Capitolo 1

I numeri complessi

1.1 I numeri complessi

Introduciamo un simbolo i, detto unità immaginaria definito dalla condizione

$$i^2 = i \cdot i = -1.$$

Il simbolo i soddisfa l'equazione

$$x^2 + 1 = 0.$$

Un numero complesso è un'espressione del tipo:

$$z = a + ib, \;\; a, b \in \mathbb{R}.$$

I numeri a e b si chiamano parte reale e parte immaginaria del numero complesso z.

Indicheremo con $\mathrm{Re}\, z$ e con $\mathrm{Im}\, z$ la parte reale e la parte immaginaria di un numero complesso z.

Sia $\mathbb{C}$ l'insieme dei numeri complessi.

L'insieme dei numeri reali $\mathbb{R}$ e un sottoinsieme di $\mathbb{C}$. Infatti i numeri reali si possono identificare con l'insieme

$$\{z = a + ib \in \mathbb{C} | \quad \mathrm{Im}(z) = b = 0\}.$$

I numeri complessi $z = a + ib$ tali che $\mathrm{Re}\, z = 0$ si dicono immaginari puri.

Due numeri complessi $z = a + ib$ e $w = c + id$ sono *uguali* se $a = c$ e $b = d$.

Definiamo la somma e la moltiplicazione di due numeri complessi $z = a + ib$ e $w = c + id$ con le solite regole del calcolo algebrico.

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

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

Dato un numero complesso $z = a + ib \neq 0$ esiste un numero complesso w, chiamato l'inverso di z tale che

$$z \cdot w = 1.$$

Denoteremo l'inverso di z con $\frac{1}{z}$.

Per trovare $\frac{1}{z}$ scriviamo $\frac{1}{z} = x + iy$. Allora

$$z \cdot \frac{1}{z} = 1 = (a + ib)(x + iy) = (ax - by) + i(bx + ay),$$

quindi

$$ax - by = 1, \ bx + ay = 0.$$

Abbiamo quindi un sistema lineare di due equazioni nelle due incognite x, y. Usando, il fatto che $z \neq 0$ e quindi $a^2 + b^2 \neq 0$ otteniamo:

$$x = \frac{a}{a^2 + b^2} \quad , \quad y = \frac{-b}{a^2 + b^2}.$$

Equivalentemente

$$\frac{1}{z} = \frac{1}{a + ib} = \frac{a - ib}{(a + ib)(a - ib)} = \frac{a - ib}{a^2 + b^2}.$$

Quindi il quoziente di due numeri complessi $w = c + id$ e $z = a + ib \neq 0$ è dato da:

$$\frac{w}{z} = w \cdot \frac{1}{z} = \frac{c + id}{a + ib} = \frac{(c + id)(a - ib)}{(a + ib)(a - ib)} = \frac{ac + bd}{a^2 + b^2} + i\frac{ad - bc}{a^2 + b^2}.$$

Osservazione 1 Le operazioni di somma e prodotto dei numeri complessi, ristrette ai numeri reali (cioè ai numeri complessi con parte immaginaria uguale a zero), coincidono con le operazioni di somma e prodotto sui numeri reali. Per questo motivo si dice che il campo dei numeri complessi $\mathbb{C}$ è un *estensione* del campo dei numeri reali $\mathbb{R}$.

Osservazione 2 Una differenza sostanziale tra il campo dei numeri reali e quello dei numeri complessi è quella dell'ordinamento.
Il campo dei numeri reali è un *campo ordinato* in quanto sussistono le seguenti proprietà:

1) se $x \leq y$ e $y \leq x$ allora $x = y$;

2) se $x \leq y$ e $y \leq z$ allora $x \leq z$;

3) $\forall x, y \in \mathbb{R}$ si ha $x \leq y$ oppure $y \leq x$;

4) se $x \leq y$, allora $x + z \leq y + z, \forall z \in \mathbb{R}$;

5) se $0 \leq x$ e $0 \leq y$, allora $0 \leq xy$.

Nel campo dei numeri complessi $\mathbb{C}$ non è possibile definire una struttura " $\leq''$ di campo ordinato.

Infatti dalle proprietà 4) e 5) seguirebbe che una relazione d'ordine $\leq$ soddisfa le seguenti proprietà:

- se un numero è positivo (negativo) il suo oppposto è negativo (positivo);
- il quadrato di un numero qualsiasi non è mai negativo.

D'altra parte $1^2 = 1$ e $i^2 = -1$. Quindi abbiamo due quadrati che sono l'uno l'opposto dell'altro. Nessuno dei due però può essere negativo (perchè sono quadrati) e questo è assurdo (perchè tra a e $-a$ uno dev'essere negativo se $a \neq 0$).

Il complesso coniugato di un numero complesso $z = a + ib$ è il numero, che si indica con $\bar{z}$, dato da $\bar{z} = a - ib$.

Un numero complesso è reale se e solo se coincide col suo coniugato.

Inoltre valgono le seguenti proprietà (esercizio!):

- $\mathrm{Re}(z) = \frac{z+\bar{z}}{2}$;
- $\mathrm{Im}(z) = \frac{z-\bar{z}}{2i}$;
- $z\bar{z} = (\mathrm{Re}(z))^2 + (\mathrm{Im}(z))^2$;

- $\frac{1}{z} = \frac{\bar{z}}{z\bar{z}}$;
- $\bar{\bar{z}} = z$; (l'operazione di coniugio è involutoria);
- $\overline{(z+w)} = \bar{z} + \bar{w}$;
- $\overline{(zw)} = \bar{z}\bar{w}, \forall z, w \in \mathbb{C}$;
- $\overline{\frac{1}{z}} = \frac{1}{\bar{z}}$.

