due piani in forma cartesiana

$$\pi : ax + by + cz = d \quad \text{e} \quad \pi' : a'x + b'y + c'z = d'$$

La posizione di un piano rispetto all'altro dipende dal calcolo del rango delle seguenti matrici

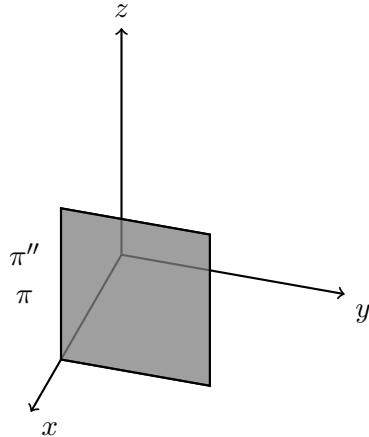
$$A = \begin{bmatrix} a & b & c \\ a' & b' & c' \end{bmatrix} \quad \text{e} \quad (A|b) = \left[\begin{array}{ccc|c} a & b & c & d \\ a' & b' & c' & d' \end{array} \right]$$

Dunque le mutue posizioni fra due piani dipendono dalla soluzione del sistema di due equazioni. Come sappiamo dal capitolo sui sistemi lineari, un sistema ha $\infty^{incognite-rk}$ soluzioni cioè, in questo caso ha ∞^{3-rk} soluzioni, i casi dipendono proprio dal rango del sistema.

Se il $\text{rk}(A) = 1$ possiamo dedurre che π e π' sono perpendicolari agli stessi vettori ma niente di più, infatti ci sono due possibili casi: Quando il $\text{rk}(A) = 1$ e la matrice completa ha lo stesso rango $\text{rk}(A|b) = 1$ oppure quando $\text{rk}(A|b) = 2$.

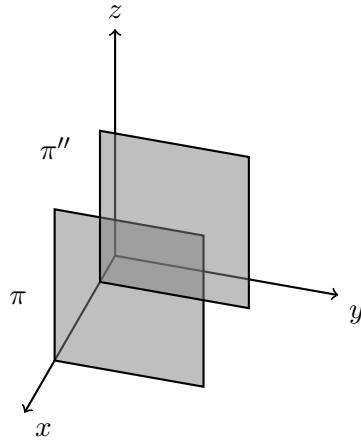
Sono uguali

Quando entrambe le matrici hanno lo stesso rango, ovvero: $\text{rk}(A) = 1$ e $\text{rk}(A|b) = 1$ i due piani risultano essere uguali $\pi = \pi'$



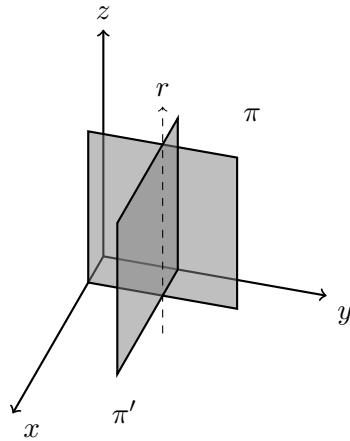
Paralleli

Con lo stesso discorso di prima se $\text{rk}(A) = 1$ e se $\text{rk}(A|b) = 2$ abbiamo per il teorema di Rouché-Capelli un sistema impossibile quindi $\pi \cap \pi' = \emptyset$ il che vuol dire che i piani non hanno alcun punto in comune. Il che obbliga, dunque, che i piani siano paralleli.



Formano una retta

È il caso più facile e succede quando $\text{rk}(A) = 2$ cioè ci sono ∞^1 soluzioni, questo succede quando $\pi \cap \pi'$ dunque la soluzione sarà una retta per definizione



4.3.2 Fra tre piani

Come si potrebbe intuire, le mutue posizioni fra tre piani dipendono dalla soluzione del sistema di equazioni:

$$\begin{cases} \pi : ax + by + cz = d \\ \pi' : a'x + b'y + c'z = d' \\ \pi'' : a''x + b''y + c''z = d'' \end{cases}$$

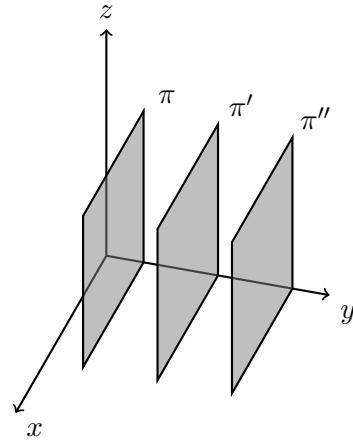
e ha $\infty^{3-\text{rk}}$ soluzioni per cui possiamo distinguere con questo sistema ben 6 casi. Come nel caso di due piani, la posizione di un piano rispetto all'altro dipende dal calcolo del rango delle seguenti matrici:

$$A = \begin{bmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \end{bmatrix} \quad \text{e } (A|b) = \left[\begin{array}{ccc|c} a & b & c & d \\ a' & b' & c' & d' \\ a'' & b'' & c'' & d'' \end{array} \right]$$

Paralleli tra loro

Quando ha $\text{rk}(A) = 1$ e la completa ha $\text{rk}(A|b) = 2$ oppure $\text{rk}(A|b) = 3$, i piani sono paralleli tra loro, cioè $\pi \parallel \pi' \parallel \pi''$ oppure $\pi = \pi' \parallel \pi''$ oppure $\pi' = \pi'' \parallel \pi$ oppure $\pi = \pi'' \parallel \pi'$.

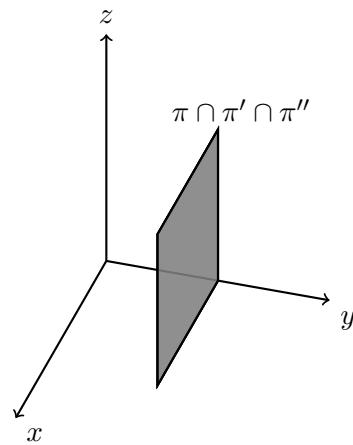
$$\text{rk}(A) = \begin{bmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \end{bmatrix} = 1 \quad \text{e} \quad \text{rk}(A|b) = \begin{bmatrix} a & b & c & d \\ a' & b' & c' & d' \\ a'' & b'' & c'' & d'' \end{bmatrix} = 2 \text{ oppure } 3$$



Uguali

Quando si ha $\text{rk}(A) = 1$ e la matrice completa ha $\text{rk}(A|b) = 1$ i piani sono tutti uguali, cioè $\pi = \pi' = \pi''$.

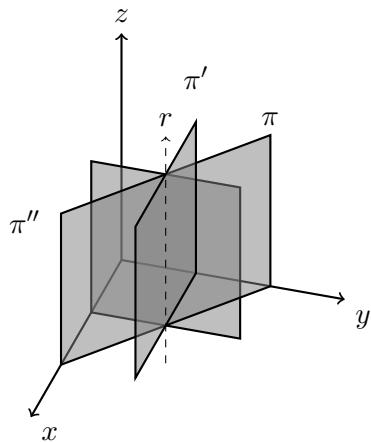
$$\text{rk}(A) = \begin{bmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \end{bmatrix} = 1$$



Intersecano in una retta

Quando ha $\text{rk}(A) = 2$ e la matrice completa ha $\text{rk}(A|b) = 2$ cioè $\pi \cap \pi' \cap \pi'' = r$.

$$\text{rk}(A) = \begin{bmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \end{bmatrix} = 2 \quad \text{e} \quad \text{rk}(A|b) = \begin{bmatrix} a & b & c & d \\ a' & b' & c' & d' \\ a'' & b'' & c'' & d'' \end{bmatrix} = 2$$

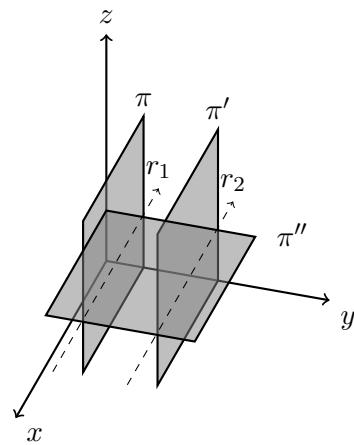


Intersecano in due rette

Questo succede se soltanto se $\pi \cap \pi' \cap \pi'' = \emptyset$ e $\begin{cases} \pi \cap \pi'' = r_1 \\ \pi' \cap \pi'' = r_2 \\ \pi \cap \pi' = \emptyset \end{cases}$

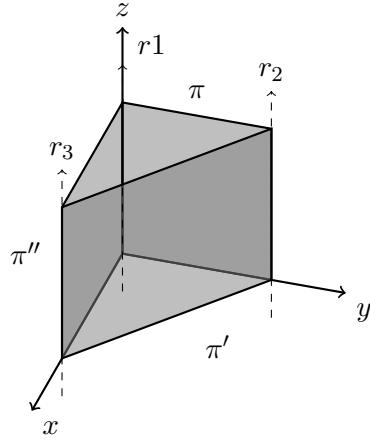
I ruoli sono certamente intercambiabili.

$$\text{rk}(A) = \begin{bmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \end{bmatrix} = 2 \quad \text{e} \quad \text{rk}(A|b) = \begin{bmatrix} a & b & c & d \\ a' & b' & c' & d' \\ a'' & b'' & c'' & d'' \end{bmatrix} = 2 \quad \text{oppure} \quad 3$$



Intersecano in tre rette

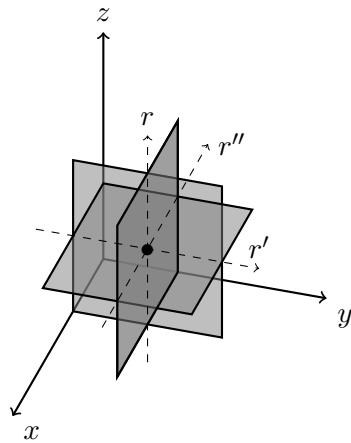
Questo succede se soltanto se $\pi \cap \pi' \cap \pi'' = \emptyset$ e

$$\begin{cases} \pi \cap \pi'' = r_1 \\ \pi \cap \pi' = r_2 \\ \pi' \cap \pi'' = r_3 \end{cases}$$


Intersecano in un solo punto

Questo succede se soltanto se

$$\text{rk}(A) = \begin{bmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \end{bmatrix} = 3 \quad \text{e logicamente} \quad \text{rk}(A|b) = \left[\begin{array}{ccc|c} a & b & c & d \\ a' & b' & c' & d' \\ a'' & b'' & c'' & d'' \end{array} \right] = 3$$

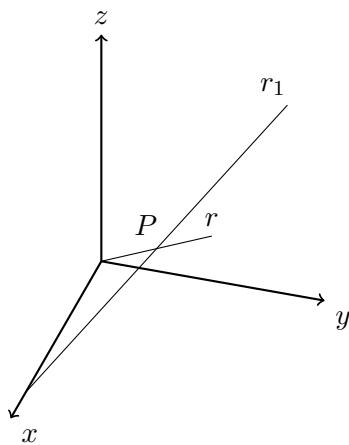


4.3.3 Fra due rette

Possiamo seguire con le rette lo stesso principio che con i piani ossia ci basta trovare la soluzione del sistema lineare formato dall'equazione delle due rette per vedere come esse si distribuiscono nello spazio, in questi casi però è meglio mettere a sistema un'equazione cartesiana e una parametrica per poi sostituire i parametri nelle due equazioni dei piani che definiscono la retta nella forma cartesiana, trovando facilmente eventuali contraddizioni che indicano che il sistema ha soluzione o no.

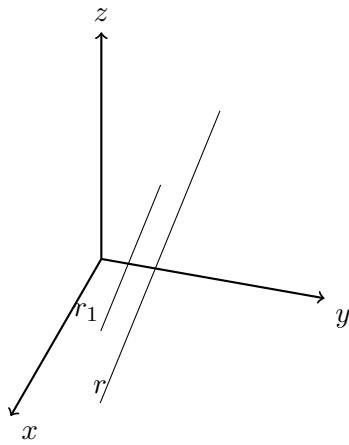
4.3.4 Incidenti

Quando il sistema fra due rette è possibile e ammette una sola soluzione, le rette si dicono incidenti, ossia, si incontrano in un punto che sarà la soluzione del sistema lineare.



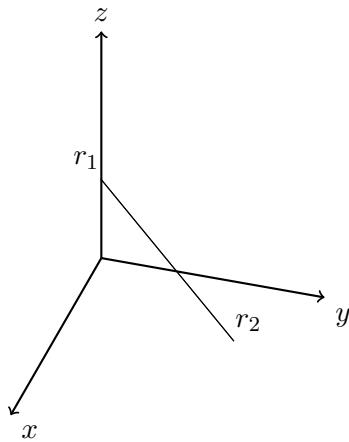
Parallele

Le rette si dicono parallele se hanno la stessa direzione ovvero se i suoi vettori direzione sono proporzionali a meno di una costante, in particolare possiamo dire che sono parallele se esiste un piano π che le contiene entrambe, questo succede se il sistema è impossibile e i vettori direzioni delle due rette sono linearmente dipendenti.



Uguali

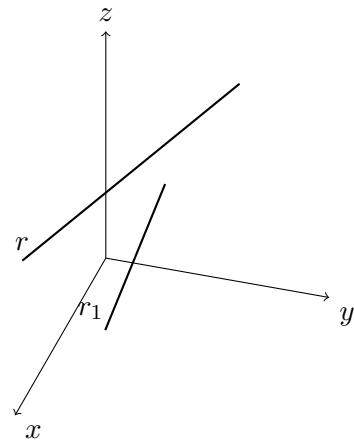
Questo succede quando il sistema è possibile e ha infinite soluzioni.



Definizione 4.6. Due rette si dicono **complanari** se sono parallele oppure si intersecano in almeno un punto.

Sghembe

Due rette che non hanno punti in comune ovvero non si intersecano e non sono parallele (non complanari) si dicono **sghembe**. Questo succede quando il sistema delle equazioni delle rette è impossibile e i loro vettori direzioni non sono proporzionali.



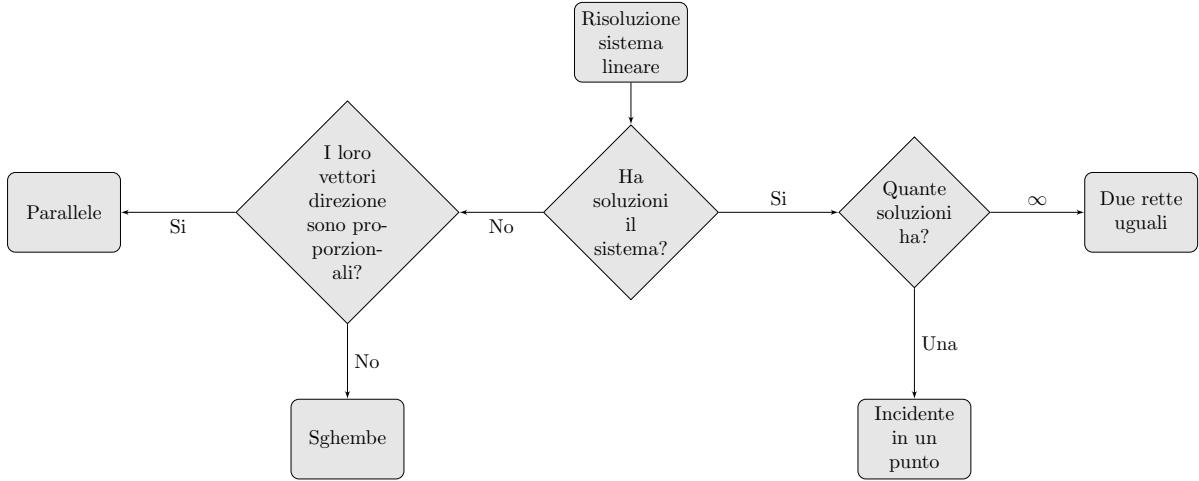
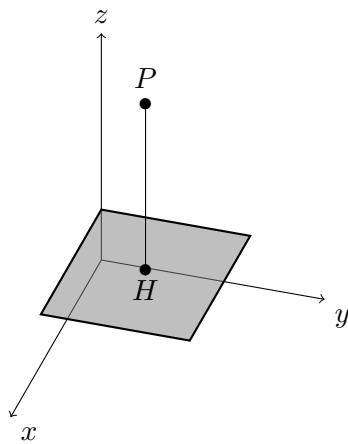


Figura 4.3: Algoritmo identificativo

4.4 Distanze fra rette, punti e piani

4.4.1 Distanza punto-piano

Consideriamo un punto $P = (x_p, y_p, z_p)$ e un piano π se $P \in \pi$, allora la distanza fra il punto e il piano è nulla e si scrive: $d(P, \pi) = 0$ se invece $P \notin \pi$, allora possiamo misurare la distanza fra il punto P e il piano, ci sono infinite distanze diverse dipendendo da che punto del piano prendiamo come riferimento, a noi ci interessa conoscere la distanza minima, per fare ciò introduciamo un nuovo punto che chiamiamo H che sarà il punto dove interseca il vettore normale (o perpendicolare) al piano passante per P e il piano stesso.



Per ottenere la distanza dobbiamo formulare l'equazione parametrica della retta con i parametri direttori del piano passante per il punto P

$$r = \begin{cases} y = x_p + at \\ x = y_p + bt \\ z = z_p + ct \end{cases}$$

avendo fatto questo mettiamo la retta e il piano a sistema per trovare la loro intersezione. Si ha:

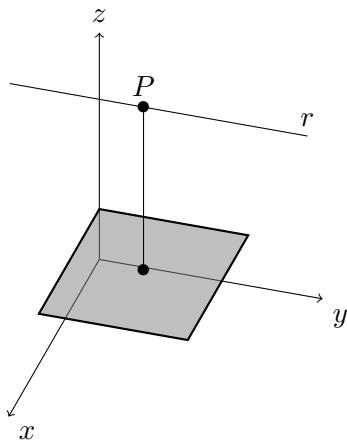
$$H = a(x_p + at) + b(y_p + bt) + c(z_p + ct) = d$$

La soluzione del sistema è il parametro t che sostituito nell'equazione parametrica della retta ci dà le coordinate del punto H . Questo procedura richiede un po' di tempo per la sua esecuzione ma per fortuna, in questo in questo caso specifico, è possibile ottenere una semplificazione derivata proprio dallo svolgimento letterale (senza sostituire i valori) del sistema, la distanza tra una retta e un piano è data dalla seguente formula:

$$\frac{|ax_p + by_p + cz_p - d|}{\sqrt{a^2 + b^2 + c^2}} \quad (4.13)$$

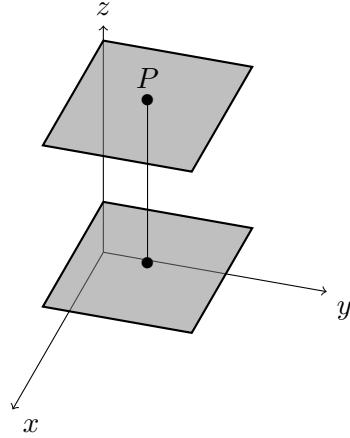
4.4.2 Distanza retta-piano

Consideriamo una retta r e un piano π se $r \cap \pi \neq \emptyset$ allora la distanza è nulla $d(r, \pi) = 0$ se invece $r \cap \pi = \emptyset$ allora possiamo dire che la retta è parallela al piano e per misurare la distanza fra il piano e la retta basta scegliere un punto a caso appartenente alla retta; questo lo si fa assegnando un valore t arbitrario nella equazione parametrica della retta, ciò fatto possiamo applicare la formula fra un punto e un piano.



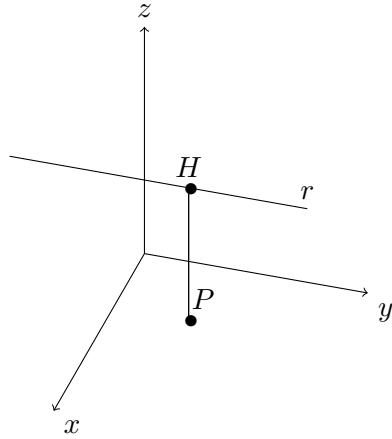
4.4.3 Distanza piano-piano

Consideriamo un piano π e un altro π' se $\pi \cap \pi' \neq \emptyset$ allora la distanza è nulla $d(\pi, \pi') = 0$ se invece $\pi \cap \pi' = \emptyset$ allora possiamo dire che i piani sono paralleli quindi ci basta prendere un punto a caso appartenente al piano e calcolare la distanza fra il punto e il piano.



4.4.4 Distanza punto-retta

Consideriamo una retta r e un punto P , come nei casi precedenti, se $P \in r$ allora la distanza è nulla cioè $d(P, r) = 0$, se invece $P \notin r$ allora possiamo misurare la distanza tra il punto e la retta ma, come nel caso del punto e il piano, sono infinite le distanze dal punto alla retta, dipendendo da quale punto della retta si prenda; consideriamo anche in questo caso la distanza minima fra entrambi, per fare ciò bisogna trovare una retta s perpendicolare a r che passi per P che ci permetta di calcolare $s \cap r$ che sarà nuovamente un punto che nominiamo H , di modo tale che la distanza tra la retta è il punto sia la distanza tra il punto H e il punto P .



Tuttavia come nei casi precedenti il punto H non viene considerato perché è possibile ottenere lo stesso risultato con una procedura più semplice dato che è meno difficoltoso (almeno in \mathbb{R}^3) trovare il piano π perpendicolare a r e passante per P , infatti tale piano contiene tutte le rette perpendicolari a r (tra cui esattamente una incidente e passante per P) per fare ciò basta porre il

vettore direzione della retta r come parametri direttori del piano e considerare il passaggio per il punto P con il calcolo di d , ossia:

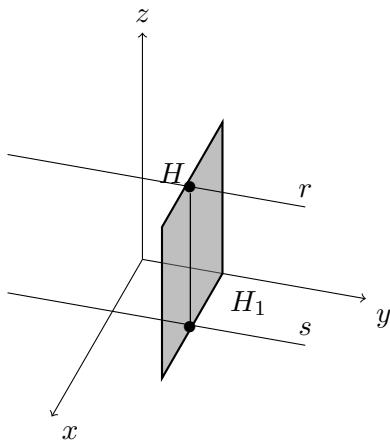
$$\pi : ax + by + cz = ax_p + by_p + cz_p \quad (4.14)$$

4.4.5 Distanza retta-retta

Consideriamo una retta r e un'altra retta s come si può intuire se $r \cap s \neq \emptyset$ allora la distanza è nulla $d(r, s) = 0$ se invece $r \cap s = \emptyset$ allora possiamo suddividere in due casi:

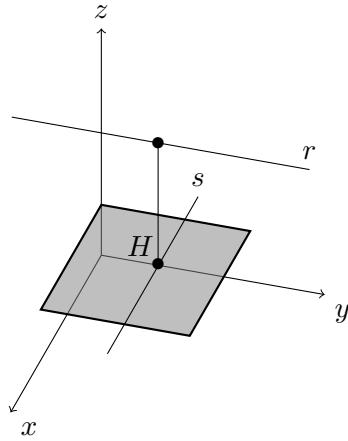
Le rette sono parallele

In questo caso troviamo un piano $\pi \perp r$, l'intersezione fra il piano e la retta sarà un punto H che nei casi precedenti è stato ignorato, ma non solo, (troviamo poi l'intersezione fra il piano $\pi \perp s$, e la retta s che chiamiamo, con grande fantasia H_1 . In questo caso i punti H non vanno ignorati dato che è l'unico metodo risolutivo, la distanza fra le due rette è la distanza tra il punto H e H_1 .



Le rette sono sghembe

Nel caso le rette siano sghembe individuiamo un piano che contiene r oppure s ed è parallelo all'altra retta, questa procedura risulta essere facile poiché niente ci vieta di usare i vettori direzione di entrambe le rette e un punto appartenente a una delle due rette, una volta trovato il piano realizziamo la stessa procedura fatta per ottenere la distanza tra una retta e un piano, basta scegliere un punto a caso appartenente alla retta non contenuta nel piano e applicare la formula.



4.5 Isometrie

Le isometrie (*iso=uguale metro=misura*) sono trasformazioni del piano o dello spazio che mantengono le distanze fra i punti cioè sono delle applicazioni lineari che conservano la distanza fra i vettori di partenza e la loro rispettiva immagine. Cioè:

$$\forall P, Q : \quad d(\overline{PQ}) = d(\overline{f(P), f(Q)})$$

Definizione 4.7. Un punto P tale che $P = f(P) = P'$, ossia che non si sposta con la funzione, è detto **punto fisso**; una retta r tale che $r = f(r) = r'$ è detta **retta fissa**.

Ci sono quattro tipi di isometrie suddivise in due categorie.

4.5.1 Isometrie dirette

Le isometrie dirette oltre a mantenere le distanze costanti, mantengono inoltre l'orientamento degli angoli. Ci sono due tipi di isometrie dirette:

- **Traslazioni:** sono isometrie che “spostano” tutti i punti nella stessa direzione, per cui non ha punti fissi.
- **Rotazioni:** sono isometrie che “ruotano” i punti rispetto un punto fisso che chiamiamo **centro di rotazione**.

4.5.2 Isometrie inverse

Le isometrie inverse mantengono le distanze ma non mantengono l'orientamento degli angoli. Ci sono due tipi di isometrie inverse:

- **Riflessioni:** fanno la simmetria o “specchiano” tutti i punti rispetto ad una retta o un piano che prende nome di **asse o piano di riflessione**.
- **Anttraslazioni o glissoriflessioni:** fanno una simmetria rispetto a una retta e inoltre fanno una traslazione parallela a questa retta. Queste isometrie si caratterizzano per non avere punti fissi.

Applicazioni Lineari

Negli spazi vettoriali, proprio come negli insieme di numeri, si può definire in essi una funzione. Come tutte le funzioni, questa avrà un dominio e un codominio, ma questi non saranno degli insiemi di numeri ma degli spazi vettoriali, dunque, siano V e W due spazi vettoriali si chiama funzione o **applicazione** l'espressione:

$$f : V \rightarrow W$$

Definizione 5.1. Diciamo che è un'**applicazione lineare** se verifica le seguenti proprietà

1. *Additività*

$$\forall \mathbf{v} \in V \quad f(\mathbf{v}_1 + \mathbf{v}_2) = f(\mathbf{v}_1) + f(\mathbf{v}_2)$$

2. *Omegeneità*

$$\forall \mathbf{v} \in V, \forall \lambda \in \mathbb{R} \quad f(\lambda \mathbf{v}) = \lambda f(\mathbf{v})$$

Possiamo sintetizzare le condizioni 1. e 2. in una sola condizione, in realtà basta che sia soddisfatta quella che chiameremo **condizione di linearità** se $\forall \alpha, \beta \in \mathbb{R}$ e $\forall \mathbf{v}_1, \mathbf{v}_2 \in V$ risulta che:

$$f(\alpha \mathbf{v}_1 + \beta \mathbf{v}_2) = \alpha f(\mathbf{v}_1) + \beta f(\mathbf{v}_2) \quad (5.1)$$

Possiamo dire dalle precedenti proprietà che una applicazione lineare è un **omomorfismo**, ovvero è un'applicazione tra due strutture algebriche dello stesso tipo che conserva le operazioni in esse definite.

Definizione 5.2. Diciamo che un'applicazione lineare è un **endomorfismo** se l'applicazione è da uno spazio vettoriale a se stesso, $f : V \rightarrow V$

Definizione 5.3. Data un'applicazione lineare $f : V \rightarrow W$ definiamo l'**immagine** di f , e la indichiamo con $\text{Im } f$, il sottoinsieme del codominio:

$$\text{Im } f := \{\mathbf{w} \in W \mid \exists \mathbf{v} \in V, f(\mathbf{v}) = \mathbf{w}\} \quad (5.2)$$

cioè tutte le immagini degli elementi di V mediante f .

Definizione 5.4. Data un'applicazione lineare $f : V \rightarrow W$ definiamo il **nucleo** o *kernel* di f , e la indichiamo con $\ker f$, il sottoinsieme del dominio:

$$\ker f := \{\mathbf{v} \in V \mid f(\mathbf{v}) = \underline{0}_W\} \quad (5.3)$$

cioè l'insieme degli elementi del dominio V che hanno immagine $\underline{0}$ mediante f .

Teorema 5.1 ($\text{Im } f$ e $\ker f$ sono sottospazi vettoriali). Sia un'applicazione lineare $f : V \rightarrow W$ fra gli spazi vettoriali V, W Allora:

$$\ker f \subseteq V \quad \text{ed è un sottospazio vettoriale di } V$$

$$\text{Im } f \subseteq W \quad \text{ed è un sottospazio vettoriale di } W$$

Teorema 5.2. Sia un'applicazione lineare $f : V \rightarrow W$ e sia $\mathcal{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ una base di V . Allora $f(\mathbf{v}_1), \dots, f(\mathbf{v}_n)$ è un sistema di generatori per $\text{Im } f$ (ma non necessariamente una base poiché i vettori possono essere linearmente dipendenti) cioè:

$$\{f(\mathbf{v}_1), \dots, f(\mathbf{v}_n)\} \subseteq \text{Im } f \quad (5.4)$$

Teorema 5.3. Sia un'applicazione lineare $f : V \rightarrow W$ allora:

- f è **suriettiva** se e solo se $\text{Im } f = W$
- f è **iniettiva** se e solo se $\ker f = \underline{0}_V$

Nel contesto delle applicazioni lineari (omomorfismi) diamo un nome particolare a questi due casi, se è iniettiva l'applicazione si dice **monomorfismo**, se è suriettiva si dice **epimorfismo**.

Definizione 5.5. Un'applicazione lineare $f : V \rightarrow W$ si dice un **isomorfismo** se è sia iniettiva che suriettiva, ovvero biettiva o biunivoca; in questo caso V, W si dicono isomorfi tra loro e si scrive $V \simeq W$.

Così come tutte le funzioni, dato che esiste una corrispondenza biunivoca tra input-output della funzione è possibile quindi assegnare all'applicazione lineare la sua applicazione inversa che indichiamo con il solito simbolo f^{-1} questa sarà $f^{-1} : W \rightarrow V$ tale che:

$$f^{-1}(\mathbf{w}) = \mathbf{v} \quad \text{ove } \mathbf{v} \text{ è l'unico vettore tale che } f(\mathbf{v}) = \mathbf{w} \quad (5.5)$$

Definizione 5.6. Un endomorfismo $f : V \rightarrow V$ (ove dominio e codominio coincidono) che risulta essere inoltre biettivo lo si chiama **automorfismo**.

Teorema 5.4 (TEOREMA DI NULLITÀ PIÙ RANGO). Sia un'applicazione lineare $f : V \rightarrow W$ fra gli spazi vettoriali V, W con $\dim V = n$ Allora vale:

$$\dim(\text{Im } f) + \dim(\ker f) = n \quad (5.6)$$

Corollario 5.4.1. Sia un'applicazione lineare $f : V \rightarrow W$ con $\dim V = \dim W = n$ allora:

- f è iniettiva se e solo se $\dim(\ker f) = 0$
- f è iniettiva se e solo se è suriettiva

Questo risultato vale soltanto perché $\dim V = \dim W$ altrimenti $\text{Im } f$ e $\ker f$ sono indipendenti tra loro. In generale possiamo affermare:

Corollario 5.4.2. Sia un'applicazione lineare $f : V \rightarrow W$ con $\dim V = n$ e $\dim W = m$ allora:

- Se $n < m$ la f non può essere suriettiva.
- Se $n > m$ la f non può essere iniettiva.

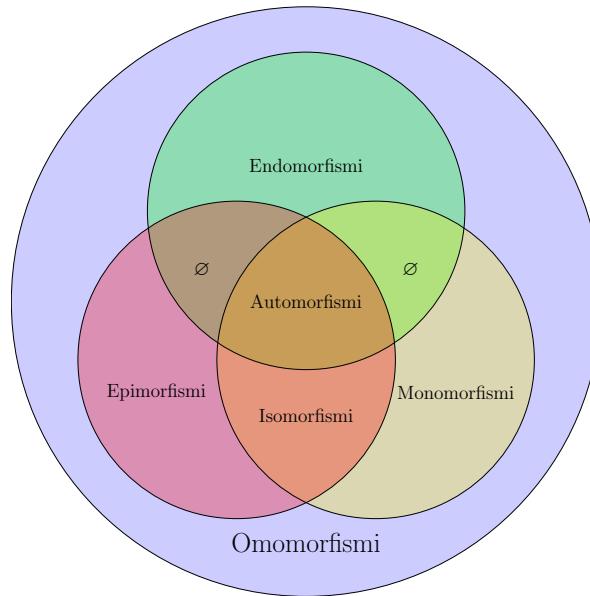


Figura 5.1: Si noti che un endomorfismo è iniettivo se e solo se è suriettivo come conseguenza del corollario 5.4.1

Teorema 5.5. Sia V uno spazio vettoriale di dimensione finita ovvero $\dim V = n$. Allora V è isomorfo a \mathbb{R}^n

$$\mathbb{R}^n = \{(x_1, x_2, \dots, x_n) \mid x_i \in \mathbb{R}, i = 1, 2, \dots, n\}$$

Teorema 5.6. Sia un isomorfismo $f : V \rightarrow W$ allora $f^{-1} : W \rightarrow V$ è lineare e un isomorfismo.

Teorema 5.7. Siano applicazioni lineari $f : U \rightarrow V$ e $g : V \rightarrow W$ allora $g \circ f : U \rightarrow W$ è lineare.

Osservazione 5.7.1. I teoremi precedenti ci fanno vedere che la linearità si conserva.

Teorema 5.8. Sia un isomorfismo $f : V \rightarrow W$ e $f^{-1} : W \rightarrow V$ un altro isomorfismo, allora:

$$\begin{array}{lll} \text{id}_W = f \circ f^{-1} & W \rightarrow W & \text{identità su } W \\ \text{id}_V = f^{-1} \circ f & V \rightarrow V & \text{identità su } V \end{array}$$

Consideriamo due spazi vettoriali, V e W con $\dim V = m$ e $\dim W = n$, una applicazione lineare $f : V \rightarrow W$ e $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m\}$ la base di V e $\mathcal{C} = \{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_n\}$ la base di W .

Consideriamo ora un generico vettore $\mathbf{v} \in V$, questo per definizione può essere espresso come combinazione lineare di una sua base ovvero:

$$\exists! \quad b_1, b_2, \dots, b_m \in \mathbb{R} \quad \text{tale che} \quad \mathbf{v} = b_1 \mathbf{v}_1 + b_2 \mathbf{v}_2 + \dots + b_m \mathbf{v}_m$$

quindi $f(\mathbf{v}) = f(b_1 \mathbf{v}_1 + b_2 \mathbf{v}_2 + \dots + b_m \mathbf{v}_m)$ per linearità diventa: $b_1 f(\mathbf{v}_1) + b_2 f(\mathbf{v}_2) + \dots + b_m f(\mathbf{v}_m)$ inoltre sappiamo che $f(\mathbf{v}) \in W$ per cui $f(\mathbf{v})$ per definizione può essere espresso come combinazione lineare di una sua base cioè:

$$f(\mathbf{v}_1) = a_{11} \mathbf{w}_1 + a_{12} \mathbf{w}_2 + \dots + a_{1n} \mathbf{w}_n$$

$$f(\mathbf{v}_2) = a_{21} \mathbf{w}_1 + a_{22} \mathbf{w}_2 + \dots + a_{2n} \mathbf{w}_n$$

⋮

$$f(\mathbf{v}_n) = a_{n1} \mathbf{w}_1 + a_{n2} \mathbf{w}_2 + \dots + a_{nn} \mathbf{w}_n$$

mettendo entrambe assieme si ha:

$$b_1(a_{11} \mathbf{w}_1 + a_{12} \mathbf{w}_2 + \dots + a_{1n} \mathbf{w}_n) + b_2(a_{21} \mathbf{w}_1 + a_{22} \mathbf{w}_2 + \dots + a_{2n} \mathbf{w}_n) + \dots + b_n(a_{n1} \mathbf{w}_1 + a_{n2} \mathbf{w}_2 + \dots + a_{nn} \mathbf{w}_n)$$

che assomiglia molto a un prodotto righe per colonne di matrici per cui considero la matrice e la colonna consideriamo allora

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \quad B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \quad \Rightarrow A \cdot B = C \in \mathbb{R}^m$$

Teorema 5.9 (TEOREMA DI RAPPRESENTAZIONE). Ogni applicazione lineare $f : V \rightarrow W$ con $\dim V = n$ e $\dim W = m$ si rappresenta in modo unico con una matrice $A \in \mathcal{M}(m, n, \mathbb{R})$ detta **matrice rappresentativa o matrice associata**, ovvero:

$$\forall \mathbf{v} \in V \quad f(\mathbf{v}) = A \cdot \mathbf{v}$$

Teorema 5.10. Sia un'applicazione lineare $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ e sia A la matrice rappresentativa di f allora $\text{Im } f$ è generata dalle colonne di A , che equivale a dire

$$\text{Im } f = \text{rk } A$$

di conseguenza per il teorema di nullità più rango:

$$\ker f = n - \operatorname{rk} A$$

quindi per trovare la dimensione dell'immagine basta ridurre la matrice associata all'applicazione lineare e calcolarne il rango, per trovare una base dell'immagine basta scrivere come vettori della base, le colonne linearmente indipendenti, bisogna fare però attenzione giacché al ridurre a scala si cambiano i coefficienti dei vettori originali, allora per sapere quali sono le colonne linearmente indipendenti basta vedere dove sono posizionati i pivot, le colonne dove compaiono i pivot nella matrice ridotta sono le stesse colonne linearmente indipendenti della matrice di partenza.

Per trovare la dimensione del nucleo basta utilizzare la formula del teorema di nullità più rango, per trovarne una base sfruttiamo la definizione di nucleo e risolviamo il sistema omogeneo associato ad A , in sostanza:

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} x \\ y \\ \vdots \\ m \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \Rightarrow \begin{cases} a_{11}x + a_{12}y + \dots + a_{1n}m = 0 \\ a_{21}x + a_{22}y + \dots + a_{2n}m = 0 \\ \vdots \\ a_{m1}x + a_{m2}y + \dots + a_{mn}m = 0 \end{cases}$$

Teorema 5.11. Sia un'applicazione lineare $f : U \rightarrow V$ e $g : V \rightarrow W$ con $\dim U = n$ $\dim V = m$ $\dim W = l$ e siano $A \in \mathcal{M}(m, n, \mathbb{R})$ la matrice rappresentativa dell'applicazione lineare f e $B \in \mathcal{M}(l, m, \mathbb{R})$ la matrice rappresentativa dell'applicazione lineare g allora $g \circ f$ è lineare e la sua matrice rappresentativa è $C \in \mathcal{M}(l, n, \mathbb{R})$ ottenuta come

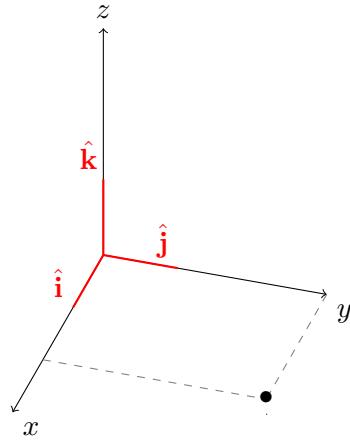
$$\underbrace{C}_{l \times n} = \underbrace{B}_{l \times m} \cdot \underbrace{A}_{m \times n}$$

5.1 Cambiamento delle basi

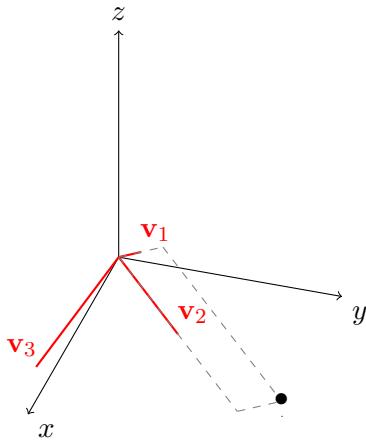
Immaginiamo di dover misurare la lunghezza di un tratto, possiamo misurarla in diversi modi, se è abbastanza piccola magari con il numero di dita che occupa, o se è lunga la possiamo misurare in confronto a quanti "campioni base" questa misura, i "campioni base" ci permettono di esprimere qualunque misura dello stesso tipo come un multiplo o combinazione lineare di quel campione base, ad esempio, oggiorno abbiamo come campioni base di lunghezza il metro (che appartiene al Sistema Internazionale di unità di misura) e la iarda (che forma parte del Sistema Imperiale) ma ce ne possono essere tanti quanti si voglia. Possiamo dire dunque che ci sono tanti modi di esprimere una grandezza, questa sarà espressa con un numero che rappresenta il numero di volte che il campione base sta nella lunghezza che si vuole misurare, certamente il numero varierà a seconda del campione base scelto inizialmente. Consideriamo questo stesso problema nell'ambito ad esempio dello spazio tridimensionale, qualunque punto è esprimibile usando un sistema di riferimento cartesiano e la base canonica $\mathcal{E} = \{\hat{i} = (1, 0, 0), \hat{j} = (0, 1, 0), \hat{k} = (0, 0, 1)\}$ Ad esempio se prendiamo un punto $P = (4, 6, 0)$ questo equivale a:

$$P = 4 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + 6 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + 0 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = (4, 6, 0)_{\mathcal{E}}$$

Rappresentato in uno spazio tridimensionale si vede facilmente come la posizione del punto P non è altro che la somma dei vettori della base opportunamente dilatati o contratti da opportuni coefficienti che prendono il nome di *coordinate*. Esprimere un punto con delle coordinate è molto utile e visto che la combinazione dei vettori base è unica per ogni punto (cfr. T 1.2), nessun altro punto avrà le stesse coordinate.



Questa non è l'unico modo di individuare un punto, in effetti, nulla ci vieta di esprimere questo stesso punto rispetto una base diversa come ad esempio $\mathcal{B} = \{\mathbf{v}_1 = (1, 1, 1) \quad \mathbf{v}_2 = (1, 2, -1) \quad \mathbf{v}_3 = (3, -1, -1)\}$, il fatto di cambiare base non comporta in alcun modo un cambio nella posizione del punto come si può apprezzare dal grafico.



Il punto P sarà dunque rappresentato nella base \mathcal{B} come:

$$P = 2 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + 2 \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} + 0 \begin{bmatrix} 3 \\ -1 \\ -1 \end{bmatrix} = (2, 2, 0)_{\mathcal{B}}$$

Dunque entrambe rappresentano lo stesso punto anche se le basi o i "campioni" base sono diversi. Il problema è che si possono formulare infinite basi e dunque infinite forme diverse per rappresentare lo stesso punto. Se, ad esempio, due persone analizzano lo stesso fenomeno e ognuna ha una

base diversa non potranno confrontare le loro misure dato che i numeri ottenuti differiranno di conseguenza, nasce dunque la necessità di fare la conversione delle "misure" o il cambiamento da una base a un'altra.

Vediamo dunque di analizzare questo problema nell'ambito delle applicazioni lineari.

Definizione 5.7. Dato uno spazio vettoriale V e due sue basi \mathcal{B} e \mathcal{B}' chiamiamo **matrice di cambiamento di base** da \mathcal{B} a \mathcal{B}' la matrice $M_{\mathcal{B}\mathcal{B}'}(\text{id}_V)$. Questa matrice ci permetterà dunque di cambiare tutte le misure espresse nella base \mathcal{B} a vettori della base \mathcal{B}' . Per ricavare questa matrice ci basta esprimere i vettori della base \mathcal{B} come combinazioni lineari dei vettori di \mathcal{B}' cioè

$$\begin{aligned}\mathbf{v}_1 &= a_{11}\mathbf{v}'_1 + a_{12}\mathbf{v}'_2 + \dots + a_{1n}\mathbf{v}'_n \\ \mathbf{v}_2 &= a_{21}\mathbf{v}'_1 + a_{22}\mathbf{v}'_2 + \dots + a_{2n}\mathbf{v}'_n \\ &\vdots \\ \mathbf{v}_n &= a_{n1}\mathbf{v}'_1 + a_{n2}\mathbf{v}'_2 + \dots + a_{nn}\mathbf{v}'_n\end{aligned}$$

La matrice di passaggio si ricava disponendo i coefficienti delle combinazioni lineari per colonna in una matrice.

$$M_{\mathcal{B}\mathcal{B}'} = \begin{bmatrix} a_{11} & a_{21} & \dots & a_{1n} \\ a_{12} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{nn} \end{bmatrix}$$

Cioè per convertire ad esempio da gradi centigradi a fahrenheit ci basta trovare il legame che c'è fra un solo grado espresso in una scala rispetto all'altra con il fine di trovare una "formula" con la quale si potranno convertire qualsivoglia temperatura all'altra scala. Consideriamo ora l'esempio di prima in cui abbiamo espresso il punto P con due basi diverse, \mathcal{E} e \mathcal{B} ; calcoliamo la matrice di cambiamento dalla \mathcal{B} alla base canonica \mathcal{E} , come abbiamo visto prima, basta porre i vettori della prima base espresi come immagine dell'altra base e posizionati come colonne della matrice di cambiamento, dato che la base "d'arrivo" è canonica i vettori rimangono invariati.

$$M_{\mathcal{B}\mathcal{E}}(\text{id}) = \begin{bmatrix} 1 & 1 & 3 \\ 1 & 2 & -1 \\ 1 & -1 & -1 \end{bmatrix}$$

Dal teorema di rappresentazione possiamo trovare il punto espresso nella base canonica come $M_{\mathcal{B}\mathcal{E}}(\text{id}) \cdot \mathbf{v}_{\mathcal{B}} = \mathbf{v}_{\mathcal{E}}$

$$\begin{bmatrix} 1 & 1 & 3 \\ 1 & 2 & -1 \\ 1 & -1 & -1 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ 2 \\ 0 \end{bmatrix} = (4, 6, 0)_{\mathcal{E}}$$

Calcoliamo ora la matrice di cambiamento dalla base canonica alla base \mathcal{B} dobbiamo quindi esprimere i vettori di \mathcal{E} come immagine della base \mathcal{B} quindi dobbiamo risolvere

$$\begin{cases} a\mathbf{v}_1 + b\mathbf{v}_2 + c\mathbf{v}_3 = (1, 0, 0) \\ d\mathbf{v}_1 + e\mathbf{v}_2 + f\mathbf{v}_3 = (0, 1, 0) \\ g\mathbf{v}_1 + h\mathbf{v}_2 + i\mathbf{v}_3 = (0, 0, 1) \end{cases}$$

che può essere visto come:

$$\begin{cases} a + b + 3c = 1 \\ a + 2b - c = 0 \\ a - b - c = 0 \end{cases} \quad \begin{cases} d + e + 3f = 0 \\ d + 2e - f = 1 \\ d - e - f = 0 \end{cases} \quad \begin{cases} g + h + 3i = 0 \\ g + 2h - i = 0 \\ g - h - i = 1 \end{cases}$$

Che ha come soluzioni:

$$\begin{cases} a = 1/4 \\ b = 0 \\ c = 1/4 \end{cases} \quad \begin{cases} d = 1/6 \\ e = 1/3 \\ f = -1/6 \end{cases} \quad \begin{cases} g = 7/12 \\ h = -1/3 \\ i = -1/12 \end{cases}$$

Quindi in conclusione abbiamo

$$M_{\mathcal{E}\mathcal{B}}(\text{id}) = \begin{bmatrix} 1/4 & 1/6 & 7/12 \\ 0 & 1/3 & -1/3 \\ 1/4 & -1/6 & -1/12 \end{bmatrix}$$

Dal teorema di rappresentazione possiamo trovare il punto espresso nella base \mathcal{B} come il prodotto della matrice di cambiamento per il vettore nella base canonica cioè $M_{\mathcal{E}\mathcal{B}}(\text{id}) \cdot \mathbf{v}_{\mathcal{E}} = \mathbf{v}_{\mathcal{B}}$:

$$\begin{bmatrix} 1/4 & 1/6 & 7/12 \\ 0 & 1/3 & -1/3 \\ 1/4 & -1/6 & -1/12 \end{bmatrix} \cdot \begin{bmatrix} 4 \\ 6 \\ 0 \end{bmatrix} = (2, 2, 0)_{\mathcal{E}}$$

Questa matrice però gode di una bellissima proprietà, infatti questa matrice risulta essere proprio l'inversa della prima matrice trovata, verifichiamo applicando Gauss-Jordan ($M_{\mathcal{B}\mathcal{E}}|I$)

$$\begin{array}{l} \left[\begin{array}{ccc|cc} 1 & 1 & 3 & 1 & 0 & 0 \\ 1 & 2 & -1 & 0 & 1 & 0 \\ 1 & -1 & -1 & 0 & 0 & 1 \end{array} \right] \xrightarrow{\substack{III-I \\ II-I}} \left[\begin{array}{ccc|cc} 1 & 1 & 3 & 1 & 0 & 0 \\ 0 & 1 & -4 & -1 & 1 & 0 \\ 0 & -2 & -4 & -1 & 0 & 1 \end{array} \right] \xrightarrow{\substack{III+2II \\ I-III}} \left[\begin{array}{ccc|cc} 1 & 1 & 3 & 1 & 0 & 0 \\ 0 & 1 & -4 & -1 & 1 & 0 \\ 0 & 0 & -12 & -3 & 2 & 1 \end{array} \right] \\ \\ \left[\begin{array}{ccc|cc} 1 & 1 & 3 & 1 & 0 & 0 \\ 0 & 1 & -4 & -1 & 1 & 0 \\ 0 & 0 & 1 & 1/4 & -1/6 & -1/12 \end{array} \right] \xrightarrow{\substack{II+4III \\ I-3III}} \left[\begin{array}{ccc|cc} 1 & 1 & 0 & 1/4 & 1/2 & 1/4 \\ 0 & 1 & 0 & 0 & 1/3 & -1/3 \\ 0 & 0 & 1 & 1/4 & -1/6 & -1/12 \end{array} \right] \xrightarrow{I-II} \\ \\ \xrightarrow{I-II} \left[\begin{array}{ccc|cc} 1 & 0 & 0 & 1/4 & 1/6 & 7/12 \\ 0 & 1 & 0 & 0 & 1/3 & -1/3 \\ 0 & 0 & 1 & 1/4 & -1/6 & -1/12 \end{array} \right] \end{array}$$

Quindi possiamo calcolare l'altra matrice di cambiamento invertendo la prima matrice, anche se questa procedura è molto utile, non la si può applicare sempre, questa la possiamo fare solo se la matrice è invertibile e si ricorda che questa risulta essere invertibile se e solo se $\text{rk } A = n$.

A volte può succedere che non si conosca la matrice rappresentativa di una applicazione lineare f nella base che abbiamo scelto, ad esempio consideriamo in questo caso la canonica \mathcal{E} , ma che si conosca la matrice rappresentativa espressa in un'altra base \mathcal{B} . Per far agire la applicazione su dei vettori espressi nella base canonica possiamo esprimere i vettori $v_{\mathcal{E}}$ come vettori di \mathcal{B} con la matrice

di passaggio $M_{\mathcal{E}\mathcal{B}}$, far agire la applicazione e poi trasformare i risultati della base \mathcal{B} di nuovo in vettori $v_{\mathcal{E}}$ con la matrice di passaggio $M_{\mathcal{B}\mathcal{E}}$ cioè:

$$M_{\mathcal{E}\mathcal{B}} \rightarrow M(f)_{\mathcal{B}\mathcal{B}} \rightarrow M_{\mathcal{B}\mathcal{E}} \rightarrow M(f)_{\mathcal{E}\mathcal{E}}$$

Invece di fare tutti questi passaggi ogni volta per ricavare l'immagine della applicazione f nella nostra base agendo con un vettore nella stessa base possiamo ricavare una volta per tutte la matrice rappresentativa dell'applicazione che si ricava come segue:

$$M(f)_{\mathcal{E}\mathcal{E}} = M_{\mathcal{B}\mathcal{E}} \cdot M(f)_{\mathcal{B}\mathcal{B}} \cdot M_{\mathcal{E}\mathcal{B}}$$

Si noti che questa espressione non è sempre la stessa poiché magari si deve cambiare non una ma due o più volte di base quindi ci saranno dei passaggi diversi.

Teorema 5.12. Possiamo riscrivere dunque il teorema 5.11 Sia un'applicazione lineare $f : U \rightarrow V$ e $g : V \rightarrow W$, siano $\mathcal{B}, \mathcal{C}, \mathcal{D}$ rispettivamente le loro basi e siano $A = M_{\mathcal{B}\mathcal{C}}$ la matrice rappresentativa dell'applicazione lineare f e $B = M_{\mathcal{C}\mathcal{D}}$ la matrice rappresentativa dell'applicazione lineare g allora $g \circ f$ è lineare e la sua matrice rappresentativa è $C = M_{\mathcal{B}\mathcal{D}}$ ottenuta come

$$M_{\mathcal{B}\mathcal{D}} = M_{\mathcal{C}\mathcal{D}} \cdot M_{\mathcal{B}\mathcal{C}} \quad (5.7)$$

Sia $f : V \rightarrow W$ un'applicazione lineare e siano $M_{\mathcal{B}\mathcal{B}'}(f)$ e $M_{\mathcal{C}\mathcal{C}'}(f)$ le matrici ad essa associate rispetto alle basi \mathcal{B}, \mathcal{C} di V e $\mathcal{B}', \mathcal{C}'$ di W . Un legame fra le matrici è:

$$M_{\mathcal{C}\mathcal{C}'}(f) = M_{\mathcal{C}\mathcal{C}'}(\text{id}_W \circ f \circ \text{id}_V) = M_{\mathcal{B}\mathcal{C}'}(\text{id}_W) \cdot M_{\mathcal{B}\mathcal{B}'}(f) \cdot M_{\mathcal{C}\mathcal{B}}(\text{id}_V)$$

In particolare sia $f : V \rightarrow V$ un endomorfismo e $\mathcal{C} = \mathcal{C}'$ e $\mathcal{B} = \mathcal{B}'$ allora avremo:

$$M_{\mathcal{C}\mathcal{C}}(f) = M_{\mathcal{B}\mathcal{C}}(\text{id}_V) \cdot M_{\mathcal{B}\mathcal{B}}(f) \cdot M_{\mathcal{C}\mathcal{B}}(\text{id}_V)$$

Quindi:

$$M_{\mathcal{C}\mathcal{C}}(f) = P \cdot M_{\mathcal{B}\mathcal{B}}(f) \cdot P^{-1}$$

Definizione 5.8. Sia A una matrice quadrata ($A \in \mathcal{M}_n(\mathbb{R})$), si definisce **traccia** di A e si rappresenta con $\text{tr}(A)$, l'applicazione lineare che somma tutti gli elementi sulla sua diagonale principale:

$$\text{tr}(A) = \sum_{i=1}^n a_{ii}$$

Proprietà 5.1. *Poiché la traccia è un'applicazione lineare:*

$$\text{tr}(A + B) = \text{tr}(A) + \text{tr}(B)$$

$$\text{tr}(kA) = k(\text{tr}(A))$$

Proprietà 5.2. *Come conseguenza immediata del procedimento di moltiplicazione di matrici si ha che: data una matrice A di dimensione $m \times n$ e una seconda matrice B di dimensione $n \times m$, si ha:*

$$\text{tr}(AB) = \text{tr}(BA)$$

Proprietà 5.3. *Una matrice A e la sua trasposta ${}^t A$ hanno la stessa traccia:*

$$\text{tr}(A) = \text{tr}({}^t A)$$

Definizione 5.9. Due matrici $A, B \in \mathbb{R}^{n,n}$ si dicono **simili** se esiste una matrice invertibile $P \in \mathbb{R}^{n,n}$, tale che $B = P^{-1}AP$ o equivalentemente $PB = AP$, in tal caso si scrive $A \sim B$.

La similitudine è una relazione di equivalenza, cioè, è riflessiva, simmetrica, e transitiva.

Proprietà 5.4. (INVARIANZA PER SIMILITUDINE DELLA TRACCIA)

$$\text{Se } A \sim B \text{ allora } \text{tr}(A) = \text{tr}(B)$$

5.2 Complementi

Se consideriamo come spazio vettoriale l'insieme dei polinomi a coefficienti reali di grado al più n , $\mathbb{R}_n[x]$.

$$\mathbb{R}_n[x] := \{a_0 + a_1x + a_2x^2 + \dots + a_nx^n \mid a_i \in \mathbb{R} \ \forall i = 1, \dots, n\}$$

Le operazioni di derivazione e integrazione possono essere espresse come applicazioni lineari che hanno le seguenti matrici:

INTEGRAZIONE	$\begin{bmatrix} 0 & 0 & 0 & 0 & \dots \\ 1 & 0 & 0 & 0 & \dots \\ 0 & 1/2 & 0 & 0 & \dots \\ 0 & 0 & 1/3 & 0 & \dots \\ \vdots & \vdots & \vdots & \ddots & \dots \end{bmatrix}$	DERIVAZIONE	$\begin{bmatrix} 0 & 1 & 0 & 0 & \dots \\ 0 & 0 & 2 & 0 & \dots \\ 0 & 0 & 0 & 3 & \dots \\ 0 & 0 & 0 & 0 & \ddots \\ \vdots & \vdots & \vdots & \vdots & \dots \end{bmatrix}$
---------------------	---	--------------------	--

Si nota facilmente che la composizione della matrice di derivazione e la matrice d'integrazione da come risultato la matrice identità solo nel caso in cui si fa per prima l'integrazione e poi la derivazione, nel altro caso è una matrice identità a meno del primo elemento nella diagonale principale, questo elemento nullo rappresenta la costante $+C$ da aggiungere nell'integrazione

Determinante

Il determinante è un numero che può essere ricavato dalle matrici quadrate $A \in \mathcal{M}_n(\mathbb{R})$ e che ci aiuta ad identificare alcune proprietà intrinseche delle matrici alle quali è applicato. Una delle ragioni principali per cui lo si calcola, è per conoscere se il rango della matrice è massimo o no, dato che esiste la seguente relazione tra i due:

$$\forall A \in \mathcal{M}_n(\mathbb{R}) : \det(A) = 0 \iff \text{rk} \neq n \quad (6.1)$$

Equivalentemente:

$$\forall A \in \mathcal{M}_n(\mathbb{R}) : \text{rk} = n \iff \det(A) \neq 0$$

Definizione 6.1. Chiameremo **matrice singolare**, una matrice quadrata con determinante nullo, oppure, analogamente, una matrice quadrata il cui rango non è massimo. Per quest'ultimo motivo nessuna matrice singolare è invertibile.

Basandoci nei capitoli precedenti, sappiamo che le matrici quadrate possono rappresentare un endomorfismo in $\mathbb{R}^{n,n}$, ragione per cui possiamo dire che il determinante è una applicazione lineare formata come segue:

$$\det : \mathbb{R}^{n,n} \rightarrow \mathbb{R}$$

Ovvero, il determinante è una applicazione che va dal dominio delle matrici quadrate all'insieme dei numeri reali.

Per definire formalmente il determinante dobbiamo dapprima definire un paio di concetti che ci aiuteranno a costruire la definizione di determinante.

Definizione 6.2. Sia $n \in \mathbb{N}$, $n \geq 1$ chiamiamo **permutazione** su n oggetti un riordinamento degli indici da 1 a n e la indichiamo con σ .

Nello specifico chiamiamo **permutazione identità** una permutazione d'oggetti senza alcun riordinamento.

ESEMPIO Consideriamo $n = 4$:

$(1, 2, 3, 4)$ è una permutazione identità

$(2, 3, 4, 1)$ è una permutazione

Indicheremo con S_n l'insieme di tutte le permutazioni su n oggetti, è doveroso sottolineare che la cardinalità di queste, coincide con il fattoriale di n ossia $|S_n| = n!$

Definizione 6.3. Data una permutazione $\sigma = \{\sigma(1), \sigma(2), \dots, \sigma(n)\}$ si chiama **inversione** una coppia $\{\sigma(i), \sigma(j)\}$ tale che $\sigma(i) > \sigma(j)$. Chiamiamo inoltre $\text{inv}(\sigma)$ il numero delle inversioni di σ . In pratica le inversioni rappresentano quanto una permutazione sia “disordinata” e la si misura con quante volte un numero più grande compare a destra degli altri numeri.

Definizione 6.4. Una permutazione dei numeri naturali $1, 2, \dots, n$ è detta di **classe pari o di classe dispari** a seconda se il numero totale di inversioni che presenta è pari o dispari (la permutazione identità, che ha 0 inversioni, si considera di classe pari in quanto lo zero è multiplo di $2 \cdot 0 = 2 \cdot 0$) .

ESEMPIO

Consideriamo l'esempio di prima $(2, 3, 4, 1)$ il numero di inversioni è 3 poiché $(2 > 1), (3 > 1), (4 > 1)$

Definizione 6.5. Chiamiamo con grande fantasia **scambio** lo scambio tra due indici di una permutazione.

ESEMPIO

Consideriamo $(3, 2, 1)$, il numero di inversioni è 3 poiché $(3 > 2), (3 > 1), (2 > 1)$; il numero di scambi però non è unico, in effetti, possiamo scambiare come segue:

$$(3, 2, 1) \xrightarrow{3,2} (2, 3, 1) \xrightarrow{3,1} (2, 1, 3) \xrightarrow{2,1} (1, 2, 3) \quad \text{scambi: 3}$$

Ma potevamo riordinare in un modo molto più efficiente:

$$(3, 2, 1) \xrightarrow{3,1} (1, 2, 3) \quad \text{scambi: 1}$$

Le definizioni di inversioni e scambi sono ben diverse, ma dall'esempio precedente si intravede un legame tra scambi e inversioni, infatti nell'esempio si possono osservare i due teoremi che seguono:

Teorema 6.1. Il numero di scambi sufficienti per riordinare una permutazione, cioè, per trasformarla nell'identità, è uguale al numero delle inversioni.

Teorema 6.2. Il numero di scambi necessari a riordinare una permutazione ha la stessa parità del numero delle inversioni.

Definizione 6.6. Sia $A = (a_{ij}) \in \mathcal{M}_n(\mathbb{R})$ ossia una matrice così fatta :

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$$

Chiamiamo **prodotto associato alla matrice**, il prodotto degli elementi della matrice con indici tutti diversi (che quindi non stanno mai sulla stessa riga o colonna), per $(-1)^{(t+u)}$ dove t e u sono le inversioni delle permutazioni degli indici degli elementi della matrice, ossia consideriamo tutti i prodotti formati da n coefficienti di A scelti in modo che due di essi non appartengano mai ne alla stessa riga ne alla stessa colonna.

Definizione 6.7. FORMULA DI LEIBNIZ

Chiamiamo determinante il numero ottenuto dalla somma dei prodotti associati di una matrice $n \times n$

$$\det(A) := \sum_{\sigma \in S_n} (-1)^{\text{inv}(\sigma)} \cdot a_{1,\sigma(1)} \cdot a_{2,\sigma(2)} \cdots \cdot a_{n,\sigma(n)}$$

Come si può osservare, il $\det(A)$ è la somma di $n!$ addendi ognuno dei quali è il prodotto di $n+1$ fattori, e S_n come era già stato detto precedentemente, è l'insieme di tutte le permutazioni σ dell'insieme numerico $\{1, 2, \dots, n\}$. Questa formula, nota come formula di Leibniz può essere riscritta come :

$$\det(A) := \sum_{\sigma \in S_n} \text{sgn}(\sigma) \prod_{i=1}^n a_{i,\sigma(i)} \quad (6.2)$$

In quest'ultima variante il simbolo $\text{sgn}(\sigma)$ denota il segno della permutazione ove consideriamo +1 se σ è una permutazione pari e -1 se è permutazione dispari.

Questa definizione vista oggettivamente è complicata, ma applicarla risulta veramente facile specialmente nei casi (2×2) o (3×3) .

Consideriamo una generica matrice $A \in \mathcal{M}_2(\mathbb{R})$ ossia 2×2 . Il determinante della matrice è allora uguale a:

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \sum_{\sigma \in S_2} \text{sgn}(\sigma) \prod_{i=1}^2 A_{i,\sigma(i)} = a_{11}a_{22} - a_{12}a_{21}$$

Nei casi 2×2 , possiamo risparmiarci sia il calcolo del numero di permutazioni per ottenere il segno che il prodotto associato, dato che si può giungere allo stesso risultato moltiplicando gli elementi della diagonale principale e sottraendo il prodotto degli elementi della diagonale secondaria:

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \implies \det(A) = ad - bc \quad (6.3)$$

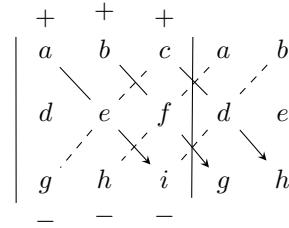
Anche nei casi 3×3 è possibile una semplificazione nel calcolo del determinante, questa semplificazione è nota come **regola di Sarrus** in onore al matematico francese Pierre-Frédéric Sarrus. Questa regola ci dice che il determinante di una matrice 3×3 è uguale a:

$$A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} \implies \det(A) = aei + bfg + cdh - gec - hfa - idb$$

Questa espressione può essere ricavata facilmente con una matrice orlata che si ottiene affiancano a destra, le prime due colonne della matrice di partenza, o alternativamente, ripetendo le prime due righe alla fine della matrice. Per ottenere l'espressione del determinante basta fare la differenza tra la somma dei prodotti delle 3 diagonali principali (che partono dall'alto a sinistra) e la somma dei prodotti delle diagonali secondarie (che partono in basso a sinistra). Infatti ripetendo a destra della matrice le sue prime due colonne si ottiene una matrice fatta così:

Come si apprezza nella figura, le diagonali principali vanno precedute da un segno più mentre le diagonali secondarie da un segno meno.

$$\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} \begin{matrix} a & b \\ d & e \\ g & h \end{matrix} \quad (6.4)$$



Se si procede con la formula di Leibniz si ottiene che nei casi 3×3 il determinante è uguale a:

$$\sum_{\sigma \in S_3} \text{sgn}(\sigma) \prod_{i=1}^3 A_{i,\sigma(i)} = a_{11}a_{22}a_{33} + a_{13}a_{32}a_{21} + a_{12}a_{23}a_{31} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}$$

Poiché queste formule semplificano tanto i conti, si potrebbe pensare che, ad esempio, la regola di Sarrus, sia generalizzabile a matrici più grandi, ma purtroppo non è così, non esistono altre regole analoghe alle precedenti per le matrici $n \times n$ con $n \geq 4$

6.1 Proprietà del determinante

Il determinante ha molte proprietà, cominciamo ad analizzare le più importanti.

Teorema 6.3. Il determinante è un invariante per trasposizione

Questo teorema ci dice che, scegliere prima la riga o la colonna per svolgere i prodotti associati non cambia il determinante. Questo è il motivo per cui tutte le seguenti proprietà valgono sia per le righe che per le colonne.

Proprietà 6.1. Sia $A \in \mathcal{M}_n(\mathbb{R})$ allora:

- (1) Se B si ottiene da A scambiando tra loro due righe/colonne allora $\det(B) = -\det(A)$ (questo comportamento lo chiamiamo di tipo **alternante**).
- (2) Se una riga o colonna è nulla allora $\det(A) = 0$
- (3) Se due righe o colonne sono uguali allora $\det(A) = 0$
- (4) Se due righe o colonne di A sono proporzionali allora $\det(A) = 0$
- (5) Moltiplicando una colonna o riga di A per uno scalare λ si ottiene una matrice B tale che $\det(B) = \lambda \cdot \det(A)$
- (6) Moltiplicando tutte le colonne o righe di A per uno scalare λ si ottiene una matrice B tale che $\det(B) = \lambda^n \cdot \det(A)$
- (7) Il det di una matrice triangolare (alta, bassa o diagonale) è il prodotto degli elementi della diagonale principale.

Corollario 6.3.1. Una matrice quadrata e la sua matrice trasposta hanno lo stesso determinante:

$$\det(A^T) = \det(A) \quad (6.5)$$

Teorema 6.4. Sia $A = [\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n] \in \mathcal{M}_n(\mathbb{R})$ con $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ colonne di A , sia inoltre $\mathbf{a}_k = B + C$, allora:

$$\det(A) = \det[\mathbf{a}_1, \dots, \mathbf{a}_{k-1}, B, \mathbf{a}_{k+1} \dots, \mathbf{a}_n] + \det[\mathbf{a}_1, \dots, \mathbf{a}_{k-1}, C, \mathbf{a}_{k+1} \dots, \mathbf{a}_n]$$

Il teorema precedente si esprime dicendo che il determinante è una funzione multilineare sulle colonne, in particolare questo teorema si può generalizzare come segue:

Teorema 6.5 (MULTILINEARITÀ DEL DETERMINANTE). $\forall 1 \leq k \leq n, \forall \lambda_1, \lambda_2 \in \mathbb{R}, \forall \mathbf{v}, \mathbf{w} \in \mathbb{R}^n$:

$$\det(A) = \det[\mathbf{a}_1, \dots, \lambda_1 \mathbf{v} + \lambda_2 \mathbf{w}, \dots, \mathbf{a}_n] = \lambda_1 \det[\mathbf{a}_1, \dots, \mathbf{v}, \dots, \mathbf{a}_n] + \lambda_2 \det[\mathbf{a}_1, \dots, \mathbf{w}, \dots, \mathbf{a}_n]$$

Quello che ci vuole dire questa proprietà, è che se una matrice quadrata ha una colonna k , che risulta essere combinazione lineare di due vettori, \mathbf{v} e \mathbf{w} , allora il determinante della matrice è uguale all'analoga combinazione lineare delle due matrici che hanno rispettivamente il vettore \mathbf{v} e il vettore \mathbf{w} nella colonna k .

ESEMPIO

Consideriamo la matrice :

$$A = \begin{pmatrix} 9 & 3 & 4 \\ 7 & 1 & 4 \\ 2 & 5 & 1 \end{pmatrix}$$

La prima colonna è una combinazione lineare di due vettori, ad esempio $\mathbf{a} = (1, 1, 1)$ e $\mathbf{b} = (8, 6, 1)$, ma non solo, è combinazione lineare anche di $\mathbf{v} = (3, 3, 1)$ e $\mathbf{w} = (3, 1, 0)$ per dirne alcuni, poiché in realtà la colonna \mathbf{a}_k può essere costruita come combinazione lineare di infiniti vettori. Per i vettori \mathbf{v} e \mathbf{w} risulta infatti che $\mathbf{a}_k = 2\mathbf{v} + 1\mathbf{w}$. Grazie alla multi-linearità possiamo dunque dire che il determinante di A è uguale a 2 volte il det della matrice $A_{\mathbf{v}}$ più 1 volta il det della matrice $A_{\mathbf{w}}$ ossia:

$$A_{\mathbf{v}} = \begin{pmatrix} 3 & 3 & 4 \\ 3 & 1 & 4 \\ 1 & 5 & 1 \end{pmatrix} \quad A_{\mathbf{w}} = \begin{pmatrix} 3 & 3 & 4 \\ 1 & 1 & 4 \\ 0 & 5 & 1 \end{pmatrix}$$

Calcolando il determinante con la regola di Sarrus abbiamo che i determinanti sono:

$$\det(A) = (9 \cdot 1 \cdot 1) + (3 \cdot 4 \cdot 2) + (4 \cdot 7 \cdot 5) - (4 \cdot 1 \cdot 2) - (9 \cdot 4 \cdot 5) - (3 \cdot 7 \cdot 1) = -36$$

$$\det(A_{\mathbf{v}}) = (3 \cdot 1 \cdot 1) + (3 \cdot 4 \cdot 1) + (4 \cdot 3 \cdot 5) - (4 \cdot 1 \cdot 1) - (3 \cdot 4 \cdot 5) - (3 \cdot 3 \cdot 1) = 2$$

$$\det(A_{\mathbf{w}}) = (3 \cdot 1 \cdot 1) + (3 \cdot 4 \cdot 0) + (4 \cdot 1 \cdot 5) - (4 \cdot 1 \cdot 0) - (3 \cdot 4 \cdot 5) - (3 \cdot 1 \cdot 1) = -40$$

vediamo che si verifica:

$$\det(A) = 2 \det(A_{\mathbf{v}}) + 1 \det(A_{\mathbf{w}}) = 2 \cdot 2 - 40 = -36$$

Corollario 6.5.1. Se si somma ad una colonna di A una combinazione lineare delle altre colonne, il determinante non cambia.

Proprietà 6.2 (NORMALIZZATA). Il determinante della matrice identità:

$$\det(I) = 1$$

Questa è una proprietà che si può osservare facilmente dato che, essendo I una matrice diagonale il det è il prodotto degli elementi della diagonale principale, ed essa è composta da tutti 1.

Teorema 6.6. Il determinante è l'unica applicazione/funzione) $\mathbb{R}^{n,n} \rightarrow \mathbb{R}$ tale che:

1. È multilineare.
2. È alternante.
3. È normalizzata.

Definizione 6.8. Sia A una matrice anche non quadrata, si chiama **minore** estratto da A il determinante di una qualunque sottomatrice quadrata di A e chiamiamo **ordine** la sua dimensione.

ESEMPIO

$$\left(\begin{array}{cccc|c} 4 & 3 & 1 & 6 \\ 5 & 2 & 4 & 3 \\ \hline 7 & 3 & 2 & 1 \\ \hline 8 & 6 & 1 & 9 \end{array} \right)$$

Due minori della stessa matrice, a sinistra un minore di ordine 2 e a destra uno di ordine 3.

$$\left(\begin{array}{cccc|c} 4 & 3 & 1 & 6 \\ 5 & 2 & 4 & 3 \\ \hline 7 & 3 & 2 & 1 \\ \hline 8 & 6 & 1 & 9 \end{array} \right)$$

Definizione 6.9. Se A è una matrice quadrata e a_{ij} è il coefficiente di A in posizione i, j , si chiama **minore complementare** di a_{ij} il determinante A_{ij} della matrice che si ottiene da A eliminando la i -esima riga e la j -esima colonna.

Nell'esempio precedente il minore a destra coincide con il minore complementare di a_{41}

Definizione 6.10. Chiamiamo **complemento algebrico** o **cofattore** di a_{ij} il numero:

$$A_{ij} = (-1)^{i+j} \det(A_{ij}) \quad (6.6)$$

Teorema 6.7 (I TEOREMA DI LAPLACE). Il determinante di una matrice A è uguale alla somma dei prodotti degli elementi di una qualsiasi colonna di A per i relativi complementi algebrici.

$$\det(A) := \sum_{i=1}^n (-1)^{i+j} a_{ij} \det(A_{ij}) \quad (6.7)$$

Questa definizione di determinante è “ricorsiva” dato che per calcolare il determinante di matrici di grandi dimensioni, questa rimanda al calcolo del determinante di una matrice quadrata di ordine $n-1$, ripetendo la procedura fintantoché le matrici diventano casi che sappiamo risolvere facilmente.

ESEMPIO

Consideriamo la matrice :

$$A = \begin{pmatrix} 5 & 1 & 2 \\ 8 & 9 & 3 \\ 1 & 7 & 2 \end{pmatrix}$$

Per calcolare il determinante con la formula di Laplace dobbiamo scegliere una colonna o riga, in questo caso sceglieremo arbitrariamente la prima colonna, il determinante sarà dunque uguale a:

$$\det(A) = 5 \cdot (-1)^{(1+1)} \cdot \det \begin{pmatrix} 9 & 3 \\ 7 & 2 \end{pmatrix} + 8 \cdot (-1)^{(2+1)} \cdot \det \begin{pmatrix} 1 & 2 \\ 7 & 2 \end{pmatrix} + 1 \cdot (-1)^{(3+1)} \cdot \det \begin{pmatrix} 1 & 2 \\ 9 & 3 \end{pmatrix}$$

$$\det(A) = (5 \cdot 1 \cdot -3) + (8 \cdot -1 \cdot -12) + (1 \cdot 1 \cdot -15) = 66$$

Teorema 6.8 (TEOREMA DI BINET). Siano $A, B \in \mathcal{M}_n(\mathbb{R})$ allora:

$$\det(A \cdot B) = \det(A) \cdot \det(B) \quad (6.8)$$

ESEMPIO

Consideriamo le matrici :

$$A = \begin{pmatrix} 2 & 3 \\ 4 & 9 \end{pmatrix} \quad \text{e} \quad B = \begin{pmatrix} 1 & 6 \\ 1 & 5 \end{pmatrix} \quad \Rightarrow A \cdot B = \begin{pmatrix} 5 & 27 \\ 13 & 69 \end{pmatrix}$$

Calcoliamo i determinanti: $\det(A) = 6$

$$\det(B) = -1$$

$$\det(A \cdot B) = 345 - 351 = -6$$

Infatti abbiamo che $6 \cdot -1 = -6$

Questo teorema può essere esteso a un numero infinito di prodotti.

$$\det(A \cdot B \cdot C \cdots) = \det(A) \cdot \det(B) \cdot \det(C) \cdots$$

Corollario 6.8.1. Se A è invertibile allora:

$$\det(A^{-1}) = \frac{1}{\det(A)} \quad (6.9)$$

Questa è una conseguenza immediata poiché se consideriamo $\det(A \cdot A^{-1})$, questo per definizione di inversa sappiamo essere uguale a $\det(I)$ che sappiamo che è uguale a 1; per il teorema di Binet abbiamo $\det(A) \cdot \det(A^{-1}) = 1$, con questa espressione si ricava la precedente.

Corollario 6.8.2. Se $A \sim B$ allora:

$$\det(A) = \det(B) \quad (6.10)$$

Questo corollario è anche immediato, in effetti per definizione di similitudine sappiamo che $A = P^{-1} \cdot B \cdot P$ inoltre per l'appena visto teorema di Binet abbiamo $\det(A) = \det(P^{-1}) \cdot \det(B) \cdot \det(P)$ che per il corollario precedente sappiamo che $\det(P^{-1}) = \frac{1}{\det(P)}$ per cui abbiamo un'espressione siffatta:

$$\det(A) = \frac{1}{\det(P)} \cdot \det(B) \cdot \det(P)$$

il $\det(P)$ e $\frac{1}{\det(P)}$ si semplificano a vicenda rimanendo come unico termine il $\det(B)$.

6.2 Significato geometrico del determinante

Dato che il determinante è una funzione che agisce su uno spazio vettoriale (applicazione) si potrebbe pensare che esista una relazione, ad esempio in \mathbb{R}^n , fra il determinante di una matrice e ciò che rappresenta la matrice; ricordiamo che una matrice $n \times n$ può rappresentare n vettori, n piani, un endomorfismo, ecc. In realtà esiste davvero un legame fra le matrici e il determinante. Consideriamo che la matrice rappresenti un insieme di vettori, dove gli elementi di ogni riga sono le coordinate di un vettore appartenente a \mathbb{R}^n .

Teorema 6.9. Siano $\mathbf{v}_1, \dots, \mathbf{v}_n \in \mathbb{R}^n$ allora il “volume con segno” del prisma di lati $\mathbf{v}_1, \dots, \mathbf{v}_n \in \mathbb{R}^n$ è uguale al $\det(\mathbf{v}_1, \dots, \mathbf{v}_n)$.

Il segno sta a dire che, se ad esempio, il vettore i -esimo si sovrappone a un altro vettore k -esimo con $k < i$, allora l’area cambia di segno, cioè se cambio l’ordine dei vettori, l’area cambia di segno.

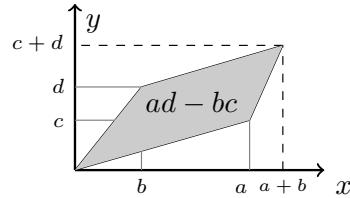
Il valore assoluto del $\det(A)$ è il fattore con cui vengono modificati i volumi degli oggetti contenuti nello spazio.

Consideriamo dunque cosa rappresenta il determinante in \mathbb{R}^2 .

Prendiamo una generica matrice 2×2 :

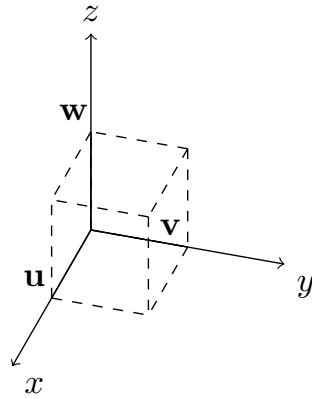
$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \text{dove:} \quad \mathbf{v} = (a, c) \text{ e } \mathbf{w} = (b, d)$$

Sappiamo che il determinante per una matrice 2×2 è uguale a $ad - bc$, in questo caso quest’espressione rappresenta l’area del parallelogramma con vertici in $(0, 0)$, (a, c) , (b, d) e $(a+b, c+d)$ ossia nei punti $O, \mathbf{v}, \mathbf{w}, \mathbf{v}+\mathbf{w}$:



Questa proprietà geometrica, come asserisce il teorema, si estende anche a spazi di dimensioni maggiori di 2.

Consideriamo dunque cosa rappresenta il determinante in \mathbb{R}^3 . In questo caso il determinante della matrice rappresenta il volume del prisma formato da 3 vettori: $\mathbf{u}, \mathbf{v}, \mathbf{w}$



In particolare gode di tre proprietà:

1. Il volume di un prisma di lati $\mathbf{e}_1, \dots, \mathbf{e}_n$ è 1 (normalizzato)
2. Il volume cambia segno se scambio due lati

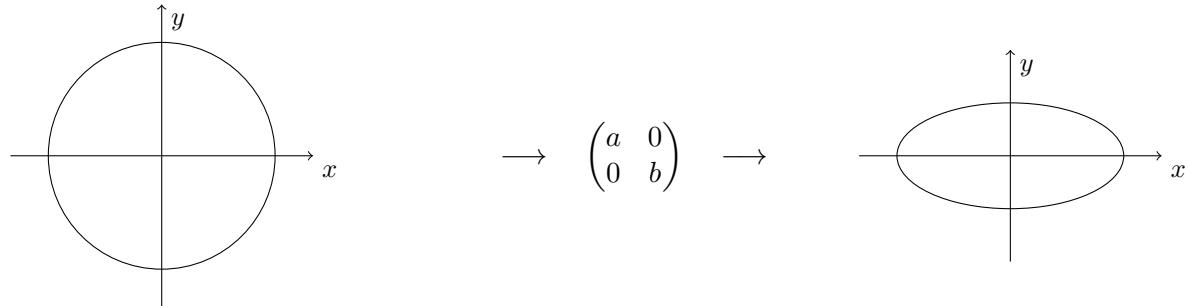
3. Se $\mathbf{v}_1, \dots, \mathbf{v}_n$ lati e $\mathbf{v}_1 = \lambda_1 \mathbf{u}_1 + \lambda_2 \mathbf{u}_2$ allora:

$$\text{vol}(\mathbf{v}_1, \dots, \mathbf{v}_n) = \lambda_1 \text{vol}(\mathbf{u}_1, \mathbf{v}_2, \dots, \mathbf{v}_n) + \lambda_2 \text{vol}(\mathbf{u}_2, \mathbf{v}_2, \dots, \mathbf{v}_n)$$

Corollario 6.9.1. Se $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ è una applicazione tale che $\mathbf{v} \rightarrow A \cdot \mathbf{v}$, dove A è la matrice rappresentativa di f rispetto ad \mathcal{E} . Sia Q un $l-n$ cubo di lati $\mathbf{e}_1, \dots, \mathbf{e}_n$. Sappiamo che il $\text{vol}(Q) = 1$ il volume dell'immagine di Q ossia $\text{vol}(f(Q))$ dipende dal $\det(A)$. In generale, se Q è un insieme di \mathbb{R}^n di volume

$$\text{vol}(Q) \quad \text{allora} \quad \text{vol}(f(Q)) = \text{vol}(Q) \cdot \det(A)$$

ESEMPIO L'area dell'ellisse è la trasformazione dell'area di un cerchio con una matrice diagonale



$$\{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\} \longrightarrow (u, v) = (ax, by) \longrightarrow \left\{ (u, v) \in \mathbb{R}^2 \mid \frac{u^2}{a^2} + \frac{v^2}{b^2} = 1 \right\}$$

6.3 Applicazioni del determinante

6.3.1 Matrici non quadrate

Come abbiamo visto prima il determinante è definito solo per le matrici quadrate, quindi se ci occorre conoscere il rango di una matrice non quadrata, in teoria, non possiamo utilizzare il concetto di determinante. Ciononostante, ci sono dei teoremi che ci permettono di estendere la sua applicazione in matrici che risultano anche non quadrate, ma non in maniera diretta poiché queste matrici non soddisfanno la condizione necessaria perché il determinante possa essere calcolato, calcoleremo dunque i determinanti non sulle matrici non quadrate ma sui minori delle stesse, ove è possibile applicarlo.

Teorema 6.10. Sia $A \in \mathcal{M}(n, m, \mathbb{R})$ allora il rango di A è l'ordine del massimo minore non nullo di A

Questo teorema ci permette di calcolare il rango di una matrice non quadrata con il solo uso del determinante, basandosi sul fatto che se le righe sono linearmente dipendenti allora anche tutti i coefficienti della sottomatrice si dovrebbero annullare. È necessario avere bene in mente che se si trova un minore di dimensione $n \times n$ con $\det \neq 0$, si può subito giungere alla conclusione che il $\text{rk} \geq n$, ma, se un minore ha $\det = 0$, non possiamo affermare nulla sul rango, bisognerebbe controllare tutti i possibili minori di dimensione $n \times n$, e, nel caso si annullino tutti, possiamo allora concludere che $\text{rk} = n - 1$

ESEMPIO

$$\left(\begin{array}{cccc} 1 & 0 & 1 & 6 \\ -1 & 2 & 1 & 0 \end{array} \right) \longrightarrow \left(\begin{array}{cc|cc} 1 & 0 & 1 & 6 \\ -1 & 2 & 1 & 0 \end{array} \right)$$

Vediamo subito che il determinante del minore è $\neq 0$, per cui si evince che il $\text{rk} = 2$ senza fare la riduzione a scala!

Il teorema precedente risulta essere molto utile per matrici relativamente piccole, ma, per matrici più grandi come una 5×5 , per verificare che $\text{rk} = 5$ dovremmo verificare che ognuno dei 25 minori 4×4 si annulla o nell'ipotesi che uno di questi determinanti fosse non singolare per verificare che ha $\text{rk}=3$ dovremmo verificare 100 minori 3×3 !

Teorema 6.11 (TEOREMA DI KRONECKER o dei minori orlati). Sia M una sottomatrice $p \times p$ di $A \in \mathcal{M}(n, m, \mathbb{R})$ con $p \leq \min\{n, m\}$ e supponiamo che $\det(M) \neq 0$. Allora, il rango di A è uguale a p se soltanto se ogni sottomatrice $(p+1) \times (p+1)$ di A ottenuta orlando (cioè aggiungendo una riga e una colonna a M) ha determinante nullo. Grazie a questo teorema, non occorre controllare tutti i minori contenuti in una matrice, ma solo quelli che orlano un minore di ordine p .

Il determinante può essere di grande aiuto quando si cerca il rango di una matrice dipendente da un parametro.

ESEMPIO. Determiniamo il rango della matrice al variare di k

$$\left(\begin{array}{cccc} k+1 & 2 & 1 & 1 \\ 3 & k & 0 & 2 \\ 3-k & 1+k & 1 & 1 \end{array} \right)$$

Vediamo subito che c'è un minore, non dipendente da un parametro, non nullo per cui si conclude rapidamente che $\text{rk} \geq 2$

$$\left(\begin{array}{cc|cc} k+1 & 2 & 1 & 1 \\ 3 & k & 0 & 2 \\ 3-k & 1+k & 1 & 1 \end{array} \right)$$

Per vedere se la matrice ha $\text{rk}=2$ dobbiamo verificare che ogni sottomatrice orlata abbia $\det = 0$, altrimenti si avrà che $\text{rk}=3$

$$\left(\begin{array}{cccc} k+1 & 2 & 1 & 1 \\ 3 & k & 0 & 2 \\ 3-k & 1+k & 1 & 1 \end{array} \right)$$

La matrice ha $\det = 0$ se ne esiste almeno un valore di k che lo annulli su tutte le matrici orlate.

$$\left(\begin{array}{cc|cc} k+1 & 2 & 1 & 1 \\ 3 & k & 0 & 2 \\ 3-k & 1+k & 1 & 1 \end{array} \right)$$

Per la prima matrice calcoliamo il determinante usando lo sviluppo di Laplace rispetto alla prima riga. $\det = -2(-2) + (-1)(k-2-2k) + 1(k) = 2k-2$.

Per la seconda matrice calcoliamo il determinante usando sempre la prima riga è $\det = (k+1)(-2) + (-1)(3-6+2k) + 3 = 4-4k$

Vediamo dunque che il $\det = 0$ in entrambi i casi per $k = 1$ quindi per $k \neq 1$ la matrice ha $\text{rk}=3$, e per $k = 1$ la matrice ha $\text{rk}=2$

6.3.2 Calcolo dell'inversa

Nei capitoli precedenti abbiamo visto che è possibile calcolare l'inversa di una matrice con il metodo di Gauss-Jordan, o con la soluzione di un sistema lineare di $n \times n$ incognite. Grazie alle proprietà del determinante possiamo determinare l'inversa con un'altro metodo descritto dal seguente teorema.

Teorema 6.12. Sia $A \in \mathcal{M}(n, m, \mathbb{R})$ invertibile (quindi deve avere rango massimo, inoltre, per quanto visto deve avere un determinante diverso da zero). $\forall (1 \leq i, j \leq n)$ sia $\mathcal{A}_{ij} = (-1)^{i+j} \det(A_{ij})$ il complemento algebrico del coefficiente a_{ij} allora:

$$A^{-1} = \frac{1}{\det(A)} \cdot \mathcal{A}_{ij}^T \quad (6.11)$$

\mathcal{A}_{ij}^T è la matrice trasposta dei complementi algebrici

ESEMPIO

$$A = \begin{pmatrix} 3 & 2 & 4 \\ 2 & -1 & 2 \\ 1 & 4 & 3 \end{pmatrix}$$

Cominciamo assicurandoci che la matrice è invertibile calcolandone il determinante : $\det(A) = -5 \neq 0$, dunque la matrice è invertibile.

Calcoliamo i complementi algebrici:

$$C(a_{11}) = (-1)^{1+1} \cdot C_{11} = \det \begin{bmatrix} -1 & 2 \\ 4 & 3 \end{bmatrix} = (-3 - 8) = -11$$

$$C(a_{12}) = (-1)^{1+2} \cdot C_{12} = \det \begin{bmatrix} 2 & 2 \\ 1 & 3 \end{bmatrix} = -1 \cdot (6 - 2) = -4$$

$$C(a_{13}) = (-1)^{1+3} \cdot C_{13} = \det \begin{bmatrix} 2 & -1 \\ 1 & 4 \end{bmatrix} = (8 + 1) = 9$$

$$C(a_{21}) = (-1)^{2+1} \cdot C_{21} = \det \begin{bmatrix} 2 & 4 \\ 4 & 3 \end{bmatrix} = -1 \cdot (6 - 16) = 10$$

$$C(a_{22}) = (-1)^{2+2} \cdot C_{22} = \det \begin{bmatrix} 3 & 4 \\ 1 & 3 \end{bmatrix} = (9 - 4) = 5$$

$$C(a_{23}) = (-1)^{2+3} \cdot C_{23} = \det \begin{bmatrix} 3 & 2 \\ 1 & 4 \end{bmatrix} = -1 \cdot (12 - 2) = -10$$

$$C(a_{31}) = (-1)^{3+1} \cdot C_{31} = \det \begin{bmatrix} 2 & -1 \\ 4 & 2 \end{bmatrix} = 4 + 4 = 8$$

$$C(a_{32}) = (-1)^{3+2} \cdot C_{32} = \det \begin{bmatrix} 3 & 4 \\ 2 & 2 \end{bmatrix} = -1 \cdot (6 - 8) = 2$$

$$C(a_{33}) = (-1)^{3+3} \cdot C_{33} = \det \begin{bmatrix} 3 & 2 \\ 2 & -1 \end{bmatrix} = (-3 - 4) = -7$$

La matrice dei complementi algebrici è quindi:

$$\begin{bmatrix} -11 & -4 & 9 \\ 10 & 5 & -10 \\ 8 & 2 & -7 \end{bmatrix} \quad \text{per cui} \quad A^{-1} = \frac{1}{-5} \begin{bmatrix} -11 & 10 & 8 \\ -4 & 5 & 2 \\ 9 & -10 & -7 \end{bmatrix}$$

Corollario 6.12.1. L'inversa di una matrice $A \in \mathcal{M}_n(\mathbb{R})$ di ordine 2×2 si ottiene a partire da A scambiando gli elementi della diagonale principale, cambiando di segno quelli della diagonale secondaria e moltiplicando la matrice per l'inverso del determinante.

6.3.3 Regola di Cramer

Sia $A \cdot \mathbf{v} = \mathbf{b}$ un sistema lineare con $A \in \mathcal{M}_n(\mathbb{R})$, $\mathbf{b} \in \mathbb{R}^n$ e sia $\det(A) \neq 0$ cioè il sistema ha una sola soluzione. Sia B_j la matrice che si ottiene da A sostituendo la j -esima colonna con il termine noto b (per $1 \leq j \leq n$) e sia x_1, \dots, x_n la soluzione del sistema, allora

$$\forall 1 \leq j \leq n \quad x_j = \frac{\det(B_j)}{\det(A)}$$

ESEMPIO Vogliamo risolvere il sistema

$$\begin{cases} 8x + 10y = -46 \\ -9x + y = -34 \end{cases}$$

Consideriamo la matrice dei coefficienti del sistema lineare

$$A = \begin{pmatrix} 8 & 10 \\ -9 & 1 \end{pmatrix}$$

che ha determinante $\det(A) = (8)(1) - (10)(-9) = 98$ per calcolare la prima componente del vettore delle soluzioni, cioè x , sostituiamo il primo vettore colonna con il vettore delle soluzioni

$$A_1 = \begin{pmatrix} -46 & 10 \\ -34 & 1 \end{pmatrix}$$

tale matrice ha determinante $\det(A_1) = (-46)(1) - (10)(-34) = 294$ quindi:

$$x = \frac{\det(A_1)}{\det(A)} = \frac{294}{98} = 3$$

Per calcolare y facciamo le stesse operazioni:

$$A_2 = \begin{pmatrix} 4 & -46 \\ -9 & -34 \end{pmatrix}$$

la matrice ha determinante $\det(A_2) = (8)(-34) - (-46)(-9) = -686$

$$y = \frac{\det(A_2)}{\det(A)} = \frac{-686}{98} = -7$$

In conclusione abbiamo trovato le soluzioni del sistema: $x = 3, y = -7$

Autovalori, autovettori e autospazi

Sia $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ un endomorfismo e sia $\mathbf{v} \in \mathbb{R}^n$. Diciamo che \mathbf{v} è un **autovettore** di f se:

- (1) $\mathbf{v} \neq 0$
- (2) $\exists \lambda \in \mathbb{R}$ tale che $f(\mathbf{v}) = \lambda\mathbf{v}$

Più di preciso chiamiamo **autovettore** di f relativo all' **autovalore** λ .

Gli autovettori sono dunque tutti quei vettori la cui immagine risulta essere proporzionale al vettore stesso: cioè, se rappresentiamo i vettori in un piano, l'immagine deve essere dunque lungo la stessa retta su cui giace il vettore originale ossia deve conservare la direzione del vettore di partenza, gli autovalori non sono altro che i coefficienti di proporzionalità, ossia quanto l'immagine del vettore si dilata o si contrae rispetto a quella originale.

ESEMPIO

Sia $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ tale che $f(x, y) = (4x, x + 3y)$

Un autovettore di f è ad esempio il vettore $\mathbf{v} = (1, 1)$ relativo a $\lambda = 4$ poiché se applichiamo $f(\mathbf{v}) = (4, 4)$ che è un multiplo di quello originale: $f(\mathbf{v}) = 4\mathbf{v}$.

Teorema 7.1. Sia $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ un endomorfismo, \mathbf{v} autovettore di f relativo all'autovalore λ , sia $k \in \mathbb{R}$ con $k \neq 0$ allora $k \cdot \mathbf{v}$ è un autovettore di f relativo a λ .

ESEMPIO

Riprendiamo l'esempio di prima, un autovettore di f come abbiamo visto è $\mathbf{v} = (1, 1)$ relativo a $\lambda = 4$ ma anche $\mathbf{u} = (3, 3)$ è un autovettore relativo a $\lambda = 4$; quello che ci dice dunque il teorema è che ci sono infiniti autovettori relativi all'autovalore $\lambda = 4$ tutti con la forma $(k, k) \forall k \in \mathbb{R}$

Teorema 7.2. Sia $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ e λ un suo autovalore allora l'insieme:

$$V_\lambda = \{\mathbf{v} \in V \mid f(\mathbf{v}) = \lambda\mathbf{v}\}$$

si dice **autospazio** relativo a λ , V_λ è l'insieme di tutti gli autovettori di f relativi a λ più il vettore nullo $\underline{0}$, quindi se λ è un autovalore per f allora V_λ è un sottospazio di \mathbb{R}^n ossia:

$$V_\lambda \subset \mathbb{R}^n$$

Corollario 7.2.1. Un endomorfismo $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ ha $\lambda = 0$ come autovalore se e solo se f non è invertibile cioè solo se non è un isomorfismo. Nel caso $\lambda = 0$ sia un autovalore, l'autospazio è il nucleo di f ossia:

$$V_0 = \ker(f)$$

Osservazione 7.1. Si potrebbe pensare che tutte le applicazioni abbiano almeno un autovalore ma purtroppo esistono applicazioni che non ne hanno, un esempio sono le rotazioni poiché proprio per come sono formate non conservano mai le direzioni.

Dal teorema di rappresentazione sappiamo che ogni applicazione lineare può essere rappresentata da una matrice A che ha per colonne le immagini dei vettori della base scelta, dato che la matrice rappresenta proprio la funzione f possiamo definire anche autovalori e autovettori, diciamo dunque che $\lambda \in \mathbb{R}$ è un autovalore, e $X \in \mathbb{R}^{n,1}$ un autovettore quando

$$A \cdot X = \lambda X$$

Proprietà degli autovalori per matrici

1. $\lambda = 0$ è un autovalore se e solo se $\det(A) = 0$ o equivalentemente $\text{rk}(A) < n$ (quello che ci dice il corollario 7.2.1).
2. $\forall k \in \mathbb{N}, k \geq 2, \lambda^k$ è un autovalore per A^k
3. Se A è invertibile, λ^{-1} è un autovalore per A^{-1}
4. Se la matrice è triangolare (bassa, alta o diagonale) gli autovalori sono gli elementi della diagonale principale.

7.1 Calcolo degli autovalori

Teorema 7.3. $\lambda \in \mathbb{R}$ è un autovalore di $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ con $f(\mathbf{v}) = A \cdot \mathbf{v}$ se soltanto se il sistema $(A - \lambda I) \cdot \mathbf{v} = \underline{0}$ ha infinite soluzioni come sappiamo dalla teoria dei sistemi lineari, è un sistema lineare omogeneo che ammette sempre una soluzione ma a noi interessa che ne abbia infinite per cui λ è autovalore se e solo se $A - \lambda I$ non ha rango massimo quindi se soltanto se:

$$\det(A - \lambda I) = 0 \tag{7.1}$$

Cioè data una matrice $A \in \mathbb{R}^{n,n}$ i suoi autovalori sono le soluzioni della equazione $\det(A - \lambda I) = 0$ questa equazione viene chiamata **equazione caratteristica**. e $\det(A - \lambda I)$ viene chiamato **polinomio caratteristico** nell'incognita λ ; gli zeri del polinomio sono gli autovalori di f .

Possiamo verificare il fatto che una matrice ha come autovalore $\lambda = 0$ solo se non è invertibile (cfr. 7.2.1) poiché sostituendo λ nell'equazione caratteristica abbiamo $\det(A) = 0$.

ESEMPIO

Consideriamo l'applicazione $f(x, y) = (2x + 2y, 7x - 3y)$, la matrice rappresentativa di f nella base canonica \mathcal{E} è dunque :

$$A = \begin{pmatrix} 2 & 2 \\ 7 & -3 \end{pmatrix}$$

Per trovare gli autovalori dobbiamo risolvere l'equazione caratteristica $(A - \lambda I) \cdot \mathbf{v} = \underline{0}$. Cominciamo togliendo λ dagli elementi della diagonale principale ossia:

$$A - \lambda I = \begin{pmatrix} 2 - \lambda & 2 \\ 7 & -3 - \lambda \end{pmatrix} \text{ calcoliamo il ora il det } = (2 - \lambda)(-3 - \lambda) - 14$$

$$\det(A) = 0 \quad \text{per cui} \quad \lambda^2 + \lambda - 20 = 0 \quad \text{le cui soluzioni sono:} \quad \lambda = -5 \quad \lambda = 4$$

Abbiamo dunque che gli autovalori sono -5 e 4, calcoliamo ora i rispettivi autospazi, sostituiamo λ nella matrice A e risolviamo l'equazione $A \cdot v = \underline{0}$

$$V_{-5} = \begin{pmatrix} 7 & 2 \\ 7 & 2 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow \begin{cases} 7x + 2y = 0 \\ 7x + 2y = 0 \end{cases}$$

Da cui si ottiene che gli autovettori sono tutti nella forma $(2k, -7k)$ per cui $V_{-5} = \langle (2, -7) \rangle$
Risolviamo ora per $\lambda = 4$

$$V_4 = \begin{pmatrix} -2 & 2 \\ 7 & -7 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow \begin{cases} -2x + 2y = 0 \\ 7x - 7y = 0 \end{cases}$$

Da cui si ottiene che gli autovettori sono tutti nella forma (k, k) per cui $V_4 = \langle (1, 1) \rangle$

Corollario 7.3.1. Un endomorfismo $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ ha al più n autovalori distinti

Questo corollario deriva proprio dal teorema fondamentale dell'algebra; dal teorema segue che il polinomio caratteristico ammette precisamente n radici complesse (contate con la loro molteplicità) che non consideriamo poiché ci siamo messi in \mathbb{R} , mentre ammette al massimo n radici reali.

Corollario 7.3.2. Se $A \in \mathcal{M}_n(\mathbb{R})$ ha n dispari allora A possiede almeno 1 autovalore.

Dall'analisi matematica sappiamo che tutte le funzioni di grado dispari sono invertibili poiché sono biunivoche; ciò implica per il teorema degli zeri che passerà almeno una volta per l'asse x ovvero che ha almeno una radice reale. Questo si poteva vedere anche dal fatto che le soluzioni complesse compaiono sempre in coppie per cui ne esiste almeno una reale.

Definizione 7.1. Siano $\{\lambda_1, \dots, \lambda_s\}$ le radici reali con $s \leq n$ allora:

$$\det(A - \lambda I) = (k - \lambda_1)^{n_1}(k - \lambda_2)^{n_2} \cdots (k - \lambda_s)^{n_s} q(k) \quad (7.2)$$

Con k parametro e $q(k)$ un polinomio privo di radici reali. Chiamiamo il valore $n_i \in \mathbb{N}$ **molteplicità algebrica** dell'autovalore λ_i e lo indichiamo con $m_a(\lambda)$ in pratica è il numero che esprime quante volte l'autovalore annulla il polinomio caratteristico.

Definizione 7.2. Chiamiamo invece **molteplicità geometrica** di un autovalore λ_i il numero di vettori che formano la base dell'autospazio V_{λ_i} ossia il $\dim(V_{\lambda_i})$ e lo indichiamo con $m_g(\lambda)$. In termini pratici si usa la seguente formula:

$$m_g(\lambda) = n - \text{rk}(A - \lambda I) \quad (7.3)$$

Teorema 7.4. Se λ_i è un autovalore allora:

$$0 < m_g(\lambda) \leq m_a(\lambda) \quad (7.4)$$

Definizione 7.3. Sia $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ e A la matrice rappresentativa di f rispetto ad una base \mathcal{B} . La matrice A si dice **diagonalizzabile** se A è simile ad una matrice diagonale D ovvero se $\exists P \in \mathcal{M}_n(\mathbb{R})$ invertibile tale che:

$$A = P^{-1} \cdot D \cdot P$$

si dice anche che f è diagonalizzabile. Cioè vogliamo trovare una base per cui la matrice rappresentativa è una matrice diagonale.

Definizione 7.4. Sia $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ essa si dice **semplice** se $\exists \mathcal{B} \in \mathbb{R}^n$ composta da autovettori di f

Teorema 7.5. Un endomorfismo $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ è semplice se soltanto se è diagonalizzabile

OSSERVAZIONE. La matrice diagonale D non è altro che la matrice rappresentativa di f espressa nella base degli autovettori, inoltre, due matrici diagonalizzabili sono associate allo stesso endomorfismo (rispetto a basi differenti) se soltanto se sono simili alla stessa matrice diagonale (ovvero hanno gli stessi autovalori). Analogamente se solamente una delle due matrici è diagonalizzabile allora non sono associate allo stesso endomorfismo.

ESEMPIO

Consideriamo sempre lo stesso esempio di prima dove $f(x, y) = (2x + 2y, 7x - 3y)$ sappiamo che

$$A = M_{\mathcal{E}}(f) = \begin{pmatrix} 2 & 2 \\ 7 & -3 \end{pmatrix} \quad \text{inoltre} \quad M_{\mathcal{B}}(f) = M_{\mathcal{E}}^{\mathcal{B}} \cdot M_{\mathcal{E}}(f) \cdot M_{\mathcal{B}}^{\mathcal{E}}$$

Consideriamo dunque come base gli autovettori di f cioè $(1, 1)$ e $(2, -7)$

$$M_{\mathcal{B}}^{\mathcal{E}} = \begin{pmatrix} 1 & 2 \\ 1 & -7 \end{pmatrix} \quad \Rightarrow \quad \text{Inversa} \quad \Rightarrow \quad (M_{\mathcal{B}}^{\mathcal{E}})^{-1} = M_{\mathcal{E}}^{\mathcal{B}} = \frac{1}{9} \begin{pmatrix} 7 & 2 \\ 1 & -1 \end{pmatrix}$$

Abbiamo dunque

$$M_{\mathcal{B}}(f) = \frac{1}{9} \begin{pmatrix} 7 & 2 \\ 1 & -1 \end{pmatrix} \cdot \begin{pmatrix} 2 & 2 \\ 7 & -3 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 \\ 1 & -7 \end{pmatrix} = \begin{pmatrix} 4 & 0 \\ 0 & -5 \end{pmatrix}$$

Abbiamo dunque trovato la matrice diagonale i cui elementi sono proprio gli autovalori di f .

Teorema 7.6. Sia $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ siano $\lambda_1, \dots, \lambda_s$ gli autovalori e $V_{\lambda_1}, \dots, V_{\lambda_s}$ i corrispondenti autospazi di f allora:

$$\dim(V_1 + \dots + V_{\lambda_s}) = \dim V_{\lambda_1} \oplus \dots \oplus \dim V_{\lambda_s} \quad (7.5)$$

Il che vuol dire che possiamo fare la somma diretta fra gli autospazi e che tutte le basi degli autospazi sono fra loro linearmente indipendenti, cosa che non era scontata.

Corollario 7.6.1. Sia $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ e siano $\lambda_1, \dots, \lambda_s$ i corrispondenti autovalori distinti di f allora f è semplice e diagonalizzabile se soltanto se la somma delle molteplicità geometriche di $\lambda_1, \dots, \lambda_s$ è n , ovvero se la somma delle dimensioni dei suoi autospazi è n , ossia è necessario che abbia n autovettori linearmente indipendenti.

Corollario 7.6.2. Se $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ ha n autovalori reali distinti allora è diagonalizzabile

ATTENZIONE! L'enunciato precedente non è una condizione necessaria affinché f sia semplice, cioè se troviamo una funzione che non abbia n autovalori distinti non implica che l'applicazione non sia semplice, quello che bisogna considerare è la somma delle molteplicità geometriche!

Definizione 7.5. Sia λ autovalore di f si dice che λ è semplice se $m_g(\lambda) = 1$

Definizione 7.6. Si dice che λ è **regolare** se $m_a(\lambda) = m_g(\lambda)$

OSSERVAZIONE. Quindi possiamo dire che una condizione necessaria e sufficiente affinché una matrice sia diagonalizzabile è che la molteplicità algebrica e geometrica dei suoi autovalori coincidano. In particolare se f ha autovalori regolari e la somma delle m_a degli autovalori è la dim del dominio, allora f è diagonalizzabile.

ESEMPIO

Analizziamo la diagonalizzabilità della matrice:

$$A = \begin{pmatrix} 2-k & 2 & 5 \\ 0 & 3 & 9 \\ 0 & 0 & k+1 \end{pmatrix}$$

Sappiamo che in una matrice triangolare gli autovalori sono gli elementi della diagonale principale quindi $(2-k-\lambda)(3-\lambda)(k+1-\lambda)$

- I. Se $k \neq -1, 2$ la matrice A è diagonalizzabile poiché ha 3 autovalori distinti.
- II. Se $k = -1$ gli autovalori sarebbero 3,3,0, dobbiamo dunque vedere se $m_a(3) \stackrel{?}{=} m_g(3)$ ossia se è regolare, sappiamo che $m_a(3) = 2$ vediamo se $m_g(3) \stackrel{?}{=} 2$ per calcolare la molteplicità algebrica dobbiamo dunque risolvere:

$$A = \begin{pmatrix} 3 & 2 & 5 \\ 0 & 3 & 9 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{da Rouché-Capelli sappiamo che ha } \infty^1 \text{ soluzioni} \Rightarrow \dim V_3 = 1$$

Questo ci basta per dire che la matrice A non è diagonalizzabile per $k = -1$

- III. Se $k = 2$ gli autovalori sarebbero 0,3,3, dobbiamo dunque vedere se $m_a(3) \stackrel{?}{=} m_g(3)$, sappiamo che $m_a(3) = 2$ vediamo se $m_g(3) \stackrel{?}{=} 2$ per calcolare la molteplicità algebrica dobbiamo dunque risolvere:

$$A = \begin{pmatrix} 0 & 2 & 5 \\ 0 & 3 & 9 \\ 0 & 0 & 3 \end{pmatrix} \Rightarrow \begin{cases} 2y + 5z = 0 \\ 3y + 9z = 0 \\ 3z = 0 \end{cases} \Rightarrow \dim V_3 = 1$$

Quindi la matrice A non è diagonalizzabile per $k = 2$

7.2 Autovalori, determinante e traccia

Gli autovalori sono strettamente correlati con il determinante e la traccia infatti seguono i seguenti teoremi:

Teorema 7.7. Una matrice A di dimensione $n \times n$ a valori reali o complessi e autovalori $\lambda_1, \dots, \lambda_n$ ha come determinante il prodotto di tutti gli autovalori della matrice:

$$\det(A) = \lambda_1 \cdot \lambda_2 \cdots \lambda_n$$

Questo teorema viene dal fatto che una qualsiasi matrice A è simile a una matrice triangolare alta, poiché come abbiamo visto nei primi capitoli possiamo applicare il procedimento di Gauss-Jordan a una matrice senza cambiare l'essenza della matrice. Alla fine di questo procedimento otteniamo una matrice triangolare alta e sappiamo che in una matrice triangolare gli autovalori sono gli elementi della diagonale principale e inoltre nelle matrici triangolari il determinante è il prodotto degli elementi della diagonale principale e dunque anche degli autovalori.

Teorema 7.8. Una matrice A di dimensione $n \times n$ a valori reali o complessi e autovalori $\lambda_1, \dots, \lambda_n$ ha come traccia la somma degli autovalori:

$$\text{tr}(A) = \lambda_1 + \lambda_2 + \dots + \lambda_n$$

Questo teorema in realtà è dimostrabile facilmente per le matrici 2×2 , infatti consideriamo una generica matrice A fatta così:

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

Gli autovalori si ottengono con la matrice:

$$A - \lambda I = \begin{pmatrix} a - \lambda & b \\ c & d - \lambda \end{pmatrix} \quad \text{e} \quad \det(A - \lambda I) = (a - \lambda)(d - \lambda) - bc$$

dopo alcuni passaggi si ottiene:

$$\det(A - \lambda I) = \lambda^2 - (a + c)\lambda + ac - b^2$$

risolvendo l'equazione di secondo grado con la formula risolutiva si ha che

$$\lambda_{1/2} = \frac{a + c \pm \sqrt{\Delta}}{2}$$

In realtà non ci interessa sapere il Δ poiché se facciamo $\lambda_1 + \lambda_2$ si ottiene:

$$\frac{a + c - \sqrt{\Delta}}{2} + \frac{a + c + \sqrt{\Delta}}{2} = \frac{2a + 2c}{2} = \frac{2(a + c)}{2} = a + c$$

che è proprio la traccia della matrice.

7.3 Calcolo veloce degli autovalori

Calcolare gli autovalori di una matrice può risultare a volte un procedimento molto lungo ma per matrici relativamente piccole e facili possiamo calcolare gli autovalori nel giro di pochi secondi ricordando principalmente il fatto che:

- Il determinante è uguale al prodotto degli autovalori
- La traccia è uguale alla somma degli autovalori
- La matrice ammette zero come autovalore per ogni riga o colonna linearmente dipendente
- Se una colonna o riga ha coefficienti nulli tranne uno nella diagonale principale, quello non nullo è l'autovalore (questa proprietà deriva dallo sviluppo del determinante con il metodo di Laplace)

ESEMPIO

Consideriamo la matrice:

$$A = \begin{pmatrix} 5 & 7 & 2 \\ 0 & 2 & 3 \\ 0 & 4 & 6 \end{pmatrix}$$

Poiché la prima colonna ha come unico coefficiente non nullo il numero 5 possiamo affermare che è un autovalore. Ci accorgiamo che la seconda e la terza riga sono linearmente dipendenti quindi ammette 0 come autovalore. Ora ci basta calcolare la traccia che è 13 e sappiamo che è inoltre la somma degli autovalori, quindi ci basta fare $\lambda + 5 + 0 = 13$ e si conclude facilmente che l'autovalore che manca è il numero 8.

Quindi gli autovalori sono: 0, 5, 8

Consideriamo la matrice:

$$B = \begin{pmatrix} 2 & 1 & 3 & 1 \\ 0 & 7 & 0 & 0 \\ 4 & 2 & 6 & 2 \\ 4 & 9 & 6 & 2 \end{pmatrix}$$

In questa matrice si apprezza chiaramente che la quarta riga è la somma della seconda e della terza, inoltre la terza è il doppio della seconda. Quindi ammette l'autovalore 0 due volte. Inoltre il numero 7 è chiaramente un autovalore e utilizzando la traccia troviamo che l'autovalore mancante è 10.

Quindi gli autovalori sono: 0, 0, 7, 10

Uno potrebbe pensare che questo tipo di matrici così facili non si incontrano mai, ma in realtà si incontrano molto spesso, in particolare questo metodo ci aiuterà in futuro ad identificare facilmente quelle che chiamiamo le quadriche.

7.4 Complementi

Teorema 7.9 (TEOREMA DI CAYLEY-HAMILTON). Se f è un endomorfismo di uno spazio vettoriale V a dimensione finita $f : V \rightarrow V$ e $p(x)$ è il suo polinomio caratteristico, allora $p(f) = 0$. Analogamente, se A è una matrice quadrata e $p(x)$ il suo polinomio caratteristico, allora $p(A) = 0$ (ogni matrice quadrata annulla il suo polinomio caratteristico).

Anche se questo teorema ci dice ben poco, le sue implicazioni sono grandissime, difatti, questo teorema ci consente di calcolare l'inversa di una matrice con la sola conoscenza del polinomio caratteristico. Per fare ciò, quello che ci occorre fare, è sostituire nel polinomio caratteristico il generico autovalore λ con la matrice A dopodiché, moltiplichiamo per la matrice identità I tutti i termini noti restanti, ottenendo così un'equazione matriciale che se manipolata adeguatamente ci da l'espressione dell'inversa.

ESEMPIO. Consideriamo la matrice:

$$A = \begin{pmatrix} 3 & 2 \\ 1 & 4 \end{pmatrix} \quad A - \lambda I = \begin{pmatrix} 3 - \lambda & 2 \\ 1 & 4 - \lambda \end{pmatrix}$$

Calcoliamo il polinomio caratteristico, ossia $\det(A - \lambda I)$:

$$\det(A - \lambda I) = (3 - \lambda)(4 - \lambda) - 2 = \lambda^2 - 7\lambda + 10$$

Il teorema ci dice che:

$$A^2 - 7A + 10I = 0$$

La prima cosa che facciamo è isolare i termini con la matrice A in un solo lato:

$$A^2 - 7A = -10I$$

Raccogliamo A al lato sinistro:

$$A(A - 7I) = -10I$$

NOTA BENE Se dal raccoglimento resta un termine noto, dobbiamo dunque moltiplicarlo per la matrice identità I . Non metterlo sarebbe un errore grave poiché stiamo ragionando con matrici e non con numeri !

Per sbarazzarci della A raccolta a sinistra possiamo moltiplicare entrambi membri dell'equazione per A^{-1} , tenendo presente che $A \cdot A^{-1} = I$ e che I è l'elemento neutro rispetto al prodotto righe per colonne, ottenendo come risultato:

$$A \cdot A^{-1}(A - 7I) = -10IA^{-1} \quad \rightarrow \quad (A - 7I) = -10A^{-1}$$

Infine dividendo per -10 si ottiene:

$$A^{-1} = -\frac{1}{10}(A - 7I)$$

Svolgendo i calcoli risulta:

$$A^{-1} = -\frac{1}{10} \begin{pmatrix} -4 & 2 \\ 1 & -3 \end{pmatrix}$$

Questo teorema ci permette di calcolare potenze ad esponente intero e matrici inverse più semplicemente che con la moltiplicazione diretta o metodi come Gauss-Jordan.

Spazi Euclidei

Definizione 8.1. Uno spazio euclideo non è altro che uno spazio vettoriale reale su cui è definito un prodotto scalare.

È importante notare che parliamo di un prodotto scalare e non del prodotto per uno scalare poiché anche se i nomi sono simili, sono cose ben diverse.

Definizione 8.2. Sia $(V, +, \cdot)$ uno spazio vettoriale e sia $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{R}$ un'operazione binaria

- ① $\forall \mathbf{u}, \mathbf{v}, \mathbf{w} \in V : \langle \mathbf{u}, \mathbf{v} + \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{u}, \mathbf{w} \rangle$ (*è bidistributiva rispetto alla somma*)
 $\forall \mathbf{u}, \mathbf{v}, \mathbf{w} \in V : \langle \mathbf{u} + \mathbf{w}, \mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{w}, \mathbf{v} \rangle$
- ② $\forall \mathbf{v}, \mathbf{w} \in V \forall \lambda \in \mathbb{R} : \langle \mathbf{v}, \lambda \mathbf{w} \rangle = \langle \lambda \mathbf{v}, \mathbf{w} \rangle = \lambda \langle \mathbf{v}, \mathbf{w} \rangle$ (*è omogenea*)
- ③ $\forall \mathbf{v}, \mathbf{w} \in V : \langle \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{w}, \mathbf{v} \rangle$ (*è simmetrica*)
- ④ $\forall \mathbf{v} \in V : \langle \mathbf{v}, \mathbf{v} \rangle \geq 0$ e $\langle \mathbf{v}, \mathbf{v} \rangle = 0$ se soltanto se $\mathbf{v} = \underline{0}$ (*è definita positiva*)

Allora $\langle \cdot, \cdot \rangle$ si dice **prodotto scalare** su V

Definizione 8.3. Sia $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$, con $\mathbf{v} = (a_1, a_2, \dots, a_n)$, e $\mathbf{w} = (b_1, b_2, \dots, b_n)$ e sia l'operazione:

$$\langle \mathbf{u}, \mathbf{w} \rangle = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n \quad (8.1)$$

chiamiamo questa operazione **prodotto scalare standard** in \mathbb{R}^n

Definizione 8.4. in particolare, chiamiamo **spazio euclideo standard** $(\mathbb{R}^n, \langle \cdot, \cdot \rangle)$ lo spazio euclideo dotato di un prodotto scalare standard e lo indichiamo con \mathbb{E}^n dove n è la dimensione.

ESEMPIO

In \mathbb{R}^2 con $\mathbf{u} = (a, b)$ e $\mathbf{w} = (c, d)$ il prodotto scalare standard non è altro che

$$\langle \mathbf{u}, \mathbf{w} \rangle = ac + bd$$

È importante notare che ci sono anche i prodotti scalari non standard come ad esempio

$$\langle \mathbf{u}, \mathbf{w} \rangle = ac + 2bd$$

poiché soddisfa tutti i “requisiti” per essere un prodotto scalare.

Definizione 8.5. Chiamiamo **norma** una funzione che assegna ad ogni vettore di uno spazio vettoriale, tranne lo zero, un numero positivo che rappresenta la lunghezza del vettore.

Definizione 8.6. Sia $\mathbf{v} = (a_1, a_2 \dots, a_n)$, chiamiamo **norma euclidea** la norma:

$$\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle} = \sqrt{\sum_{i=1}^n (a_i)^2} \quad (8.2)$$

La norma gode le seguenti proprietà:

1. $\|\mathbf{v}\| \geq 0$ e in particolare $\|\mathbf{v}\| = 0$ se e solo se $\mathbf{v} = \underline{0}$ (*la norma è definita positiva*)
2. $\|\lambda\mathbf{v}\| = |\lambda|\|\mathbf{v}\|$ (*omogeneità*)
3. $\|\mathbf{v}\| - \|\mathbf{w}\| \leq \|\mathbf{v} + \mathbf{w}\| \leq \|\mathbf{v}\| + \|\mathbf{w}\|$ (*disuguaglianza triangolare*)
4. $\|\mathbf{v} + \mathbf{w}\|^2 = \|\mathbf{v}\|^2 + 2\langle \mathbf{v}, \mathbf{w} \rangle + \|\mathbf{w}\|^2$
5. $\langle \mathbf{v}, \mathbf{w} \rangle \leq \|\mathbf{v}\| \|\mathbf{w}\|$ (*disuguaglianza di Cauchy-Schwartz*)

Definizione 8.7. Sia $\mathbf{v} = (a_1, a_2 \dots, a_n)$ e $\mathbf{w} = (b_1, b_2 \dots, b_n)$ Chiamiamo **distanza euclidea** il numero reale $d(\mathbf{v}, \mathbf{w})$ non negativo:

$$d(\mathbf{v}, \mathbf{w}) = \|\mathbf{v} - \mathbf{w}\| = \sqrt{\sum_{i=1}^n (a_i - b_i)^2} \quad (8.3)$$

Possiamo definire l'angolo θ fra due vettori con la seguente formula:

$$\theta = \arccos \frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\|\mathbf{v}\| \|\mathbf{w}\|}$$

ESEMPIO

Vogliamo determinare la distanza così come l'angolo fra i vettori $\mathbf{v} = (6, 3, 2)$ e $\mathbf{w} = (2, 2, 1)$. Calcoliamo per prima l'angolo:

$$\begin{aligned} \langle \mathbf{v}, \mathbf{w} \rangle &= 6(2) + 3(2) + 2(1) = 20 \\ \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle} &= \sqrt{36 + 9 + 4} = 7 & \sqrt{\langle \mathbf{w}, \mathbf{w} \rangle} &= \sqrt{4 + 1 + 4} = 3 \\ \theta &= \arccos \left(\frac{20}{21} \right) \sim 17^\circ 45' \end{aligned}$$

Calcoliamo ora la distanza fra i due, $\|\mathbf{v} - \mathbf{w}\| = (6 - 2, 3 - 2, 2 - 1) = (4, 1, 1)$

$$\sqrt{\langle \mathbf{v} - \mathbf{w}, \mathbf{v} - \mathbf{w} \rangle} = \sqrt{16 + 1 + 1} = 3\sqrt{2}$$

Definizione 8.8. Diciamo che $\mathbf{v}, \mathbf{w} \in V$ sono **ortogonal**i o perpendicolari e lo scriviamo come $\mathbf{v} \perp \mathbf{w}$ se e solo se $\langle \mathbf{v}, \mathbf{w} \rangle = 0$

ESEMPIO

Consideriamo il prodotto scalare standard è i vettori $\mathbf{v} = (1, 6, k)$ e $\mathbf{w} = (4, k, k - 1)$ vogliamo determinare il parametro k affinché i due vettori siano perpendicolari. Calcoliamo il prodotto scalare standard:

$$\langle \mathbf{v}, \mathbf{w} \rangle = 4 + 6(k) + k(k - 1) = k^2 + 5k + 4 = (k + 4)(k + 1)$$

Per cui, affinché i vettori siano perpendicolari, $k \in \{-4, -1\}$

Teorema 8.1. Sia $(V, \langle \cdot, \cdot \rangle)$ uno spazio euclideo e siano $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ un insieme di vettori a due a due ortogonali, cioè il prodotto scalare fra i vettori è un delta di Kronecker $\forall (1 \leq i, j \leq n, i \neq j) \langle \mathbf{v}_i, \mathbf{v}_j \rangle = 0$, δ_{ij} , per costruzione i vettori $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ sono tra loro linearmente indipendenti (poiché ortogonali) e formano una **base ortogonale**.

Definizione 8.9. Sia ora V uno spazio euclideo, chiamiamo **ortogonale** di \mathbf{u} :

$$\mathbf{u}^\perp = \{\mathbf{v} \in V \mid \langle \mathbf{u}, \mathbf{v} \rangle = 0\}$$

ESEMPIO

Vogliamo calcolare tutti i vettori in \mathbb{E}^3 perpendicolari al vettore $\mathbf{u} = (3, 2, -1)$. Consideriamo un generico vettore $\in \mathbb{E}^3$ con componenti (x, y, z) , affinché i vettori siano perpendicolari il loro prodotto scalare deve essere nullo, ossia, deve essere soddisfatta l'equazione $3x + 2y - z = 0$, da questa si ricava:

$$\begin{cases} x = t \\ y = s \\ z = 3t + 2s \end{cases} \quad \text{tutti i vettori ortogonali appartengono al sottospazio } \mathbf{u}^\perp = \{(t, 0, 3t), (0, s, 2s)\}$$

Sia ora $U \subseteq V$ un sottospazio, chiamiamo **complemento ortogonale** di U in V e lo indichiamo con U^\perp il sottoinsieme di V così definito:

$$U^\perp = \{\mathbf{v} \in V \mid \forall \mathbf{u} \in U, \langle \mathbf{u}, \mathbf{v} \rangle = 0\}$$

ESEMPIO

Consideriamo ora un sottospazio vettoriale $U = \{(x, y, z) \in \mathbb{E}^3 \mid 2x + y - z = 0, 3y + z = 0\}$ di cui vogliamo calcolarne il complemento ortogonale. Per fare ciò ci occorre risolvere il sistema lineare omogeneo

$$\begin{cases} 2x + y - z = 0 \\ 3y + z = 0 \end{cases} \quad \text{sommando le equazioni si ottiene: } 2x + 4y = 0 \rightarrow x = -2y$$

Dalla seconda si ricava facilmente $z = -3y$

$$\text{quindi } U^\perp = \{\mathbf{v} \in \mathbb{E}^3 \mid \forall t \in \mathbb{R}, (-2t, t, -3t)\}$$

Vale la pena osservare, inoltre che se \mathcal{B} è una base del sottospazio U e \mathcal{B}' è una base di U^\perp allora $(\mathcal{B} \cup \mathcal{B}' = V)$ (l'unione delle basi è una base che genera tutto lo spazio V), questo è una osservazione più che ovvia poiché i vettori ortogonali alla base non potranno mai essere una loro combinazione lineare proprio per il fatto di essere perpendicolari! In effetti, nell'esempio precedente $(2, 1, -1), (0, 3, 1)$ e $(-2, 1, -3)$ formano una base di \mathbb{E}^3 .

Teorema 8.2. Sia $(V, \langle \cdot, \cdot \rangle)$ uno spazio euclideo di dimensione finita ($\dim(V) = n$) e sia U un suo sottospazio allora:

- $\dim(U^\perp) = \dim(V) - \dim(U)$
- $(U^\perp)^\perp = U$

Definizione 8.10. Sia $(V, \langle \cdot, \cdot \rangle)$ uno spazio euclideo e chiamiamo $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ un insieme di vettori ortonormali se $\forall (1 \leq i, j \leq n, i \neq j)$, $\langle \mathbf{v}_i, \mathbf{v}_j \rangle = 0$ e $\|\mathbf{v}_i\| = 1$, cioè se si tratta di un insieme di vettori ortogonali e di lunghezza unitaria.

Sia $(V, \langle \cdot, \cdot \rangle)$ uno spazio euclideo e \mathcal{B} una base di V si dice che \mathcal{B} è una **base ortonormale** di V se è un insieme di vettori ortonormali.

Osservazione 8.1. Tutte le basi canoniche \mathcal{E}_n sono basi ortonormali in \mathbb{R}^n .

Utilizzeremo spesso le basi ortonormali, dato che sono quelle più semplici da usare, infatti sia $(V, \langle \cdot, \cdot \rangle)$ uno spazio euclideo e $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ una sua base ortonormale e sia $\mathbf{v} = a_1\mathbf{v}_1 + \dots + a_n\mathbf{v}_n$ Allora $\forall i = 1 \dots n$ abbiamo che:

$$\langle \mathbf{v}, \mathbf{v}_i \rangle = a_1\langle \mathbf{v}_1, \mathbf{v}_i \rangle + \dots + a_i\langle \mathbf{v}_i, \mathbf{v}_i \rangle + \dots + a_n\langle \mathbf{v}_n, \mathbf{v}_i \rangle = a_i$$

Poiché i vettori sono ortogonali fra loro, per definizione di ortogonalità il prodotto scalare è nullo in tutti i casi, eccetto nel caso dove si fa il prodotto scalare con il vettore stesso.

Corollario 8.2.1. Sia $(V, \langle \cdot, \cdot \rangle)$ uno spazio euclideo e $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ una sua base ortonormale e siano $\mathbf{v} = \{a_1\mathbf{v}_1 + \dots + a_n\mathbf{v}_n\}$ e $\mathbf{w} = \{b_1\mathbf{v}_1 + \dots + b_n\mathbf{v}_n\}$, allora qualunque sia il prodotto scalare vale:

$$\langle \mathbf{v}, \mathbf{w} \rangle = a_1b_1 + a_2b_2 + \dots + a_nb_n \quad (8.4)$$

Questo corollario ci permette di lavorare come se fossimo in \mathbb{E}^n

Teorema 8.3 (ESISTENZA E COMPLETAMENTO DI BASI ORTONORMALI). Sia $(V, \langle \cdot, \cdot \rangle)$ uno spazio euclideo non nullo, allora:

1. Ogni insieme di vettori ortogonali in V si può completare a una base ortonormale.
Se si ha un insieme di vettori ortogonali, ma non ortonormale, si può ottenere una base ortonormale facendo una **normalizzazione**, che consiste in dividere ogni vettore per la sua norma euclidea.
2. $(V, \langle \cdot, \cdot \rangle)$ ha almeno una base ortonormale.

ESEMPIO

Consideriamo la base $\mathcal{B} = \{(3, 4), (4, -3)\}$ di \mathbb{E}^2 vogliamo calcolare una base ortonormale \mathcal{C} di \mathbb{E}^2 basandoci su i vettori della base \mathcal{B} .

Cominciamo col verificare che la base \mathcal{B} è una base ortogonale di \mathbb{E}^2

$$\langle \mathbf{b}_1, \mathbf{b}_2 \rangle = 12 - 12 = 0$$

poiché il prodotto scalare standard è nullo possiamo dire che in effetti è una base ortogonale, per renderla una base ortonormale ci basta dividere le componenti di ogni vettore per la sua norma, la norma di $\mathbf{b}_1 = \sqrt{3^2 + 4^2} = 5$ e $\mathbf{b}_2 = \sqrt{4^2 + (-3)^2} = 5$, infine si ottengono i vettori.

$$\mathbf{c}_1 = \left(\frac{3}{5}, \frac{4}{5} \right) \quad \mathbf{c}_2 = \left(\frac{4}{5}, \frac{-3}{5} \right)$$

Definizione 8.11. Dati due vettori $\mathbf{v}, \mathbf{w} \in V$ si chiama proiezione ortogonale di \mathbf{w} su \mathbf{v} il vettore:

$$\text{pr}_{\mathbf{v}}(\mathbf{w}) = \frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\|\mathbf{v}\|^2} \cdot \mathbf{v} = \frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle} \cdot \mathbf{v}$$

È importante sottolineare:

- $\text{pr}_{\mathbf{v}}$ è un vettore parallelo a \mathbf{v} ,
- $\mathbf{v} - \text{pr}_{\mathbf{v}}$ è un vettore ortogonale a \mathbf{v} ,
- $\mathbf{v} = (\mathbf{v} - \text{pr}_{\mathbf{v}}) + \text{pr}_{\mathbf{v}}$, ovvero ogni vettore \mathbf{v} può sempre essere scritto come somma di un vettore ortogonale e uno parallelo ad un altro vettore \mathbf{w} .

ESEMPIO

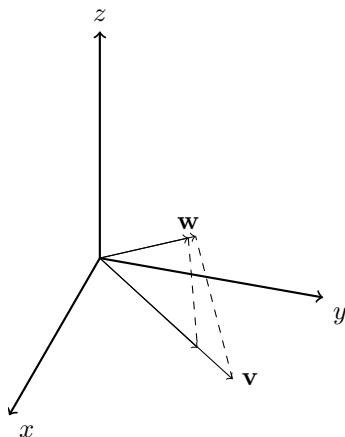
Consideriamo i vettori $\mathbf{v} = (3, 3, 1)$ e $\mathbf{w} = (2, 2, 2)$, vogliamo calcolare le loro proiezioni reciproche. Cominciamo calcolando la proiezione di \mathbf{v} su \mathbf{w} ossia $\text{pr}_{\mathbf{v}}(\mathbf{w})$:

$$\text{pr}_{\mathbf{v}}(\mathbf{w}) = \frac{14}{19}(3, 3, 1) = \left(\frac{42}{19}, \frac{42}{19}, \frac{14}{19} \right)$$

la proiezione di \mathbf{w} su \mathbf{v} :

$$\text{pr}_{\mathbf{w}}(\mathbf{v}) = \frac{14}{12}(2, 2, 2) = \left(\frac{7}{3}, \frac{7}{3}, \frac{7}{3} \right)$$

Rappresentato in un piano cartesiano si ha:



Teorema 8.4 (ORTOGONALIZZAZIONE/ORTONORMALIZZAZIONE DI GRAM-SCHMIDT). Sia $(V, \langle \cdot, \cdot \rangle)$ uno spazio euclideo e sia $\mathcal{B}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\} \in V$ dei vettori linearmente indipendenti allora esistono $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_n \in V$ tali che:

1. $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_n \in V$ sono ortonormali.
2. $\langle \mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n \rangle = \langle \mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_n \rangle$ (generano lo stesso spazio)

Questo teorema però non si limita ad affermare che esistono dei vettori ortonormali che generano tutto lo spazio euclideo, ma ci propone un algoritmo su come trovarli. Supponiamo di avere una base di fatta dai vettori $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$, per trovare una base ortonormale, ci teniamo come “buono” il primo vettore, è costruiamo i successivi sottraendo ai vettori \mathcal{B} la loro proiezione ortogonale rispetto ai nuovi vettori ortogonali:

$$\begin{aligned}\mathbf{u}_1 &= \mathbf{v}_1 \\ \mathbf{u}_2 &= \mathbf{v}_2 - \text{pr}_{\mathbf{u}_1}(\mathbf{v}_2) \quad \text{ossia} \quad = \mathbf{v}_2 - \frac{\langle \mathbf{v}_2, \mathbf{u}_1 \rangle}{\langle \mathbf{u}_1, \mathbf{u}_1 \rangle} \cdot \mathbf{u}_1 \\ \mathbf{u}_3 &= \mathbf{v}_3 - \text{pr}_{\mathbf{u}_1}(\mathbf{v}_3) - \text{pr}_{\mathbf{u}_2}(\mathbf{v}_3) \quad \text{ossia} \quad = \mathbf{v}_3 - \frac{\langle \mathbf{v}_3, \mathbf{u}_1 \rangle}{\langle \mathbf{u}_1, \mathbf{u}_1 \rangle} \cdot \mathbf{u}_1 - \frac{\langle \mathbf{v}_3, \mathbf{u}_2 \rangle}{\langle \mathbf{u}_2, \mathbf{u}_2 \rangle} \cdot \mathbf{u}_2\end{aligned}$$

In generale l' n -esimo vettore può essere calcolato come :

$$\mathbf{u}_n = \mathbf{v}_n - \sum_{i=1}^{n-1} \text{pr}_{\mathbf{u}_i}(\mathbf{v}_n) = \mathbf{v}_n - \sum_{i=1}^{n-1} \frac{\langle \mathbf{v}_n, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \cdot \mathbf{u}_i \quad (8.5)$$

Questo algoritmo funziona poiché se sceglioamo un vettore di riferimento, e “togliamo”, al vettore successivo, la parte parallela al vettore di riferimento, quello che ci resta è la sola parte ortogonale al vettore, così facendo possiamo togliere all' n -esimo vettore le parti parallele ai vettori ortogonali ricavati precedentemente, inoltre siamo sicuri che troveremo sempre una base ortogonale, poiché, essendo i vettori linearmente indipendenti (per il fatto di costituire una base), il vettore successivo avrà sempre una parte che non è parallela a nessun altro vettore, altrimenti sarebbe un vettore linearmente dipendente. È importante sottolineare che di conseguenza, la procedura deve essere applicata in forma strettamente sequenziale, giacché per conoscere \mathbf{u}_n dovremo aver già calcolato gli $n-1$ precedenti.

A questo punto abbiamo già trovato una base ortonormale con n vettori a due a due perpendicolari, ora per renderli ortonormali, ci basta normalizzare ciascun vettore:

$$\mathbf{w}_i = \frac{\mathbf{u}_i}{\|\mathbf{u}_i\|}$$

ESEMPIO

Consideriamo la base $\mathcal{B} = \{(-1, 0, 1), (2, 0, 0), (0, 2, 0)\}$, vogliamo calcolare una base ortonormale \mathcal{C} a partire da \mathcal{B} con il procedimento di Gram-Schmidt.

$$\mathbf{u}_1 = (-1, 0, 1) \quad \text{non modifichiamo il primo vettore}$$

$$\mathbf{u}_2 = \mathbf{b}_2 - \text{pr}_{\mathbf{u}_1}(\mathbf{b}_2) = (2, 0, 0) - \frac{-2}{2}(-1, 0, 1) = (1, 0, 1)$$

$$\mathbf{u}_3 = \mathbf{b}_3 - \text{pr}_{\mathbf{u}_1}(\mathbf{b}_3) - \text{pr}_{\mathbf{u}_2}(\mathbf{b}_3) = (0, 2, 0) - \frac{0}{2}(-1, 0, 1) - \frac{0}{2}(1, 0, 1) = (0, 2, 0)$$

Ora che abbiamo ottenuto vettori dividiamo ciascuno per la sua norma:

$$\begin{aligned}\mathbf{c}_1 &= \frac{\mathbf{u}_1}{\|\mathbf{u}_1\|} = \left(\frac{-1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}} \right) \\ \mathbf{c}_2 &= \frac{\mathbf{u}_2}{\|\mathbf{u}_2\|} = \left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}} \right) \\ \mathbf{c}_3 &= \frac{\mathbf{u}_3}{\|\mathbf{u}_3\|} = (0, 1, 0)\end{aligned}$$

8.1 Isometrie

Definizione 8.12. Siano V, W spazi euclidei e sia $f : V \rightarrow W$ un'applicazione su essi definita, diciamo che f è una **isometria** se conserva il prodotto scalare, ovvero:

$$\forall (\mathbf{v}, \mathbf{w} \in V) \text{ si ha } \langle \mathbf{v}, \mathbf{w} \rangle = \langle f(\mathbf{v}), f(\mathbf{w}) \rangle$$

Questa caratteristica delle isometrie ci permette di concludere che la norma dell'immagine è uguale alla norma del vettore di partenza:

$$\|f(\mathbf{v})\| = \sqrt{\langle f(\mathbf{v}), f(\mathbf{v}) \rangle} = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle} = \|\mathbf{v}\|$$

Inoltre le isometrie conservano gli angoli fra i vettori poiché:

$$\theta = \arccos \frac{\langle f(\mathbf{v}), f(\mathbf{w}) \rangle}{\|f(\mathbf{v})\| \|f(\mathbf{w})\|} = \arccos \frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\|\mathbf{v}\| \|\mathbf{w}\|}$$

OSSERVAZIONE. Un'isometria è sempre iniettiva (cioè il suo $\ker(f) = \underline{0}$) infatti se esistesse un vettore $\mathbf{v} \in \ker(f)$ diverso dal vettore nullo tale che $f(\mathbf{v}) = \underline{0}$ ciò implicherebbe che:

$$\langle \mathbf{v}, \mathbf{v} \rangle = \langle f(\mathbf{v}), f(\mathbf{v}) \rangle = \langle \underline{0}, \underline{0} \rangle = 0$$

ossia avrei trovato un vettore diverso dal vettore nullo il cui prodotto scalare è zero (che è contro la proprietà del prodotto scalare di essere definito positivo).

Teorema 8.5. Sia V uno spazio euclideo e sia $f : V \rightarrow V$ un endomorfismo; f è una isometria solo se manda basi ortonormali in basi ortonormali, cioè se $\mathcal{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ è una base di V , allora anche $\mathcal{B}' = \{f(\mathbf{v}_1), \dots, f(\mathbf{v}_n)\}$ è una base di V .

Definizione 8.13. Sia $Q \in \mathcal{M}_n(\mathbb{R})$ si dice che Q è una **matrice ortogonale** se:

$${}^t Q \cdot Q = I \quad \text{ossia se} \quad {}^t Q = Q^{-1}$$

Il determinante di una matrice ortogonale Q può essere $\{+1, -1\}$ poiché sapendo che $\det(I) = 1$ e applicando il teorema di Binet si ottiene:

$$\det({}^t Q \cdot Q) = \det(I) \quad \rightarrow \quad \det({}^t Q) \cdot \det(Q) = 1 \quad \rightarrow \quad [\det(Q)]^2 = 1$$

le cui uniche soluzioni possibili sono $\{+1, -1\}$. In particolare, diciamo che una matrice ortogonale Q è una matrice **ortogonale speciale** se $\det(Q) = 1$.

OSSERVAZIONE. Se Q è una matrice ortogonale le sue colonne formano una base ortonormale di \mathbb{R}^n .

Teorema 8.6. Sia V uno spazio euclideo e sia $f : V \rightarrow V$ un endomorfismo, la condizione necessaria e sufficiente affinché f sia una isometria è che scelta una qualsiasi base \mathcal{B} ortonormale la matrice rappresentativa dell'endomorfismo rispetto a quella base sia una matrice ortogonale, ossia:

$${}^t [M_{\mathcal{B}, \mathcal{B}}(f)] \cdot M_{\mathcal{B}, \mathcal{B}}(f) = I$$

Consideriamo gli automorfismi nello spazio euclideo $f : \mathbb{E}^2 \rightarrow \mathbb{E}^2$, per quanto visto dal teorema precedente, le isometrie in esso devono avere come matrice rappresentativa, in una certa base, una matrice ortogonale Q . Per determinare le isometrie dobbiamo dunque risolvere ${}^t Q \cdot Q = I$:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \begin{pmatrix} a & c \\ b & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \rightarrow \quad \begin{pmatrix} a^2 + c^2 & ab + cd \\ ab + cd & b^2 + d^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Le possibili soluzioni a questo sistema sono:

$$Q_1 = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \quad \text{e} \quad Q_2 = \begin{pmatrix} \cos \theta & \sin \theta \\ \sin \theta & -\cos \theta \end{pmatrix} \quad \text{per } \theta \in [0, 2\pi)$$

La matrice Q_1 prende il nome di **matrice di rotazione** è ciò che rappresenta nel piano è una rotazione antioraria di un angolo θ con centro di rotazione nell'origine. La matrice di rotazione, come ci si poteva aspettare (poiché è un'isometria), conserva tutte le lunghezze e tutti gli angoli fra i vettori, inoltre il suo determinante è:

$$\det(Q_1) = \cos^2 \theta + \sin^2 \theta = 1$$

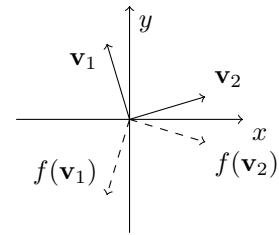
questo ci dice che la matrice non cambia l'orientamento dei vettori poiché se ricordiamo quanto detto nel capitolo del determinante, questo rappresenta il fattore con cui vengono modificati i volumi degli oggetti (in questo caso essendo 1 non vengono modificati) e il segno rappresenta l'orientamento dei vettori (non essendo negativo non c'è alcun cambio nel loro orientamento).

La matrice Q_2 è una matrice che rappresenta la composizione tra una simmetria ortogonale e una rotazione antioraria, cioè, la **matrice di simmetria** fa la riflessione rispetto alla retta passante per l'origine e formante un angolo θ con l'asse x .

In particolare, se $\theta = 0$:

Graficamente rappresenta la simmetria rispetto all'asse delle ascisse:

$$Q_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$



Di conseguenza

$$\begin{pmatrix} \cos \theta & \sin \theta \\ \sin \theta & -\cos \theta \end{pmatrix} = \underbrace{\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}}_{\text{matrice di rotazione}} \cdot \underbrace{\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}}_{\text{matrice di simmetria}}$$

poiché il suo determinante è:

$$\det(Q_1) = -\cos^2 \theta - \sin^2 \theta = -(\cos^2 \theta + \sin^2 \theta) = -1$$

Per cui questa matrice cambia l'orientamento dei vettori.

Teorema Spettrale

In questo capitolo parleremo soltanto del teorema spettrale, che è uno dei teoremi più importanti dell'algebra lineare, ma per capire la sua formulazione, ci servono dei concetti e teoremi preliminari che enunceremo nelle prossime pagine.

Teorema 9.1. Una matrice $A \in \mathcal{M}_n(\mathbb{R})$ si dice **ortogonalmente diagonalizzabile** se esistono una matrice diagonale D e una matrice ortogonale Q in $\mathcal{M}_n(\mathbb{R})$ tali che $A = Q \cdot D \cdot {}^t Q$.

Quello che afferma questo teorema è che una matrice ortogonalmente diagonalizzabile ha come matrici di passaggio una isometria.

Facciamo una piccola pausa per introdurre un nuovo operatore che ci aiuterà a dimostrare il teorema spettrale, questo operatore, detto **prodotto Hermitiano**, è analogo al prodotto scalare ma invece di agire sugli spazi vettoriali reali agisce sugli spazi vettoriali complessi ($\in \mathbb{C}$).

Sia \mathbb{C}^n lo spazio vettoriale complesso di dimensione n , il prodotto hermitiano è un'operazione binaria $\langle \cdot, \cdot \rangle : \mathbb{C}^n \times \mathbb{C}^n \rightarrow \mathbb{C}$ definita in modo standard come:

$$\langle \mathbf{v}, \mathbf{w} \rangle = \mathbf{v}_1 \overline{\mathbf{w}_1} + \mathbf{v}_2 \overline{\mathbf{w}_2} + \dots + \mathbf{v}_n \overline{\mathbf{w}_n} \quad (9.1)$$

dove $\mathbf{v}_n \overline{\mathbf{w}_n}$ Inoltre gode le seguenti proprietà:

$$① \quad \forall \mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{C}^n : \langle \mathbf{u} + \mathbf{w}, \mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{w}, \mathbf{v} \rangle$$

$$\forall \mathbf{v}, \mathbf{w} \in \mathbb{C}^n \quad \forall \lambda \in \mathbb{R} : \langle \lambda \mathbf{v}, \mathbf{w} \rangle = \lambda \langle \mathbf{v}, \mathbf{w} \rangle$$

(lineare rispetto alla prima variabile)

$$② \quad \forall \mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{C}^n : \langle \mathbf{u}, \mathbf{v} + \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{u}, \mathbf{w} \rangle$$

$$\forall \mathbf{v}, \mathbf{w} \in \mathbb{C}^n \quad \forall \lambda \in \mathbb{R} : \langle \mathbf{v}, \lambda \mathbf{w} \rangle = \bar{\lambda} \langle \mathbf{v}, \mathbf{w} \rangle \quad (\bar{\lambda} = \text{coniugato di } \lambda)$$

(sequilineare rispetto alla seconda variabile)

$$③ \quad \forall \mathbf{v}, \mathbf{w} \in \mathbb{C}^n \quad \langle \mathbf{v}, \mathbf{w} \rangle = \overline{\langle \mathbf{w}, \mathbf{v} \rangle}$$

(è sequisimmetrica)

Come conseguenza immediata si ha che $\langle \mathbf{v}, \mathbf{v} \rangle = \overline{\langle \mathbf{v}, \mathbf{v} \rangle}$ quest'ultima relazione è vera se soltanto se $\text{Im } \mathbf{v} = 0$ per cui il prodotto hermitiano di un qualunque vettore per se stesso è un numero reale.

④ $\forall \mathbf{v} \in \mathbb{C}^n \quad \langle \mathbf{v}, \mathbf{v} \rangle \geq 0$ e $\langle \mathbf{v}, \mathbf{v} \rangle = 0$ se soltanto se $\mathbf{v} = \underline{0}$

(è definita positiva)

Inoltre gode anche del teorema di Carnot, norma e ortogonalità:

⑤ $\forall \mathbf{v} \in \mathbb{C}^n$ norma $\rightarrow \|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}$

⑥ $\forall \mathbf{v}, \mathbf{w} \in \mathbb{C}^n \quad \mathbf{v} \perp \mathbf{w} \quad \text{se} \quad \langle \mathbf{v}, \mathbf{w} \rangle = 0$

⑦ $\|\mathbf{v} + \mathbf{w}\|^2 = \|\mathbf{v}\|^2 + \|\mathbf{w}\|^2 + 2\operatorname{Re}(\langle \mathbf{v}, \mathbf{w} \rangle)$

Osservazione 9.1.1. Consideriamo una matrice $A \in \mathcal{M}_n(\mathbb{R})$ e $\mathbf{v}, \mathbf{w} \in \mathbb{R}^{n,1}$ (colonne) allora:

$$\langle A\mathbf{v}, \mathbf{w} \rangle = {}^t(A\mathbf{v}) \cdot \mathbf{w} = {}^t\mathbf{v} \cdot {}^tA \cdot \mathbf{w} = \langle \mathbf{v}, {}^tA\mathbf{w} \rangle$$

consideriamo ora una matrice $A \in \mathcal{M}_n(\mathbb{C})$ e $\mathbf{v}, \mathbf{w} \in \mathbb{C}^{n,1}$ (colonne) allora seguendo la stessa logica:

$$\langle A\mathbf{v}, \mathbf{w} \rangle = {}^t(A\mathbf{v}) \cdot \overline{\mathbf{w}} = {}^t\mathbf{v} \cdot {}^tA \cdot \overline{\mathbf{w}} = {}^t\mathbf{v} \cdot \overline{{}^tA \cdot \mathbf{w}} = \langle \mathbf{v}, A^H\mathbf{w} \rangle$$

dove A^H è il coniugato della trasposta $(\overline{{}^tA})$

Grazie a questa osservazione possiamo dimostrare il seguente lemma che ci aiuterà a dimostrare il teorema spettrale.

Lemma 9.1. Gli autovalori di una matrice reale simmetrica sono tutti reali.

$$A \in \mathcal{M}_n(\mathbb{R}) \quad \lambda \in A \quad \therefore \quad \lambda \in \mathbb{R} \tag{9.2}$$

Ricordiamo che le matrici simmetriche sono tutte le matrici tali che $A = {}^tA$, per il teorema fondamentale dell'algebra sappiamo che la matrice A ha n autovalori complessi contati con la loro molteplicità. Consideriamo \mathbf{v} un autovettore relativo a $\lambda \in \mathbb{C}$ ossia tale che $(A\mathbf{v} = \lambda\mathbf{v})$.

Consideriamo i seguenti prodotti scalari:

$$\langle A\mathbf{v}, \mathbf{v} \rangle = \langle \lambda\mathbf{v}, \mathbf{v} \rangle = \lambda \langle \mathbf{v}, \mathbf{v} \rangle$$

$$\langle \mathbf{v}, A\mathbf{v} \rangle = \langle \mathbf{v}, \lambda\mathbf{v} \rangle = \bar{\lambda} \langle \mathbf{v}, \mathbf{v} \rangle$$

Ma per quanto osservato prima $\langle A\mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{v}, A^H\mathbf{w} \rangle$, sappiamo che $A^H = (\overline{{}^tA})$ e per definizione di matrice simmetrica abbiamo $A^H = (\overline{A})$, per ipotesi A è reale, quindi $\overline{A} = A$ per cui $\langle \mathbf{v}, A\mathbf{v} \rangle = \langle A\mathbf{v}, \mathbf{v} \rangle$ ciò implica $\lambda \langle \mathbf{v}, \mathbf{v} \rangle = \bar{\lambda} \langle \mathbf{v}, \mathbf{v} \rangle$ e ciò è solo possibile se $\lambda = \bar{\lambda}$, la uguaglianza risulta verificata soltanto se $\operatorname{Im}(\lambda) = 0$ per cui λ è sempre un numero reale!

Lemma 9.2. Sia $A \in \mathcal{M}_n(\mathbb{R})$ simmetrica ($A = {}^tA$) e siano λ e μ due autovalori distinti della matrice A , siano \mathbf{v} e \mathbf{w} gli autovettori relativi a λ e μ rispettivamente. Allora $\mathbf{v} \perp \mathbf{w}$ cioè $\langle \mathbf{v}, \mathbf{w} \rangle = 0$

Infatti seguendo la stessa logica di prima abbiamo che $\langle A\mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{v}, A\mathbf{w} \rangle$ ma:

$$\langle A\mathbf{v}, \mathbf{w} \rangle = \langle \lambda\mathbf{v}, \mathbf{w} \rangle = \lambda\langle \mathbf{v}, \mathbf{w} \rangle$$

$$\langle \mathbf{v}, A\mathbf{w} \rangle = \langle \mathbf{v}, \mu\mathbf{w} \rangle = \mu\langle \mathbf{v}, \mathbf{w} \rangle$$

Poiché i due autovalori sono diversi almeno uno non è zero per cui sia $\mu \neq 0$:

$$\lambda\langle \mathbf{v}, \mathbf{w} \rangle = \mu\langle \mathbf{v}, \mathbf{w} \rangle \quad \rightarrow \quad \frac{\lambda}{\mu}\langle \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{v}, \mathbf{w} \rangle \quad \text{che ha come soluzioni} \quad \begin{cases} \frac{\lambda}{\mu} = 1 & \text{assurdo! } \lambda \neq \mu \\ \langle \mathbf{v}, \mathbf{w} \rangle = 0 & \end{cases}$$

Corollario 9.2.1. Sia $A \in \mathcal{M}_n(\mathbb{R})$ simmetrica ($A = {}^t A$), se A ha n autovalori distinti cioè se $m_a(\lambda_i) = 1$ allora A è ortogonalmente diagonalizzabile.

Teorema 9.2 (TEOREMA SPETTRALE). Una matrice reale simmetrica è ortogonalmente diagonalizzabile.

Essendo $A \in \mathcal{M}_n(\mathbb{R})$ simmetrica, la domanda che ci poniamo ora è: come facciamo a trovare la matrice Q ortogonale tale che $A = QD^tQ$? Quello che ci occorre fare è:

1. Trovare gli autovalori $\lambda_1, \dots, \lambda_s$ con $s \leq n$ (tutti $\in \mathbb{R}$ per il lemma 9.1).
2. Trovare i corrispondenti autospazi $V_{\lambda_1}, V_{\lambda_2}, \dots, V_{\lambda_s}$ (che sono tutti \perp tra loro per il lemma 9.2)
3. Trovare una base per ogni autospazio $\mathcal{B}_{V_{\lambda_1}}, \mathcal{B}_{V_{\lambda_2}}, \dots, \mathcal{B}_{V_{\lambda_s}}$
4. Trasformare ogni base trovata in basi ortonormali con l'algoritmo di Gram-Schmidt

$$\mathcal{B}_{V_{\lambda_1}}, \mathcal{B}_{V_{\lambda_2}}, \dots, \mathcal{B}_{V_{\lambda_s}} \longrightarrow \mathcal{B}'_{V_{\lambda_1}}, \mathcal{B}'_{V_{\lambda_2}}, \dots, \mathcal{B}'_{V_{\lambda_s}}$$

5. La matrice Q che cerchiamo avrà per colonne i vettori che compongono le basi normalizzate

$$\mathcal{B}'_{V_{\lambda_1}}, \mathcal{B}'_{V_{\lambda_2}}, \dots, \mathcal{B}'_{V_{\lambda_s}}$$

Corollario 9.2.2. Sia $A \in \mathcal{M}_n(\mathbb{R})$ simmetrica e $\lambda_1, \dots, \lambda_s$ autovalori distinti e $V_{\lambda_1}, V_{\lambda_2}, \dots, V_{\lambda_s}$ i corrispondenti autospazi e $\mathcal{B}'_{V_{\lambda_1}}, \mathcal{B}'_{V_{\lambda_2}}, \dots, \mathcal{B}'_{V_{\lambda_s}}$ le loro rispettive basi ortonormali. Allora:

- (i) L'unione delle basi ortonormali è una base di \mathbb{R}^n

$$\mathcal{B}'_{V_{\lambda_1}} \cup \mathcal{B}'_{V_{\lambda_2}} \cup \dots \cup \mathcal{B}'_{V_{\lambda_s}} = <\mathbb{R}^n>$$

- (ii) La somma degli autospazi è diretta ed è tutto \mathbb{R}^n

$$V_{\lambda_1} \oplus V_{\lambda_2} \oplus \dots \oplus V_{\lambda_s} = \mathbb{R}^n$$

Inoltre, $\forall \mathbf{v} \in \mathbb{R}^n$, \mathbf{v} si scrive in modo unico come combinazione lineare di autovettori:

$$\mathbf{v} = \mathbf{v}_{\lambda_1} + \dots + \mathbf{v}_{\lambda_s} \text{ con } \mathbf{v}_{\lambda_i} \in V_{\lambda_i}$$

ove \mathbf{v}_{λ_i} è la proiezione ortogonale di \mathbf{v} su V_{λ_i}

Definizione 9.1. Sia V_{λ_i} l'autospazio relativo a λ_i e sia $\mathcal{B}'_{\lambda_i} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ una sua base ortonormale. Consideriamo una matrice B che ha per colonne i vettori della base ortonormale $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$. Allora chiamiamo **matrice di proiezione ortogonale** su V_{λ_i} la matrice ottenuta così:

$$P_i = B \cdot {}^t B$$

Un endomorfismo è una proiezione se vale $f \circ f = f$.

Corollario 9.2.3. Sia $A \in \mathcal{M}_n(\mathbb{R})$ simmetrica, $\lambda_1, \dots, \lambda_s$ i suoi autovalori, $V_{\lambda_1}, V_{\lambda_2}, \dots, V_{\lambda_s}$ i corrispondenti autospazi, inoltre, siano P_1, P_2, \dots, P_s matrici di proiezione ortogonale su V_{λ_i} . Si hanno le seguenti proprietà:

$$\textcircled{i} \quad A = \lambda_1 P_1, \lambda_2 P_2 \dots \lambda_s P_s$$

DECOMPOSIZIONE SPETTRALE

$\lambda_1, \dots, \lambda_s$ è lo **spettro** di A

$$\textcircled{ii} \quad P_1 + P_2 + \dots + P_s = I$$

$$\textcircled{iii} \quad \forall (1 \leq i \leq s) \quad P_i = {}^t P_i \quad \text{e} \quad P_i^n = P_i \quad \text{È SIMMETRICA E IDEMPOTENTE}$$

$$\textcircled{iv} \quad \forall (1 \leq i, j \leq s) \quad \text{se } i \neq j \text{ si ha } P_i \cdot P_j = 0 \quad \text{È UN DELTA DI KRONECKER}$$

\textcircled{v} Gli autovalori possibili di P_i sono +1 e 0.

Le proprietà appena viste ci permettono di calcolare A^n , infatti, sia $A = \lambda_1 P_1, \lambda_2 P_2 \dots \lambda_s P_s$ la decomposizione spettrale di A , allora:

$$\begin{aligned} A^n &= (\lambda_1 P_1, \lambda_2 P_2 \dots \lambda_s P_s)^n = (\lambda_1 P_1)^n, (\lambda_2 P_2)^n \dots (\lambda_s P_s)^n = \lambda_1^n P_1^n, \lambda_2^n P_2^n \dots \lambda_s^n P_s^n \\ &= \lambda_1^n P_1, \lambda_2^n P_2 \dots \lambda_s^n P_s \end{aligned} \quad (9.3)$$

Questo vale anche per trovare l'inversa nel caso in cui A sia invertibile (se tutti gli autovalori $\neq 0$):

$$A^{-1} = \frac{1}{\lambda_1} P_1 + \frac{1}{\lambda_2} P_2 + \dots + \frac{1}{\lambda_s} P_s$$

Di conseguenza possiamo anche calcolare la funzione esponenziale della matrice A con l'aiuto della serie esponenziale:

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} \quad \forall x \in \mathbb{R}$$

Si ha

$$e^A = \sum_{n=0}^{\infty} \frac{A^n}{n!} = \sum_{n=0}^{\infty} \frac{(\lambda_1 P_1, \lambda_2 P_2 \dots \lambda_s P_s)^n}{n!} = \dots = \sum_{n=0}^{\infty} \frac{\lambda_1^n P_1, \lambda_2^n P_2 \dots \lambda_s^n P_s}{n!}$$

La decomposizione spettrale ci permette di estendere da \mathbb{R} a $\mathcal{M}_n(\mathbb{R})$ qualunque funzione se gli autovalori appartengono al dominio della funzione, sempre con l'aiuto delle serie possiamo calcolare la radice quadrata di A , il $\sin(A)$, ecc.

Forme quadratiche

Definizione 10.1. Una forma quadratica da \mathbb{R}^n in \mathbb{R} è un polinomio omogeneo di secondo grado nelle n variabili di \mathbb{R}^n [per omogeneo intendiamo un polinomio di secondo grado che non contenga termini di grado uno o zero (termini misti come xy sono di grado 2)]. In generale una forma quadratica è scritta così:

$$q : \mathbb{R}^n \rightarrow \mathbb{R}$$

$$q : (x_1, \dots, x_n) = \sum_{i=1}^n a_{11} x_i^2 + \sum_{1 \leq i < j \leq n} 2 a_{ij} x_i x_j$$

È importante notare che $i < j$ ciò è dovuto al fatto che se i e j sono ordinati, si evitano ripetizioni. I termini non quadrati, ossia quelli presenti nella seconda sommatoria, prendono il nome di **termini rettangolari**.

OSSERVAZIONE Una forma quadratica q ha $\frac{n(n+1)}{2}$ termini.

ESEMPIO

$$q : \mathbb{R}^3 \rightarrow \mathbb{R}$$

$$q : (x, y, z) = ax^2 + by^2 + cz^2 + dxy + exz + fyz$$

Le forme quadratiche sono di grande interesse perché si possono rappresentare come un prodotto matriciale, infatti:

$$q(\mathbf{v}) = {}^t \mathbf{v} \cdot A \cdot \mathbf{v} \quad (10.1)$$

Dove $\mathbf{v} \in \mathbb{R}^n$ e $A \in \mathcal{M}_n(\mathbb{R})$ è simmetrica; più in generale si rappresenta così:

$$q : (x_1, x_2, \dots, x_n) = \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{12} & \ddots & & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad (10.2)$$

Le forme quadratiche godono di due particolari proprietà:

1. $q(\underline{0}) = 0$
2. $\forall \lambda \in \mathbb{R} : q(\lambda \mathbf{v}) = \lambda^2 q(\mathbf{v})$

ESEMPIO

Consideriamo le forme quadratiche in con $n = 3$ quello che otteniamo è:

$$q(x, y, z) = \begin{bmatrix} x & y & z \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

$$q(x, y, z) = a_{11}x^2 + a_{22}y^2 + a_{33}z^2 + 2a_{12}xy + 2a_{13}xz + 2a_{23}yz$$

Quest'operazione non è unidirezionale, in realtà è completamente reversibile, infatti, da un'espressione quadratica possiamo risalire facilmente alla matrice simmetrica rappresentativa della forma quadratica.

Definizione 10.2. Sia $q : \mathbb{R}^n \rightarrow \mathbb{R}$ una forma quadratica, allora la possiamo classificare a seconda del suo **segno**, in generale può essere:

- ① Definita positiva se $\forall \mathbf{v} \in \mathbb{R}^n - \{\underline{0}\}$ $q(\mathbf{v}) > 0$
- ② Semidefinita positiva se $\forall \mathbf{v} \in \mathbb{R}^n - \{\underline{0}\}$ $q(\mathbf{v}) \geq 0$
- ③ Definita negativa se $\forall \mathbf{v} \in \mathbb{R}^n - \{\underline{0}\}$ $q(\mathbf{v}) < 0$
- ④ Semidefinita negativa se $\forall \mathbf{v} \in \mathbb{R}^n - \{\underline{0}\}$ $q(\mathbf{v}) \leq 0$
- ⑤ Indefinita se $\exists \mathbf{v}, \mathbf{w} \in \mathbb{R}^n - \{\underline{0}\}$ $q(\mathbf{v}) > 0$ e $q(\mathbf{w}) < 0$

Trovare il segno di una forma quadratica può risultare molto facile se i termini sono solo dei quadrati.

ESEMPIO

$$q(x, y, z) = 3x^2 + y^2 + 5z^2 \quad \text{è definita positiva}$$

$$q(x, y, z) = -3x^2 - y^2 - 5z^2 \quad \text{è definita negativa}$$

In questi primi casi possiamo affermare ciò, poiché il quadrato di qualunque numero $\neq 0$ è positivo.

$$q(x, y, z) = 3x^2 + y^2 \quad \text{è semidefinita positiva}$$

$$q(x, y, z) = -3x^2 - 5z^2 \quad \text{è semidefinita negativa}$$

Questi casi sono semidefiniti perché in entrambi i casi esiste un vettore non nullo che finisce in zero (per ogni $k \in \mathbb{R}$ $\mathbf{v} = (0, 0, k)$ e $\mathbf{w} = (0, k, 0)$ rispettivamente)

$$q(x, y, z) = 3x^2 + y^2 - 5z^2 \quad \text{è indefinita}$$

Possiamo dire che è indefinita poiché risulta positiva per tutti i vettori $\mathbf{v} = (t, s, 0)$ mentre è negativa per i vettori $\mathbf{w} = (0, 0, t)$, essendo t, s parametri reali.

Se i termini non sono quadrati, ovvero ci sono termini “misti”, non è facile stabilire il segno di una forma quadratica, per fortuna ci vengono in aiuto diversi teoremi che vedremo in seguito.

Un approccio al problema e che data una forma quadratica possiamo cercare un cambiamento di sistema di riferimento (di base) tale che nel nuovo sistema $q(\mathbf{v}')$ non abbia termini misti, e dunque si possa calcolare facilmente il segno della forma. Se \mathbf{x} è il vettore delle incognite e \mathbf{X} è il vettore delle incognite nel nuovo sistema di riferimento allora $\exists S$ invertibile tale che $\mathbf{x} = S\mathbf{X}$ e di conseguenza:

$$q(\mathbf{x}) = {}^t \mathbf{x} \cdot A \cdot \mathbf{x} = {}^t (S\mathbf{X}) \cdot A \cdot S\mathbf{X} = {}^t \mathbf{X} \cdot {}^t S A S \cdot \mathbf{X} = q'(\mathbf{X})$$

dunque Q e q' sono la stessa forma quadratica espressa in due sistemi di riferimento diversi, A è la matrice che rappresenta q , B quella che rappresenta q' e S è la matrice di passaggio per cui $B = {}^t S A S$

Definizione 10.3. Siano, $A, B \in \mathcal{M}_n(\mathbb{R})$ si dice che A è **congruente** a B se esiste una matrice S invertibile tale che :

$$B = {}^t S \cdot A \cdot S$$

Se S è ortogonale allora $A \sim B$ e si dice **ortogonalmente congruente**.

Teorema 10.1. Le matrice congruenti hanno le seguenti proprietà:

- (1) Se A e B sono congruenti allora le forme quadratiche associate ad A e B hanno lo stesso segno.
- (2) Sia $A \in \mathcal{M}_n(\mathbb{R})$ simmetrica e $q(\mathbf{x}) = {}^t \mathbf{x} \cdot A \cdot \mathbf{x}$ e siano λ_{min} e λ_{max} rispettivamente il più piccolo e il più grande autovalore di A . Allora $\forall \mathbf{v} \in \mathbb{R}^n$ si ha:

$$\lambda_{min} \|\mathbf{v}\|^2 \leq q(\mathbf{v}) \leq \lambda_{max} \|\mathbf{v}\|^2$$

inoltre se \mathbf{v}_{min} è autovettore di λ_{min} è \mathbf{v}_{max} è autovettore di λ_{max} si ha:

$$q(\mathbf{v}_{min}) = \lambda_{min} \|\mathbf{v}\|^2 \quad e \quad q(\mathbf{v}_{max}) = \lambda_{max} \|\mathbf{v}\|^2$$

Corollario 10.1.1. Il segno di $q(\mathbf{x}) = {}^t \mathbf{x} \cdot A \cdot \mathbf{x}$ dipende solo dal segno degli autovalori di A

Definizione 10.4. Sia A una matrice simmetrica di ordine n . Si definisce **segnatura** di A la terna di numeri:

$$(n_+, n_-, n_0) \quad \text{ove ovviamente} \quad n_+ + n_- + n_0 = \text{ordine della matrice } (n)$$

n_+ si dice *indice di positività* e indica il numero degli autovalori strettamente positivi contati con la loro molteplicità algebrica.

n_- è l'*indice di negatività*, ossia il numero degli autovalori strettamente negativi contati sempre con la loro molteplicità algebrica.

n_0 è denominato *indice di nullità* e indica il numero degli autovalori nulli, ossia se e quante volte la matrice A ammette l'autovalore $\lambda = 0$, questo come sappiamo è la dimensione del nucleo dell'applicazione lineare definita dalla matrice.

Teorema 10.2 (LEGGE DI INERZIA DI SYLVESTER). Due matrici reali diagonali congruenti hanno lo stesso rango, e la stessa *segnatura*, ovvero lo stesso numero di autovalori positivi, lo stesso numero di autovalori negativi, e lo stesso numero di autovalori nulli.

Teorema 10.3 (PRIMO TEOREMA DI DEBREU). Sia data una forma quadratica $q(\mathbf{v}) = {}^t \mathbf{v} \cdot A \cdot \mathbf{v}$ con $A \in \mathcal{M}_n(\mathbb{R})$ simmetrica, allora:

- La forma quadratica è definita positiva se e solo se tutti gli autovalori della matrice A sono maggiori di 0, ossia la matrice è definita positiva solo se ha una segnatura $(n, 0, 0)$
- La forma quadratica è definita negativa se e solo se tutti gli autovalori della matrice A sono minori di 0, ossia la matrice è definita negativa solo se ha una segnatura $(0, n, 0)$
- La forma quadratica è semidefinita positiva, se e solo se tutti gli autovalori della matrice A sono maggiori di 0 e ne esiste almeno uno uguale a 0, ossia la matrice ha segnatura $(a, 0, b)$ con $a + b = n, a, b > 0$
- La forma quadratica è semidefinita negativa, se e solo se tutti gli autovalori della matrice A sono minori di 0 e ne esiste almeno uno uguale a 0, ossia la matrice ha segnatura $(0, a, b)$ con $a + b = n, a, b > 0$

Definizione 10.5. Chiamiamo **minori principali dominanti** i minori contenenti a_{11} che possono essere estratti dalla diagonale principale.

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{bmatrix} \quad \text{i minori principali dominanti sono: } [1] \quad \begin{bmatrix} 1 & 2 \\ 5 & 6 \end{bmatrix} \quad \begin{bmatrix} 1 & 2 & 3 \\ 5 & 6 & 7 \\ 9 & 10 & 11 \end{bmatrix} \dots$$

Teorema 10.4 (SECONDO TEOREMA DI DEBREU). Sia data una forma quadratica $q(\mathbf{v}) = {}^t \mathbf{v} \cdot A \cdot \mathbf{v}$ con $A \in \mathcal{M}_n(\mathbb{R})$ simmetrica, allora:

- La forma quadratica è definita positiva se soltanto se tutti i minori principali dominanti hanno $\det > 0$.
- La forma quadratica è semidefinita positiva se soltanto se tutti i minori principali hanno $\det \geq 0$.
- La forma quadratica è definita negativa se soltanto se i minori principali dominanti di ordine pari hanno $\det > 0$ e quelli di ordine dispari hanno un $\det < 0$
- La forma quadratica è semidefinita negativa se soltanto se i minori principali di ordine pari hanno $\det \geq 0$ e quelli di ordine dispari hanno un $\det \leq 0$
- In tutti gli altri casi è indefinita.

Questo teorema fornisce le basi per la dimostrazione di un teorema di vitale importanza del corso di Analisi 2, la matrice hessiana.

10.1 Quadriche

Sia $p(x)$ un polinomio di II grado in n variabili, ossia $\mathbf{x} = (x_1, \dots, x_n)$, chiamiamo **quadriche** il luogo degli zeri di $p(x)$, cioè :

$$Q = \{(x_1, \dots, x_n) \in \mathbb{R}^n \mid p(x_1, \dots, x_n) = 0\}$$

In particolare chiamiamo **coniche** il luogo dei punti che annullano il polinomio di II grado in \mathbb{R}^2 :

$$C = \{(x, y) \in \mathbb{R}^2 \mid p(x, y) = 0\}$$

Le coniche prendono questo nome poiché possono essere ottenute intersecando opportunamente un cono e un piano.

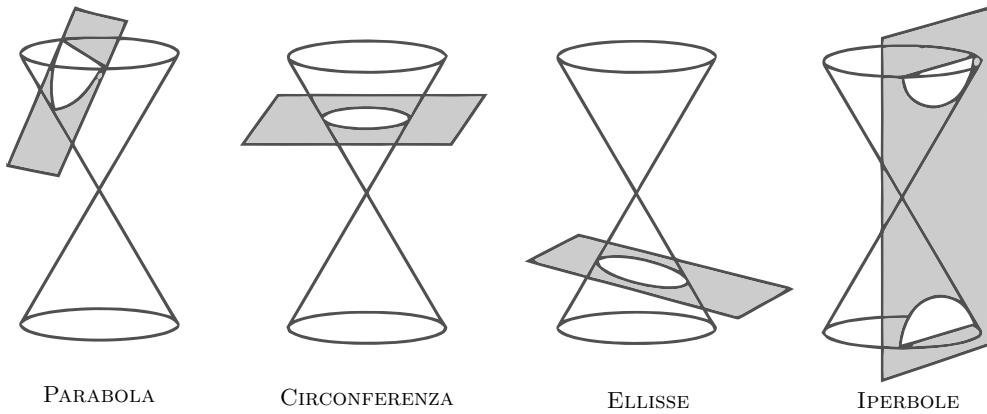


Figura 10.1: Il motivo per cui si chiamano coniche

Le coniche rivestono un'importanza fondamentale in diversi ambiti scientifici come la fisica, ad esempio Eulero scoprì che tutti i corpi celesti nello spazio si muovono su traiettorie coniche, le circonference e le ellissi descrivono le traiettorie chiuse e, l'iperbole e la parabola le traiettorie aperte [Teoria del movimento dei pianeti e delle comete (1744)].

Sia $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ tale che $\exists S \in \mathcal{M}_n(\mathbb{R})$ invertibile e $\mathbf{v} \in \mathbb{R}^n$ se:

$$F(\mathbf{x}) = S\mathbf{x} + \mathbf{v}$$

allora F si chiama **affinità**. Un'affinità non è altro che la composizione di una applicazione lineare con una traslazione.

Definizione 10.6. Sia $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ un'affinità $F(\mathbf{x}) = S\mathbf{x} + \mathbf{v}$, se S è una matrice ortogonale con $\det(S) = 1$ allora F si chiama **rototraslazione**.

Osservazione 10.1. Se $p(\mathbf{x})$ è un polinomio di II grado in $\mathbf{x} = (x_1, x_2, \dots, x_n)$ allora $\exists A \in \mathcal{M}_n(\mathbb{R})$ simmetrica, $\mathbf{b} \in \mathbb{R}^n$, $c \in \mathbb{R}$ tale che la forma matriciale di una quadrica può essere scritta come:

$$p(\mathbf{x}) = \underbrace{{}^t \mathbf{x} \cdot A \cdot \mathbf{x}}_{\text{parte quadratica}} + \underbrace{2 {}^t \mathbf{b} \cdot \mathbf{x}}_{\text{parte lineare}} + \underbrace{c}_{\text{costante}}$$

Si può scrivere anche con quella che chiamiamo **matrice completa della quadrica** (B).

$$p(\mathbf{x}) = [\mathbf{x} \quad 1] \underbrace{\begin{bmatrix} A & {}^t \mathbf{b} \\ \mathbf{b} & c \end{bmatrix}}_B \begin{bmatrix} \mathbf{x} \\ 1 \end{bmatrix}$$

Teorema 10.5 (DEL CAMBIAMENTO DI EQUAZIONI PER AFFINITÀ). Sia $q(\mathbf{x}) = {}^t \mathbf{x} \cdot A \cdot \mathbf{x} + 2 {}^t \mathbf{b} \cdot \mathbf{x} + c$ un polinomio di II grado e sia $\mathbf{x} = S\mathbf{X} + \mathbf{v}$ un'affinità, allora $\hat{q}(\mathbf{x}) = q(S\mathbf{X} + \mathbf{v})$ e si ha:

$$\hat{q} = {}^t \mathbf{X} \cdot \hat{A} \cdot \mathbf{X} + 2 {}^t \hat{\mathbf{b}} \cdot \mathbf{X} + \hat{c}$$

Ove

$$\hat{A} = {}^t S \cdot A \cdot S \quad \hat{\mathbf{b}} = {}^t S \cdot (A \cdot \mathbf{v} + \mathbf{b}) \cdot S \quad \hat{c} = {}^t \mathbf{v} \cdot A \cdot \mathbf{v} + 2 {}^t \mathbf{b} \cdot \mathbf{v} + c$$

Teorema 10.6 (Forma canonica dei polinomi di II grado). Sia $q(\mathbf{x}) = {}^t \mathbf{x} \cdot A \cdot \mathbf{x} + 2 {}^t \mathbf{b} \cdot \mathbf{x} + c = {}^t \mathbf{z} \cdot B \cdot \mathbf{z}$ con $\mathbf{x} = x_1, \dots, x_n$ e $\mathbf{z} = \begin{bmatrix} \mathbf{x} \\ 1 \end{bmatrix}$ con $A \in \mathcal{M}_n(\mathbb{R})$, $B \in \mathcal{M}_{n+1}(\mathbb{R})$, $\mathbf{b} \in \mathbb{R}^n$ e $c \in \mathbb{R}$, allora gli autovalori di A , $\text{rk}(A)$, $\text{rk}(B)$, $\det(B)$ sono tutti invarianti per rototraslazione. Inoltre se $\text{rk}(A) = r$ e $\lambda_1, \dots, \lambda_r$ sono gli autovalori non nulli di A allora:

$$r \leq \text{rk}(B) \leq (r + 2)$$

ed esiste una rototraslazione tale che il polinomio $q(\mathbf{x})$ si trasforma in $\hat{q}(\mathbf{x})$ fatto così:

1 Se $\text{rk}(B) = r$ allora:

$$\hat{q}(\mathbf{x}) = \lambda_1 \mathbf{X}_1^2 + \lambda_2 \mathbf{X}_2^2 + \dots + \lambda_r \mathbf{X}_r^2$$

2 Se $\text{rk}(B) = r + 1$ allora:

$$\hat{q}(\mathbf{x}) = \lambda_1 \mathbf{X}_1^2 + \lambda_2 \mathbf{X}_2^2 + \dots + \lambda_r \mathbf{X}_r^2 + \hat{c} \quad \hat{c} \neq 0$$

3 Se $\text{rk}(B) = r + 2$ allora:

$$\hat{q}(\mathbf{x}) = \lambda_1 \mathbf{X}_1^2 + \lambda_2 \mathbf{X}_2^2 + \dots + \lambda_r \mathbf{X}_r^2 + 2p \mathbf{X}_{r+1} \quad p \neq 0$$

10.2 Classificazione delle coniche

Ora analizzeremo le coniche, che come detto prima è un caso particolare delle quadriche in cui il luogo degli zeri si trova in \mathbb{R}^2 . Per quanto detto precedentemente, possiamo associare ad una conica $p(x, y) = 0$ una espressione del tipo

$$a_{11}x^2 + 2a_{12}xy + a_{22}y^2 + 2a_{13}x + 2a_{23}y + a_{33} = 0$$

che si ottiene da $p(\mathbf{x}) = {}^t\mathbf{x} \cdot A \cdot \mathbf{x} + 2 {}^t\mathbf{b} \cdot \mathbf{x} + c$. Ove A è una matrice simmetrica di dimensione 2×2 , fatta così:

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} \text{coeff. } x^2 & 1/2 \text{ coeff. } xy \\ 1/2 \text{ coeff. } xy & \text{coeff. } y^2 \end{bmatrix}$$

e \mathbf{b} è il vettore colonna dei termini lineari, c è il termine costante

$$\mathbf{b} = \begin{bmatrix} a_{13} \\ a_{23} \end{bmatrix} = \begin{bmatrix} 1/2 \text{ coeff. } x \\ 1/2 \text{ coeff. } y \end{bmatrix} \quad c = a_{33}$$

l'espressione della conica può essere ottenuta anche dall'espressione matriciale

$$p(\mathbf{x}) = [\mathbf{x} \ 1] \cdot B \cdot \begin{bmatrix} \mathbf{x} \\ 1 \end{bmatrix}$$

con la matrice completa della conica fatta in questo modo:

$$B = \begin{bmatrix} A & \mathbf{b} \\ {}^t\mathbf{b} & c \end{bmatrix}$$

cioè la matrice B dei coefficienti è la matrice:

$$B = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{bmatrix}$$

Per classificare le coniche dobbiamo definire quelli che vengono chiamati invarianti, che prendono questo nome proprio perché non cambiano con le eventuali rotazioni e traslazioni. Gli invarianti sono numeri che ci aiutano a identificare il tipo di conica con il solo uso delle matrici rappresentative ricavate dall'equazione associata alla conica.

Gli invarianti sono tre e gli chiamiamo:

- **Invariante cubico:** $I_3 = \det(B)$
- **Invariante quadratico:** $I_2 = \det(A) = \lambda_1 \cdot \lambda_2$ (prodotto degli autovalori di A)
- **Invariante lineare:** $I_1 = \text{tr}(A) = \lambda_1 + \lambda_2$ (somma degli autovalori di A)

In generale le coniche si suddividono in due grandi categorie, la prima prende il nome di **coniche degeneri** e la seconda prende il nome di **coniche non degeneri**, una conica viene classificata in una di queste categorie a seconda che la sua matrice B abbia rango massimo o meno, cioè come criterio per la classificazione utilizziamo l'invariante cubico.

10.2.1 Coniche degeneri

Le coniche degeneri sono tutte quelle coniche che hanno la matrice B singolare ossia:

$$\text{Coniche DEGENERI} \iff \text{rk}(B) < 3 \iff \det(B) = 0 \iff I_3 = 0$$

- (1) $I_2 \neq 0$ in questo caso si dice che la conica è **semplicemente degenera** e $\text{rk}(B) = 2 = \text{rk}(A)$.

Questo vuol dire che A ha 2 autovalori reali (per il teorema spettrale) non nulli ($\lambda = 0$ se e soltanto se A non è invertibile o equivalentemente non ha rango massimo). La sua forma canonica è dunque:

$$\lambda_1 X^2 + \lambda_2 Y^2 = 0$$

Questo da origine a due casi:

- (a) Se $I_2 < 0$ vuol dire che λ_1 e λ_2 sono discordi e si tratta di *due rette reali incidenti* di equazione canonica:

$$\text{RETTE REALI} \quad X = \sqrt{\frac{-\lambda_2}{\lambda_1}} Y \quad X = -\sqrt{\frac{-\lambda_2}{\lambda_1}} Y$$

Se inoltre $I_1 = 0$ (equivalentemente $\lambda_1 + \lambda_2 = 0$) le rette sono perpendicolari.

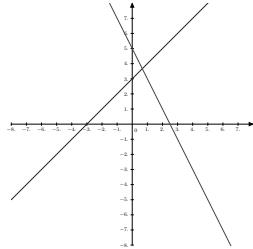


Figura 10.2: Rette incidenti

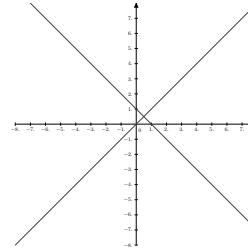


Figura 10.3: Rette perpendicolari

- (b) Se $I_2 > 0$ vuol dire che λ_1 e λ_2 sono concordi e si tratta di *due rette immaginarie incidenti* di forma canonica:

$$\text{RETTE IMAGINARIE} \quad X = i\sqrt{\frac{-\lambda_2}{\lambda_1}} Y \quad X = -i\sqrt{\frac{-\lambda_2}{\lambda_1}} Y$$

per cui stiamo in pratica sono due rette complesse coniugate il cui punto di intersezione è l'origine degli assi e risulta l'unico punto reale della conica.

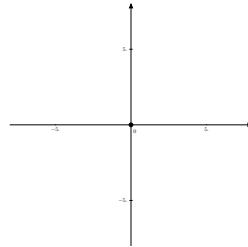


Figura 10.4: Rette immaginarie coincidenti

② Se $I_2 = 0$ la conica si dice **doppiamente degenera** e ciò implica che $\text{rk}(B) = 1$. Anche qui possiamo incontrare due casi:

- (a) Se $\text{rk}(A) = 1 = \text{rk}(B)$ questo vuol dire che l'unica riga linearmente indipendente è compresa nella matrice A per cui la matrice A avrà un autovalore nullo e $\lambda_1 \neq 0$, da cui la forma canonica della quadrica:

$$\text{RETTE COINCIDENTI} \quad \lambda_1 X^2 = 0$$

in realtà si tratta di due rette coincidenti, cioè è una retta contata due volte.

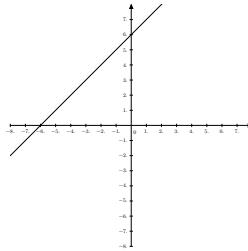


Figura 10.5: Rette coincidenti

- ③ Se $\text{rk}(A) = 1$ e $\text{rk}(B) = 2$ si tratta di due rette parallele che potrebbero essere complesse coniugate ed in tal caso la conica non avrebbe punti reali. In generale:

$$\lambda_1 X^2 + \hat{c} = 0$$

Le rispettive forme canoniche sono:

$$\text{RETTE REALI PARALLELE} \quad X = \sqrt{\frac{-\hat{c}}{\lambda_1}} \quad X = -\sqrt{\frac{-\hat{c}}{\lambda_1}}$$

$$\text{RETTE IMMAGINARIE PARALLELE} \quad X = i\sqrt{\frac{-\hat{c}}{\lambda_1}} \quad X = -i\sqrt{\frac{-\hat{c}}{\lambda_1}}$$

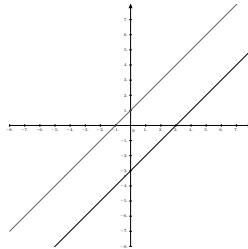


Figura 10.6: Rette reali parallele

ESEMPIO

Consideriamo la conica di equazione $2x^2 - 2y^2 - 2x + 2y = 0$

La conica avrà le matrici associate A e B fatte così:

$$A = \begin{bmatrix} 2 & 0 & -1 \\ 0 & -2 & 1 \\ -1 & 1 & 0 \end{bmatrix} \quad B = \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix}$$

Calcoliamo gli invarianti:

$$I_3 = -(-2) - 2 = 0$$

poiché $I_3 = 0$ si conclude che è una conica degenere.

$$I_2 = -4$$

poiché $I_2 < 0$ si conclude si tratta di una coppia di rette reali.

$$I_1 = 18 + 12 - 20 = 10$$

poiché $I_1 = 0$ si tratta di due rette perpendicolari.

Con qualche manipolazione algebrica ci possiamo ricondurre alla forma canonica delle rette:

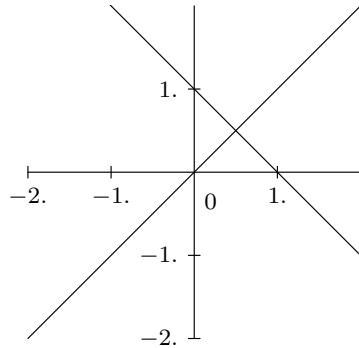
$$2x^2 - 2y^2 - 2x + 2y = 0 \quad 2(x^2 - y^2 - x + y) = 0$$

$$x^2 - y^2 - x + y = 0 \quad (x + y)(x - y) - x + y = 0$$

$$(x + y)(x - y) - (x - y) = 0 \quad (x + y - 1)(x - y) = 0$$

le rette sono: $\begin{cases} y = x \\ y = -x + 1 \end{cases}$

In un piano la conica è fatta così:



10.2.2 Coniche non degeneri

Le coniche non degeneri sono tutte quelle coniche la cui matrice B ha rango massimo ossia:

$$\text{Coniche NON DEGENERI} \iff \text{rk}(B) = 3 \iff \det(B) \neq 0 \iff I_3 \neq 0$$

① Se $I_2 \neq 0$ questo significa che A ha due autovalori $\neq 0$ in generale si ha:

$$\lambda_1 x^2 + \lambda_2 y^2 + \hat{c} = 0 \quad \text{con} \quad \hat{c} = \frac{I_3}{I_2} \quad \text{poiché} \quad \det(B) = \det \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \hat{c} \end{bmatrix}$$

dividendo per \hat{c} si ha:

$$eX^2 + fY^2 = 1 \quad \text{ove} \quad e = \frac{-\lambda_1 I_2}{I_3} \quad f = \frac{-\lambda_2 I_2}{I_3}$$

Possiamo dunque distinguere 3 casi:

(a) Se $I_2 < 0$ ciò significa che e ed f sono discordi in segno e si tratta di un'**iperbole**.

Se inoltre $I_1 = 0$ si tratta di un'**iperbole equilatera**. In forma canonica:

$$\text{IPERBOLE} \quad \frac{X^2}{a^2} - \frac{Y^2}{b^2} = 1$$

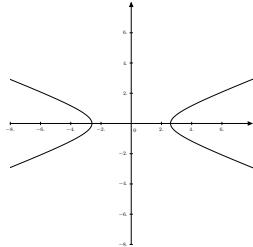


Figura 10.7: Iperbole

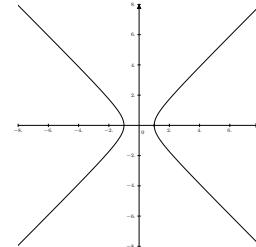


Figura 10.8: Iperbole equilatera

(b) Se $I_2 > 0$ significa che e ed f hanno lo stesso segno.

Se $(I_1 \cdot I_3) < 0$ allora e ed f sono entrambi positivi e si tratta di un'**ellisse a punti reali**. In forma canonica:

$$\text{ELISSE REALE:} \quad \frac{X^2}{a^2} + \frac{Y^2}{b^2} = 1$$

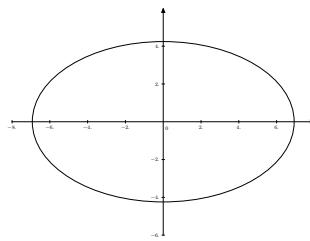


Figura 10.9: Ellisse reale

Se $(I_1 \cdot I_3) > 0$ allora e ed f sono entrambi negativi e si tratta di un'**ellisse immaginaria**. In forma canonica:

$$\text{ELISSE IMMAGINARIA: } \frac{X^2}{a^2} + \frac{Y^2}{b^2} = -1$$

Il caso particolare dell'ellisse è la **circonferenza** che si ottiene quando $a_{11} = a_{22}$ e $a_{12} = 0 = a_{21}$. In forma canonica:

$$\text{CIRCONFERENZA } X^2 + Y^2 = 1$$

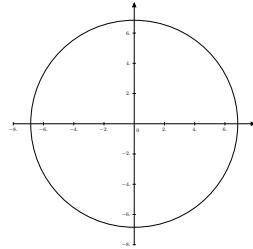


Figura 10.10: Circonferenza

- ② Se $I_2 = 0$ allora $\text{rk}(B) = 3$ e $\text{rk}(A) = 1$, ciò implica che A abbia un autovalore $\lambda_1 \neq 0$ è un altro nullo. Questa conica prende il nome di *parabola* e in forma canonica:

$$\text{PARABOLA } \lambda_1 X^2 + 2pY = 0 \quad p \neq 0$$

Così come per il calcolo di \hat{c} , per calcolare p si sfrutta il fatto che il $\det(B)$ sia un invariante e si trova il valore di p per il quale la matrice:

$$\begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & 0 & p \\ 0 & p & 0 \end{bmatrix} \quad p = \sqrt{\frac{I_3}{-\lambda_1}}$$

abbia lo stesso determinante della matrice B .

Le parabole possono essere inoltre rivolte verso l'alto o verso il basso a seconda del segno dell'autovalore λ_1 se è positivo la parabola sarà rivolta verso l'alto e se è negativo la parabola sarà rivolta verso il basso.

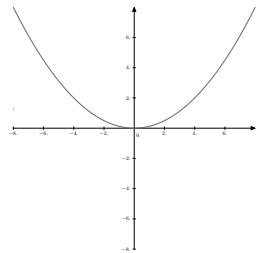


Figura 10.11: Parabola rivolta verso l'alto

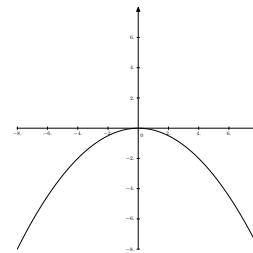


Figura 10.12: Parabola rivolta verso il basso

ESEMPIO

Consideriamo la conica di equazione $18x^2 + 8xy + 12y^2 = 20$

La conica avrà le matrici associate A e B fatte così:

$$B = \begin{bmatrix} 18 & 4 & 0 \\ 4 & 12 & 0 \\ 0 & 0 & -20 \end{bmatrix} \quad A = \begin{bmatrix} 18 & 4 \\ 4 & 12 \end{bmatrix}$$

Calcoliamo gli invarianti:

$$I_3 = -20 \cdot [18(12) - 4(4)] = -4000$$

poiché $I_3 \neq 0$ si conclude che è una conica non degenere.

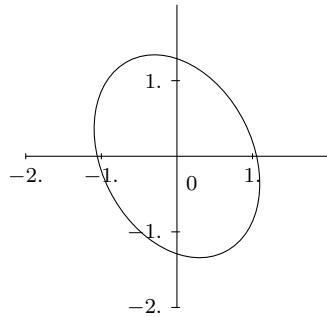
$$I_2 = [18(12) - 4(4)] = 200$$

poiché $I_2 > 0$ si conclude che è un'ellisse.

$$I_1 = 18 + 12 - 20 = 10$$

poiché $I_1 \cdot I_3 < 0$ si tratta di un'ellisse reale.

In un piano è un'ellisse fatta così:



10.3 Centro, assi e rototraslazioni

Centro

Definizione 10.7. Se $I_2 \neq 0$ allora si dice che sono **coniche a centro**. Il centro può essere trovato risolvendo il sistema:

$$A \cdot \begin{bmatrix} x \\ y \end{bmatrix} = -\mathbf{b}$$

Nel caso si abbia una coppia di rette incidenti, allora lo stesso sistema fornisce il punto di incidenza delle due rette, per cui si considera anch'essa una conica a centro.

Assi

Per le coniche non degeneri con $I_2 \neq 0$, ossia per l'ellisse e le iperbole gli assi sono le rette passanti per il centro, e parallele agli autovettori di A , per cui gli assi saranno esprimibili facilmente in forma parametrica.

ESEMPIO

Se il centro è $c = (1, 3)$ e gli autovettori sono $\mathbf{v} = (2, 1)$ e $\mathbf{w} = (1, -2)$ gli assi sono:

$$a_1 = \begin{cases} x = 1 + 2t \\ y = 3 + t \end{cases} \rightarrow 2y = x + 5$$

$$a_2 = \begin{cases} x = 1 + t \\ y = 3 - 2t \end{cases} \rightarrow y = 5 - 2x$$

Per le coniche non degeneri con $I_2 = 0$, ossia per le parabole, esiste solo un asse ed è la retta passante per il vertice di direzione parallela all'autovettore relativo all'autovalore nullo. Per trovare l'asse quello che si fa è trovare il fascio di rette perpendicolari all'asse e imponendo il passaggio per un punto qualsiasi ossia

$$\begin{cases} x = x_0 + lt \\ y = y_0 + mt \end{cases} \quad \text{ove } l, m \text{ sono i coefficienti relativi all'autovalore non nullo (equivalente alla direzione perpendicolare all'autovettore relativo all'autovalore nullo)}$$

Per rendere i conti più agevoli scegiamo x_0 e y_0 uguali a 0 e avremo un'equazione del tipo

$$\frac{x}{l} = \frac{y}{m} \rightarrow mx - lm = 0$$

Mettendo a sistema quest'ultima equazione con l'equazione della parabola in forma quadratica otterremo i due punti dove la parabola interseca la retta perpendicolare all'asse, anche se questi punti possono essere immaginari, quello che ci permette di calcolare l'asse è il punto medio, e il punto medio è sempre reale

$$x_M = \frac{x_1 + x_2}{2} \in \mathbb{R} \quad y_M = \frac{y_1 + y_2}{2} \in \mathbb{R}$$

Per trovare l'asse quello che ci resta è scrivere l'equazione dell'asse in forma parametrica imponendo il passaggio per il punto medio è con direzione parallela all'autovettore relativo all'autovalore nullo. Ossia

$$a = \begin{cases} x = x_M + t \\ y = y_M + s \end{cases}$$

Inoltre possiamo determinare il **vertice** della parabola che è determinato dal punto d'intersezione tra l'asse e la parabola.

Rototraslazione

Come sappiamo dalla teoria, possiamo portare l'equazione della conica in forma canonica con un opportuno cambio di sistema di riferimento, per fare ciò abbiamo bisogno di effettuare una rototraslazione, ovvero un'affinità che abbia come matrice rappresentativa una matrice ortogonale speciale (con $\det = 1$), questa affinità "raddrizza" la conica e porta il suo centro nell'origine del nuovo sistema di riferimento. Per capire se la matrice ha bisogno di una traslazione, rotazione o entrambe, ci basta vedere la matrice completa della conica, se la matrice completa ha termini "misti" allora è necessario una rotazione, ciò si traduce nel fatto che A abbia termini non nulli al di

fuori della diagonale principale, di conseguenza se la matrice A è diagonale non c'è bisogno di una rotazione. Per quanto riguarda la traslazione ci basta vedere il vettore colonna \mathbf{b} , se è nullo allora non ha bisogno di una traslazione, altrimenti sì. La conica sarà esprimibile in forma canonica con un cambio di sistema di riferimento rappresentato da una matrice Q tale che

$$\begin{bmatrix} x \\ y \end{bmatrix} = Q \cdot \begin{bmatrix} x' \\ y' \end{bmatrix} + \begin{bmatrix} ax' \\ by' \end{bmatrix}$$

Traslazione

Ci sono due modi per effettuare la traslazione, con il completamento del quadrato o basandoci sulla posizione del centro:

Completamento del quadrato

Questo è un metodo molto facile ed efficace, consiste semplicemente nel completare algebricamente i quadrati aggiungendo e togliendo dei termini noti. Ossia normalmente si ha un'espressione nella forma

$$x^2 + ax$$

dobbiamo sommare e sottrarre una stessa quantità, in modo da non alterare l'espressione

$$x^2 + ax + c - c$$

e poi ricondurci alla sviluppo del binomio che ha forma

$$(x + x_0)^2 = x^2 + 2x_0x + x_0^2$$

ESEMPIO

Troviamo la traslazione da effettuare nella conica di equazione $x^2 + 6x + 2y^2 - 8y + 2 = 0$.

Aggiungendo e togliendo termini opportunamente si ha:

$$x^2 + 6x + 9 - 9 + 2(y^2 - 4y + 4 - 4) + 2 = 0$$

$$x^2 + 6x + 9 + 2(y^2 - 4y - 4) - 9 - 8 + 2 = 0$$

$$(x + 3)^2 + 2(y - 2)^2 - 15 = 0$$

Per cui la traslazione da effettuare è:

$$X = x + 3 \quad Y = y - 2$$

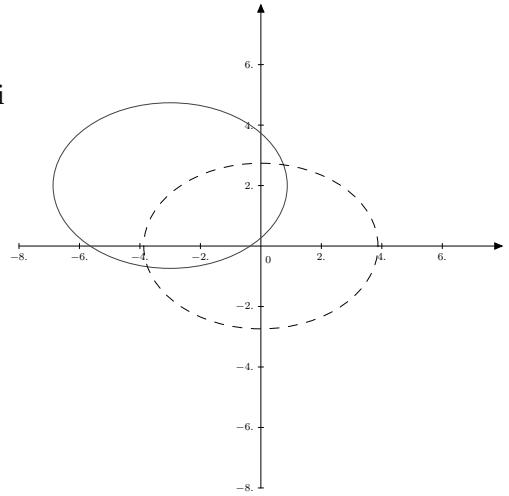


Figura 10.13: Traslazione dell'ellisse

Con l'utilizzo del centro

Questo metodo è quello più immediato se si deve fare soltanto la traslazione e non la rotazione, poiché ci basta vedere di quanto dobbiamo trascinare il centro per portarlo nell'origine, ossia basta sommare $-c$; questo metodo si può applicare anche ai casi che hanno bisogno di traslazione e rotazione ma bisogna fare attenzione poiché la traslazione cambia i coefficienti con le rotazioni. È importante notare che il metodo non è sempre applicabile poiché per esempio la parabola non è una conica a centro.

ESEMPIO

Riprendendo l'esempio precedente, la conica ha come matrice

$$B = \begin{bmatrix} 1 & 0 & 3 \\ 0 & 2 & -4 \\ 3 & -4 & 2 \end{bmatrix}$$

Per trovare il centro risolviamo

$$\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = - \begin{bmatrix} 3 \\ -4 \end{bmatrix} \quad \text{da cui} \quad \begin{cases} x = -3 \\ y = 2 \end{cases}$$

Per cui possiamo subito dire che:

$$X = x + 3 \quad Y = y - 2$$

Rotazione

Per fare la rotazione quello che dobbiamo fare è calcolare gli autovalori della matrice A per trovare i rispettivi autovettori che costituiranno le colonne della matrice di passaggio P , poi rendiamo ortonormale la matrice P con il procedimento di *Gram – Schmidt* (ricordando che se P ha autovalori tutti diversi ci basta normalizzarli) una volta fatto questo, l'unica cosa che ci resta è renderla “speciale” ovvero cerchiamo la configurazione che dia $\det = 1$ poiché il determinante potrebbe essere anche -1 poiché, se ricordiamo, una matrice ortogonale può avere come determinante ± 1 .

Nel caso sia già 1 lasciamo intatta la matrice, altrimenti se il $\det = -1$ per una delle proprietà alternate del determinante ci basta scambiare le colonne per cambiare il segno del determinante, una volta fatti tutti questi passaggi si ha la matrice Q che rappresenta la rotazione necessaria per “raddrizzare” la conica. Il cambio di sistema di riferimento che effettua la rotazione è dato da:

$$\begin{bmatrix} x \\ y \end{bmatrix} = Q \cdot \begin{bmatrix} x' \\ y' \end{bmatrix}$$

ESEMPIO

Consideriamo la conica

$$x^2 + 6xy + y^2 - 3$$

la matrice associata alla conica è:

$$\begin{bmatrix} 1 & 3 & 0 \\ 3 & 1 & 0 \\ 0 & 0 & -3 \end{bmatrix} \quad I_3 = -3(-8) = 24 \text{ (conica non degenere)} \quad I_2 = -8 \text{ (conica a centro)}$$

In particolare si tratta di una iperbole non equilatera. Per determinare la rotazione diagonalizziamo la matrice A

$$\det \begin{bmatrix} 1-\lambda & 3 \\ 3 & 1-\lambda \end{bmatrix} = (1-\lambda)^2 - 9 = 0 \rightarrow (\lambda - 4)(\lambda + 2) = 0$$

Gli autovalori sono dunque 4 e -2. Calcoliamo gli autovettori:

$$\text{Per } \lambda = 4 \quad \begin{bmatrix} 1-4 & 3 \\ 3 & 1-4 \end{bmatrix} \rightarrow \begin{bmatrix} -3 & 3 \\ 3 & -3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \rightarrow \mathbf{v} = (1, 1)$$

$$\text{Per } \lambda = -2 \quad \begin{bmatrix} 1+2 & 3 \\ 3 & 1+2 \end{bmatrix} \rightarrow \begin{bmatrix} 3 & 3 \\ 3 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \rightarrow \mathbf{w} = (1, -1)$$

Poiché sono autovettori relativi a autovalori diversi, per il teorema spettrale possiamo affermare che sono già perpendicolari per cui ci basta normalizzarli. La matrice P , che ha come colonne gli autovettori, è dunque:

$$\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \xrightarrow{\text{normalizzo}} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}$$

Ora dobbiamo renderla speciale ossia deve essere $\det = 1$ in questo caso il determinante è -1 per cui l'unica cosa da fare è scambiare le colonne.

$$\begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \xrightarrow{\text{la rendo speciale}} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

La rotazione è:

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix} \rightarrow \begin{cases} x = \frac{1}{\sqrt{2}}x' + \frac{1}{\sqrt{2}}y' \\ y = -\frac{1}{\sqrt{2}}x' + \frac{1}{\sqrt{2}}y' \end{cases}$$

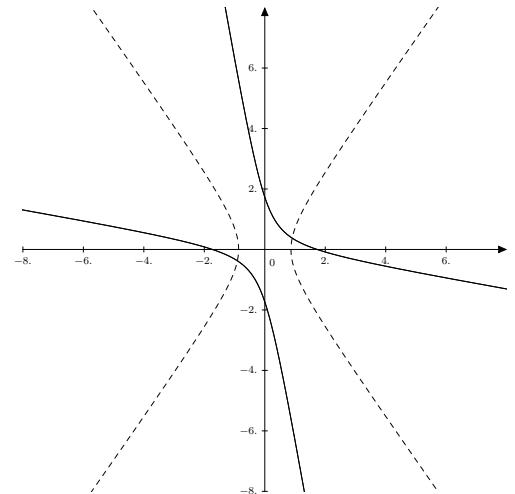


Figura 10.14:
Rotazione dell'iperbole

10.4 Fascio di coniche

Siano date due coniche diverse:

$$\Gamma_1 : f_1(x, y) = 0$$

$$\Gamma_2 : f_2(x, y) = 0$$

chiamiamo **fascio di coniche** tutte le coniche che si ottengono come combinazione lineare di Γ_1 e Γ_2 ossia:

$$F : \lambda\Gamma_1 + \mu\Gamma_2 = 0 \quad \text{con } \lambda, \mu \text{ parametri omogeni}$$

Se supponiamo che nell'equazione del fascio sia $\lambda \neq 0$ possiamo porre $k = \frac{\mu}{\lambda}$ che ci permette di riscrivere l'equazione come:

$$F_k : \Gamma_1 + k\Gamma_2 = 0 \quad \text{con } k \text{ parametro non omogeneo}$$

Se poniamo $k = 0$ si ottiene Γ_1 , convenzionalmente si dice che per $k = \infty$ si ottiene Γ_2 .

NOTA BENE Si noti che per essere un fascio di coniche il parametro k deve essere lineare ossia se l'equazione ha k^2 non può essere in alcun modo un fascio di coniche.

Le coniche Γ_1 e Γ_2 prendono il nome di **coniche base**, queste coniche si intersecano fra loro in un certo numero di punti detti **punti base del fascio**, questi punti sono al minimo 1 e al massimo quattro.

Quello che si fa solitamente con i fasci di coniche è lo studio del fascio, che consiste nel distinguere le possibili coniche che si formano al variare del parametro k

ESEMPIO

Si consideri il fascio:

$$F_t : x^2 + (1 - 2t)y^2 + 2tx + (4 - 2t)y + 1 - t = 0$$

Scrivendo il fascio in forma di matrice si ottiene:

$$B = \begin{bmatrix} 1 & 0 & t \\ 0 & 1 - 2t & 2 - t \\ t & 2 - t & 1 - t \end{bmatrix} \quad A = \begin{bmatrix} 1 & 0 \\ 0 & 1 - 2t \end{bmatrix}$$

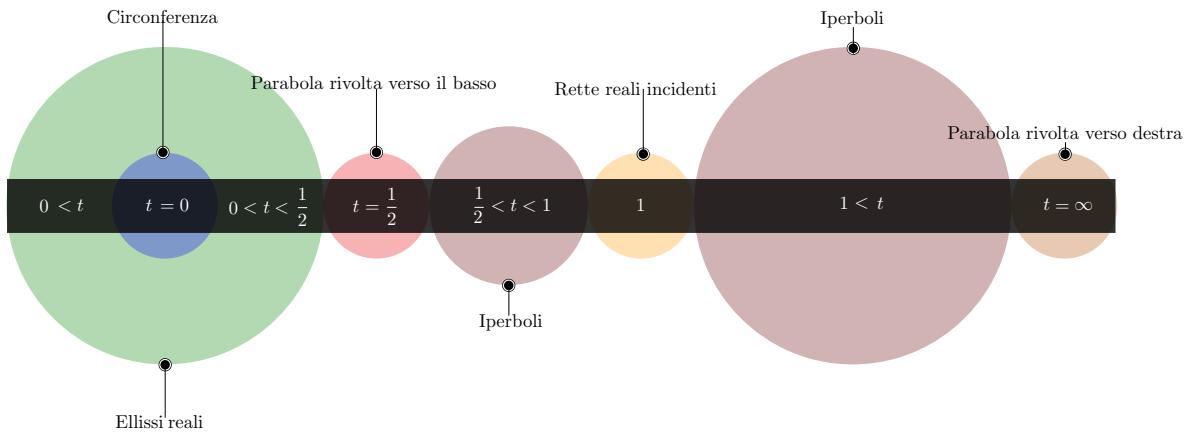
Per distinguere le coniche degeneri dalle non degeneri calcoliamo l'invariante $I_3 = \det(B)$ che con lo sviluppo di Laplace rispetto alla prima colonna risulta:

$$(1 - 2t)(1 - t) - (2 - t)^2 - t^2(1 - 2t) = (t - 1)(2t^2 + 2t + 3)$$

Il determinante si annulla solo per $t = 1$ quindi per $t \neq 1$ si hanno coniche non degeneri, nello specifico per $t = 1$ si ha $x^2 - y^2 + 2x + 2y + 1$ che può essere riscritto come $(x - y + 2)(x + y) = 0$ che non sono altro che due rette reali incidenti in un unico punto, inoltre poiché la traccia è nulla possiamo concludere che le rette sono fra loro perpendicolari.

Analizziamo ora l'invariante I_2 ossia il $\det(A)$, si vede facilmente che il determinante è uguale a $(1 - 2t)$ e che si annulla se $t = \frac{1}{2}$; in questo caso, in particolare, la conica non degenere è una parabola rivolta verso il basso. Per $1 > t > \frac{1}{2}$ e $t > 1$ risulta $I_2 < 0$ per cui si hanno delle iperbole poiché la traccia si annulla solo se $t = \frac{1}{2}$ possiamo concludere che non ci sono iperbole equilateri; analogamente troviamo che per $t < \frac{1}{2}$ si ha $I_2 > 0$ per cui per questi valori si hanno dell'ellissi, in questo caso svolgendo i conti si trovano che sono tutte reali e che per $t = 0$ si ha una circonferenza.

Per concludere dobbiamo vedere cosa succede per “ $t = \infty$ ”, per fare ciò ci basta considerare soltanto i termini dipendenti dal parametro t come una conica, che per il nostro esercizio si perviene all'equazione: $2x - 2y^2 - 2y - 2 = 0$; che è una parabola rivolta verso destra.

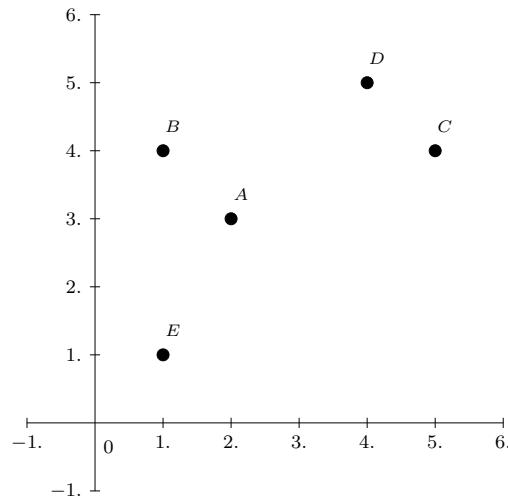


A volte non interessa analizzare una conica o studiare un fascio ma vogliamo piuttosto trovare l'equazione di una determinata conica con la sola conoscenza di alcuni punti, questo non è affatto difficile, infatti basta la sola conoscenza di quelli che abbiamo chiamato punti basi e di un qualunque altro punto che appartenga alla conica.

ESEMPIO

Si considerino i seguenti punti:

- $A = (2, 3)$
- $B = (1, 4)$
- $C = (5, 4)$
- $D = (4, 5)$
- $E = (1, 1)$



Per determinare l'equazione della conica basta sapere che se la conica passa per tutti i 5 punti, in particolare è una delle tante coniche che appartiene a un determinato fascio ove le coniche passano per 4 dei 5 punti, ma è l'unica di quel fascio che tocca il quinto punto. Quindi per trovare l'equazione della conica dobbiamo trovare per prima l'equazione del fascio a cui appartiene, per fare ciò dobbiamo trovare almeno 2 coniche che passino entrambe per gli stessi 4 punti.

Fortunatamente, in questo esempio non è difficile trovare un paio di coniche che passino per gli stessi quattro punti, infatti la coppia di rette parallele che è, come visto nelle pagine precedenti, una conica degenere, si trova senza difficoltà, giacché la retta che attraversa \overline{AB} è parallela alla retta che attraversa \overline{CD} così come la retta che passa per \overline{AC} è parallela alla retta che attraversa \overline{BD} .

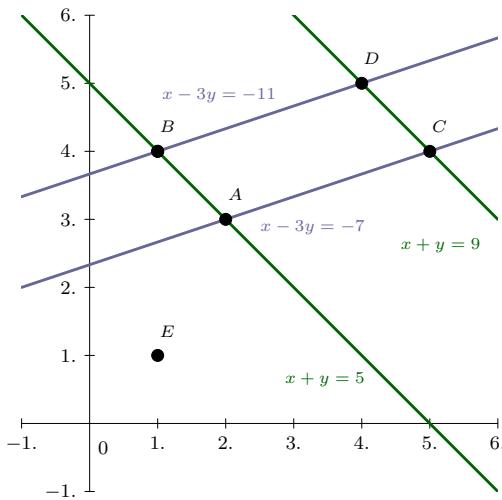


Figura 10.15: Le due coppie di rette parallele

La conica cercata appartiene al fascio:

$$(x - 3y + 11)(x - 3y + 7) + t[(x + y - 5)(x + y - 9)] = 0$$

A questo punto abbiamo già trovato l'equazione del fascio, per trovare qual'è la conica appartenente a questo fascio che incrocia il punto $E = (1, 1)$ basta impostare i valori delle coordinate del punto nell'equazione e risolvere rispetto a t ; si ottiene il risultato:

$$t = -\frac{45}{21}$$

Sostituendo il valore di t nell'equazione del fascio di coniche trovato precedentemente si determina l'equazione della conica che tocca tutti e cinque i punti, che risulta essere :

$$x^2 - 6y^2 + 9xy - 42x + 21y + 17 = 0$$

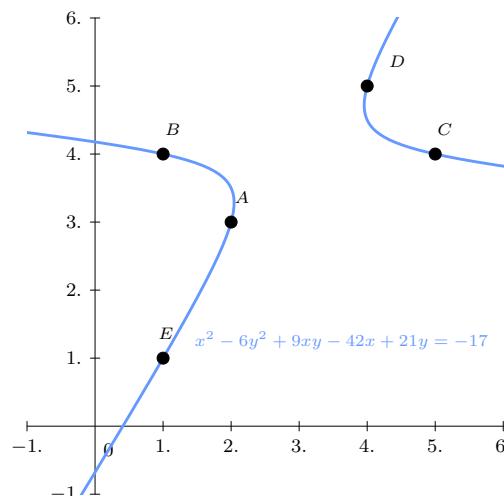


Figura 10.16: La conica cercata è un'iperbole

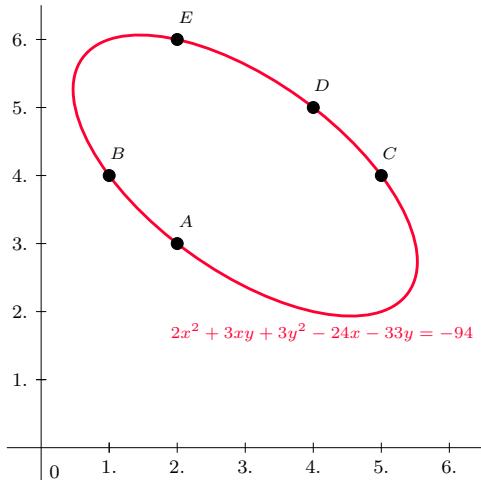


Figura 10.17: Conica passante per un nuovo punto E

Se avessimo scelto diversamente il punto E , non solo cambia l'equazione ma può anche succedere che il tipo di conica cambi oppure non esista una conica che incroci tutti i cinque punti.

Se nell'esempio precedente scegliamo un altro punto $E = (2, 6)$ e imponiamo il passaggio nell'equazione del fascio troviamo in questo caso $t = 15$ e l'equazione della conica è:

$$2x^2 + 3xy + 3y^2 - 24x - 33y + 94 = 0$$

10.5 Quadriche in \mathbb{R}^3

Come abbiamo già visto, una quadrica è il luogo degli zeri di un polinomio di secondo grado in n incognite, dato che al luogo degli zeri di un polinomio in due incognite su \mathbb{R}^2 gli viene dato il nome di conica e dato che il luogo degli zeri dei polinomi su \mathbb{R}^n con dimensione n superiore a tre non ha molto senso giacché non è possibile rappresentare graficamente il luogo, diamo colloquialmente il nome di quadrica soltanto al luogo degli zeri di un polinomio in tre incognite (x, y, z) -ossia nello spazio \mathbb{R}^3 - e coniche a quello in due incognite anche se anch'esse sono quadriche. Ad una quadrica $p(x, y, z) = 0$ viene associata una espressione del tipo:

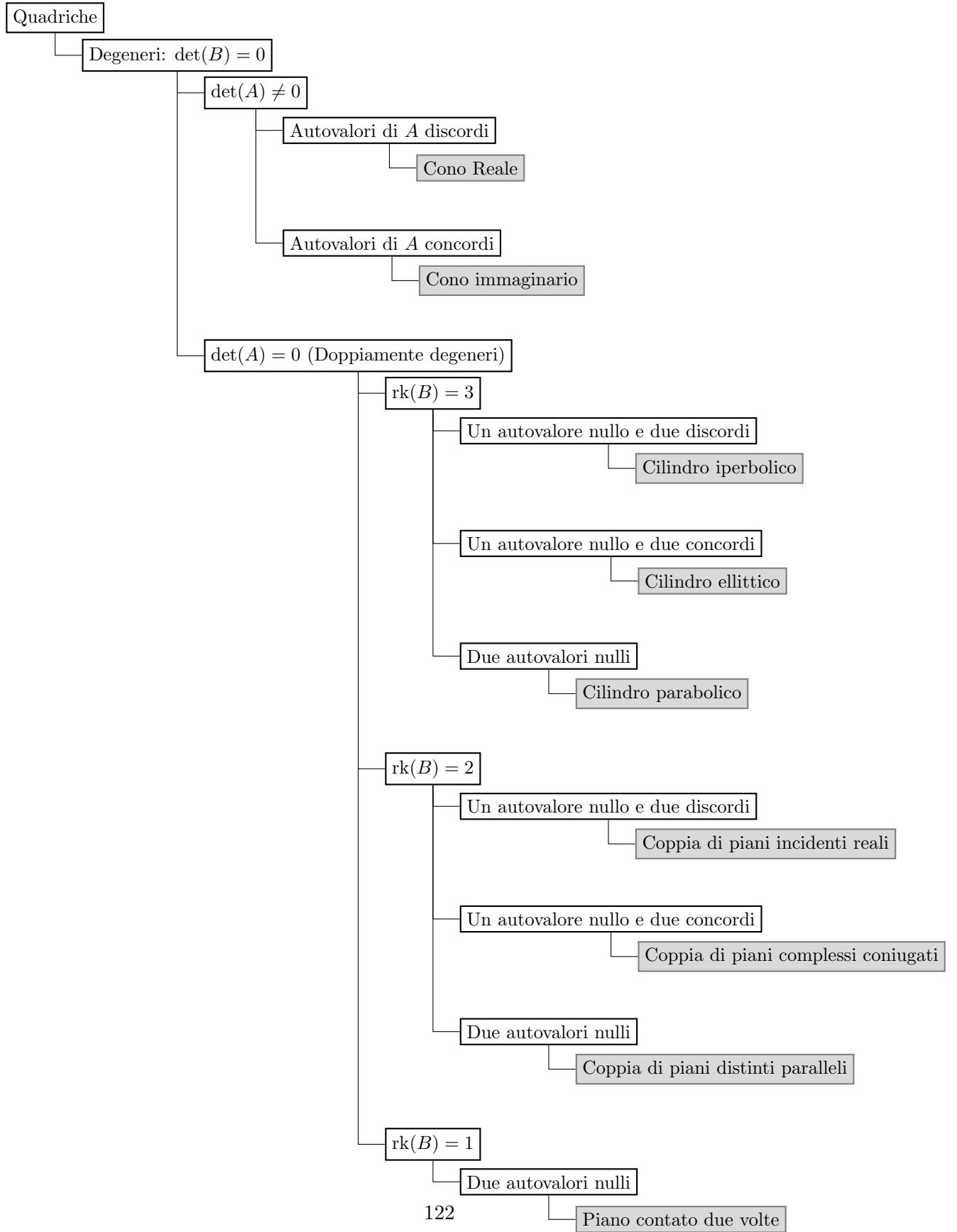
$$a_{11}x^2 + a_{22}y^2 + a_{33}z^2 + 2a_{12}xy + 2a_{13}xz + 2a_{23}yz + 2a_{14}x + 2a_{24}y + 2a_{34}z + a_{44} = 0$$

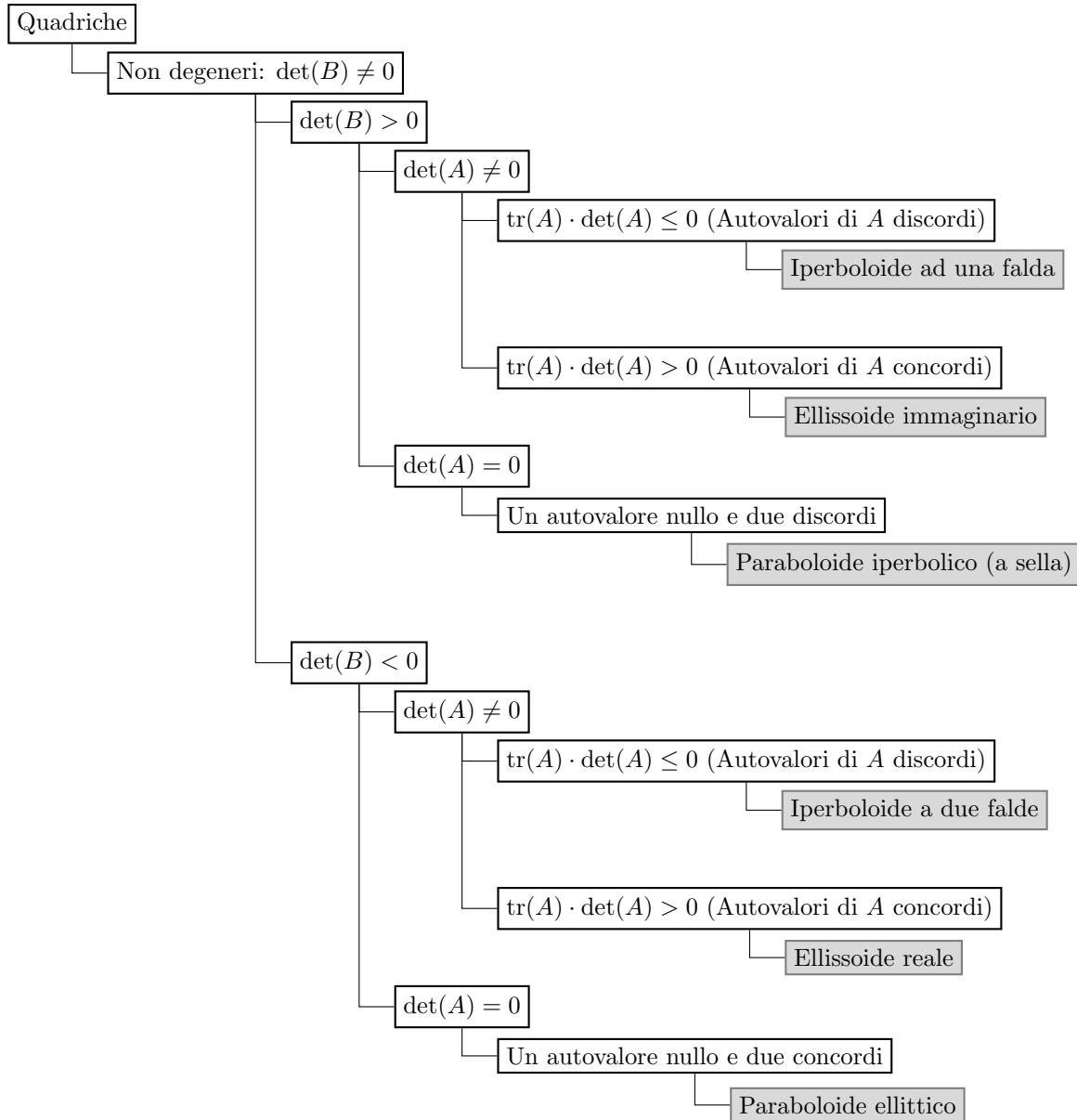
Così come nelle coniche lo studio delle quadriche si riduce allo studio della sua matrice rappresentazione e i suoi invarianti. La matrice associata alla quadrica è una matrice simmetrica data da:

$$B = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix} \quad A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

Per classificare qualsiasi quadrica è sufficiente conoscere il determinante della matrice B , della matrice A e il segno degli autovalori di A .

Proprio come nella classificazione delle coniche le quadriche si dividono in due grandi gruppi: le coniche degeneri e non. Nelle prossime pagine vengono riportate le quadriche e i criteri per la loro rispettiva classificaziond.





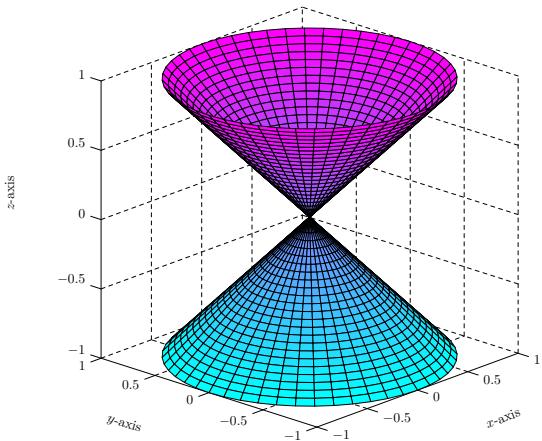
10.5.1 Forme canoniche delle quadriche

Cono Reale

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 0$$

Quando $a = b$ il cono si dice circolare

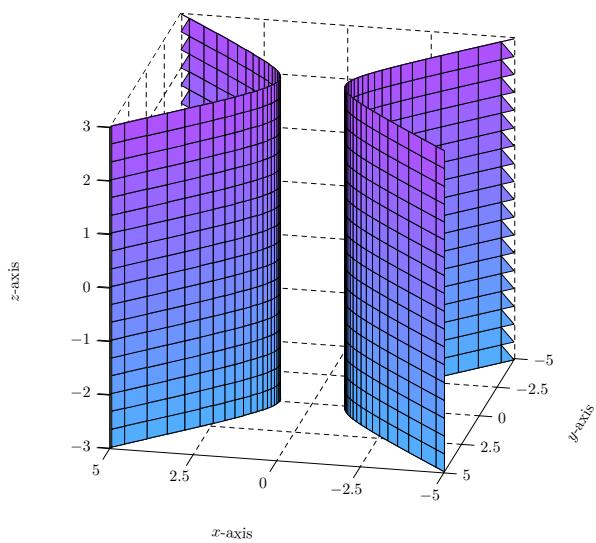
Quando $a \neq b$ il cono si dice ellittico



Cono Immaginario / Complesso

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 0$$

Ha un unico punto reale che è laddove si uniscono le due metà, gli altri punti stanno nel campo complesso.



Cilindro iperbolico

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$$

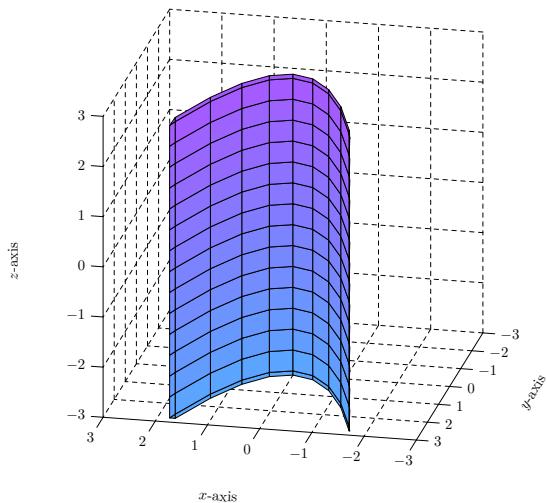
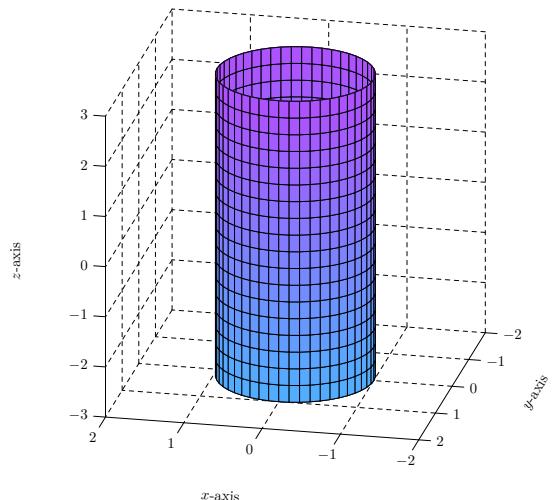
È una coppia simmetrica di piani curvi le cui intersezioni con i piani $z = h$ sono tutte delle iperboli.

Cilindro ellittico

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

Quando $a = b$ il cilindro si dice circolare

Quando $a \neq b$ il cilindro si dice ellittico



Cilindro parabolico

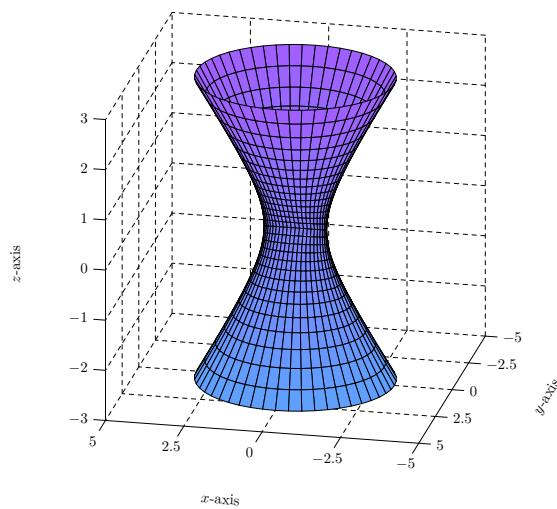
$$\frac{x^2}{a^2} - 2y = 0$$

È una superficie le cui intersezioni con i piani $z = h$ sono tutte delle parabole se $a > 0$.

Iperboloide iperbolico “ad una falda”

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1$$

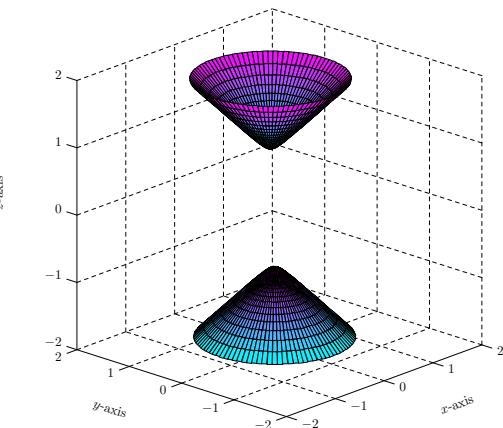
Le sue intersezioni con i piani $x = h$ e $y = h$ sono tutte delle iperboli, le intersezioni con i piani $z = h$ sono invece tutte delle ellissi.



Iperboloide ellittico “a due falde”

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = -1$$

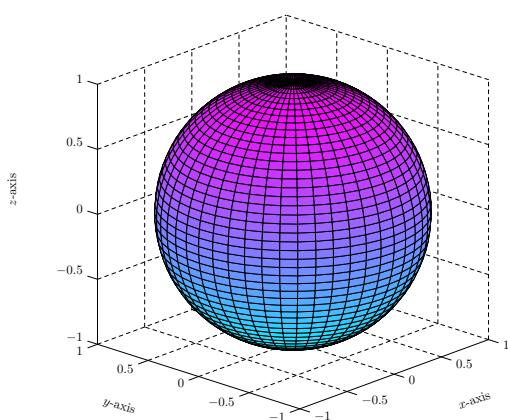
Le intersezioni con i piani $z = h$ sono tutte delle ellissi.



Ellissoide Immaginario

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = -1$$

Non ha punti reali.



Ellissoide Reale

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$$

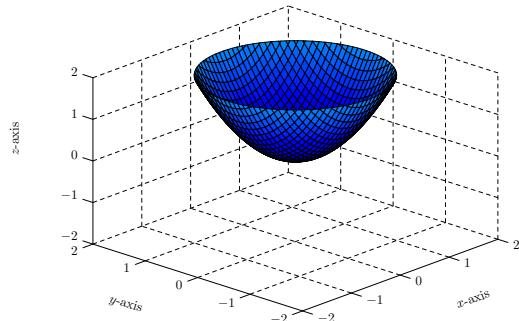
Si chiama **sferoide** quando
 $a = b \neq c \quad \vee \quad b = c \neq a \quad \vee \quad a = c \neq b$

Si chiama **sfera** quando $a = b = c$

Paraboloide Ellittico

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z}{c} = 0$$

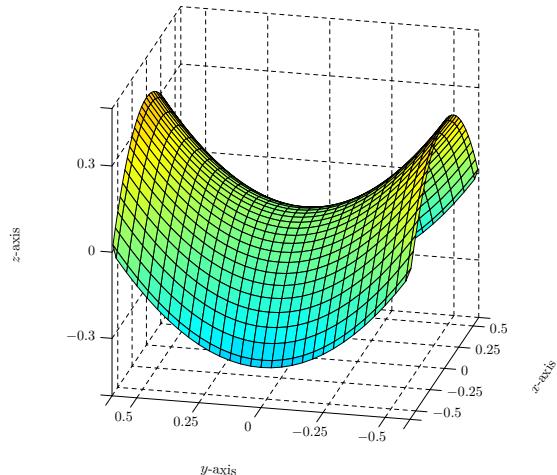
È una superficie le cui intersezioni con i piani $x = h$ e $y = h$ sono tutte delle parbole se $a > 0$.



Paraboloidi iperbolici

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} - \frac{z}{c} = 0$$

Prende anche il nome di paraboloidi “a sella”



Nei paraboloidi iperbolici esiste un punto che viene chiamato **punto di sella** in cui la prima derivata si annulla, ma la seconda derivata dipende dalla direzione in cui viene calcolata.

Figura 10.18: le patatine Pringles™ sono un perfetto esempio di paraboloidi iperbolici.

10.5.2 Altri criteri di classificazione

In alcuni testi le quadriche sono classificate in accordo con quello che chiamiamo **indice di specializzazione** che non è altro che $\text{rk}(B)$ e al variare del rango si classificano in diverse categorie:

- Si dicono quadriche non specializzate se $\text{rk}(B) = 4$
- Si dicono quadriche specializzate ma irriducibili se $\text{rk}(B) = 3$
- Si dicono quadriche specializzate e riducibili in due piani distinti se $\text{rk}(B) = 2$
- Si dicono quadriche specializzate e riducibili in due piani coincidenti se $\text{rk}(B) = 1$

10.5.3 Superfici rigate

Diciamo che una qualunque superficie S è **rigata** se per ogni punto appartenente alla superficie passa una retta che sia tutta contenuta nella superficie, cioè $\forall s \in S \exists r_s \in S$ I paraboloidi iperbolici, l'iperboloido iperbolico, tutti i cilindri reali e il cono sono delle superfici rigate.

10.5.4 Quadriche di rotazione

Quando una quadrica possiede due autovalori uguali, la quadrica si dice di rotazione, queste quadriche prendono tale nome poiché sono solidi di rotazione che possono essere ottenute ruotando opportunamente una curva, attorno ad una retta (*asse di rotazione*). Nel caso delle quadriche non specializzate a centro (ossia con $\det A, \det B \neq 0$), l'asse di rotazione coincide con la retta che passa per il centro e ha come direzione quella individuata dall'autospazio relativo all'autovalore semplice.

L'animazione qui sotto ci mostra come l'iperboloido iperbolico è una quadrica di rotazione che si ottiene ruotando la retta rispetto all'asse z (si consiglia di usare Foxit Phantom PDF oppure Acrobat Reader.)

Figura 10.19: Animazione

La sfera può essere ottenuta ruotando opportunamente una circonferenza.

Il paraboloide ellittico può essere ottenuto ruotando opportunamente una parabola.

Il cilindro ellittico può essere ottenuto ruotando opportunamente una retta perpendicolare al piano xy .

Dati curiosi

Pierre-Simon Laplace, celebre matematico francese che abbiamo incontrato nel capitolo 6, lasciò contributi a vari campi della matematica, dell'astronomia e della teoria della probabilità. Le sue nuove scoperte sono state di così grande rilevanza che dopo la sua morte è stato onorato più e più volte. Il suo nome è uno dei 72 nomi iscritti sulla torre Eiffel.

A riconoscimento delle sue famose equazioni delle maree fu onorato nel 1996 dai creatori del popolare videogioco *Pokémon* i cui nominarono una creatura marina Laplace. [Pokémon è un videogioco che ha come obiettivo catturare creature fantastiche con nomi particolari, i loro nomi risaltano le origini delle stesse, ad esempio *Charizard* viene da *char* (carbonizzare) e *lizard* (lucertola)]

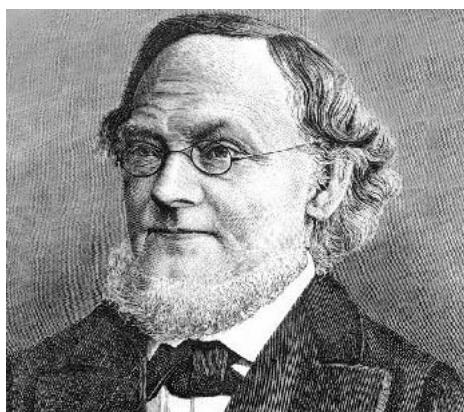


Figura 9.2:
Herman Grassmann ebbe undici figli.



Figura 9.1:
Pokémon #131: Lapras (traslitterazione giapponese di Laplace)

Il matematico Hermann Günther Grassmann (1809-1877) che abbiamo incontrato nel capitolo 1, fu conosciuto in vita principalmente per i suoi studi linguistici; non fu sino pochi primi anni della sua morte che i scienziati cominciarono a capire le sue quantiose scoperte nel ambito della matematica. Ad oggi è considerato il padre dell'algebra multilineare.

Carl Friedrich Gauss ha rivoluzionato la matematica e non a caso ha ricevuto il titolo di *princeps mathematicorum*, "signore dei matematici", addirittura è considerato da molti il più grande matematico della modernità.

Sin dai tre anni seppe come fare i calcoli a mente e ai dieci scopre la formula della somma dei primi n numeri naturali.

$$\frac{n(n+1)}{2}$$



Figura 9.3: Immagine più nota di Gauss

Il procedimento di Gram-Schmidt incontrato al capitolo 8 è chiamato così in onore del matematico danese Jørgen Gram (1850-1916) e del matematico tedesco Erhard Schmidt (1876-1959); esso però è stato utilizzato per prima da Laplace e Cauchy.

Il paraboloide iperbolico rivestì un'importanza fondamentale nei lavori dell'architetto spagnolo Felix Candela (1910-1997), che a causa della guerra civile decisi di trasferirsi in Messico ove rimase per ben trent'anni; lì ebbe l'opportunità di sperimentare nei suoi progetti con i paraboloidi iperbolici e alcune sue deformazioni che gli diedero fama internazionale.



Gli spagnoli non contenti del fatto che costruì quasi tutte le sue opere in Messico, si fecero progettare da lui l'iconico Oceanogràfic di Valencia (2002), che purtroppo non riuscì a vedere.



Figura 9.4: (1959)
Cappella aperta, Lomas de Cuernavaca,
Messico $y^2/100 - x^2/4.679 = z - 6$