Esempio 3 Vogliamo scrivere la forma algebrica, cioè la forma $z = a + ib$ del numero complesso

$$z = \frac{2+5i}{1-3i}.$$

Per fare questo moltiplichiamo numeratore e denominatore per $1+3i$ (il complesso coniugato di $1-3i$) e otteniamo:

$$z = \frac{2+5i}{1-3i} = \frac{(2+5i)(1+3i)}{(1-3i)(1+3i)} = \frac{1}{10}(-13+11i).$$

Quindi $a = \mathrm{Re}(z) = -\frac{13}{10}$ e $b = \mathrm{Im}(z) = \frac{11}{10}$.

Esempio 4 Scriviamo in forma algebrica il numero complesso

$$i(5+6i)^2 - \frac{4}{1-i}.$$

Si ha:

$$i(25+60i+36i^2)-\frac{4(1+i)}{(1-i)(1+i)} = -60-11i-2(1+i) = -62-13i$$

1.2 Rappresentazione geometrica e trigonometrica dei numeri complessi

Ricordiamo che i numeri reali si possono rappresentare come i punti di una retta una volta scelto un sistema di riferimento su di essa. Analogamente i numeri complessi si possono rappresentare con i punti di un piano, che prende il nome di *piano di Gauss*.

Fissato un sistema di coordinate ortogonali, associamo al numero complesso $z = a+ib$ il punto P del piano di coordinate (a, b).

In questo modo ai numeri reali vengono associati i punti dell'asse x, che viene detto *asse reale*, mentre gli immaginari puri sono rappresentati dai punti dell'asse y, detto *asse immaginario*.

L'unità immaginaria viene quindi rappresentata come la coppia $(0, 1)$.

Il punto che rappresenta il complesso coniugato $\bar{z}$ è il simmetrico rispetto all'asse x del punto che rappresenta z.

La distanza dall'origine O (che rappresenta lo zero 0 in $\mathbb{C}$) del punto P che rappresenta z si chiama <u>modulo</u> di z e si indica con $\rho = |z|$.

Per il Teorema di Pitagora

$$\rho = |z| = \sqrt{z\bar{z}} = \sqrt{a^2 + b^2}.$$

Supponiamo $z \neq 0$ (e quindi $P \neq O$). Denotiamo con

$$\theta = \operatorname{Arg}(z)$$

la misura in radianti di una qualsiasi rotazione che sovrappone l'asse delle x al segmento orientato OP, presa con il segno positivo se avviene in senso antiorario, negativo nel caso opposto, si chiama *argomento* di z.

Osserviamo che l'argomento di z è definito a meno di multipli di 2π. Inoltre se $z = 0$, l'argomento è indeterminato.

Per assegnare un ben determinato argomento basta assegnare un intervallo di ampiezza 2π nel quale far variare l'angolo θ. Noi fisseremo l'intervallo $[0, 2\pi)$.

Dato un numero complesso $z = a+ib$ sussistono le relazioni:

$$a = \mathrm{Re}(z) = \rho\cos\theta, \;\; b = \mathrm{Im}(z) = \rho\sin\theta. \tag{1.1}$$

Quindi

$$\boxed{z = \rho(\cos\theta + i\sin\theta)}, \tag{1.2}$$

che prende il nome di <u>rappresentazione trigonometrica</u> del numero complesso z.

Per trovare la forma trigonometrica (1.2) di un numero complesso $z = a + ib$ si usano le (1.1). Prima di tutto si calcola il modulo di z dato da

$$\boxed{\rho = \sqrt{a^2 + b^2}}.$$

Inoltre

$$\boxed{\cos\theta = \tfrac{a}{\rho} \quad \sin\theta = \tfrac{b}{\rho}}.$$

e quindi possiamo trovare univocamente θ restringendosi all'intervallo $[0, 2\pi)$.

Esempio 5 Scriviamo la forma trigonometrica del numero complesso $z = 1 - i$. Si ha

$$\rho = \sqrt{1^2 + 1^2} = \sqrt{2}$$

mentre

$$\cos\theta = \frac{1}{\rho} = \frac{1}{\sqrt{2}} \Rightarrow \theta = \frac{\pi}{4} \text{ oppure } \theta = \frac{7\pi}{4}$$

$$\sin\theta = \frac{-1}{\rho} = \frac{-1}{\sqrt{2}} \Rightarrow \theta = \frac{5\pi}{4} \text{ oppure } \theta = \frac{7\pi}{4}$$

Segue che $\theta = \frac{7\pi}{4}$ e la forma trigonometrica di z è data da

$$z = \sqrt{2}(\cos\frac{7\pi}{4} + i\sin\frac{7\pi}{4}).$$

Due numeri complessi

$$z_1 = \rho_1(\cos\theta_1 + i\sin\theta_1) \text{ e } z_2 = \rho_2(\cos\theta_2 + i\sin\theta_2)$$

sono uguali se e solo se

$$\rho_1 = \rho_2, \quad \theta_1 = \theta_2 + 2k\pi, k \in \mathbb{Z}$$

Dati

$$z_1 = \rho_1(\cos\theta_1 + i\sin\theta_1) \text{ e } z_2 = \rho_2(\cos\theta_2 + i\sin\theta_2)$$

usando le relazioni trigonometriche

$$\cos(\theta_1 + \theta_2) = \cos\theta_1 \cos\theta_2 - \sin\theta_1 \sin\theta_2$$

e

$$\sin(\theta_1 + \theta_2) = \sin\theta_1 \cos\theta_2 + \sin\theta_2 \cos\theta_1$$

si ottiene

$$\boxed{z_1 z_2 = \rho_1 \rho_2 [\cos(\theta_1 + \theta_2) + i \sin(\theta_1 + \theta_2)]}. \tag{1.3}$$

Proprietà *il prodotto di due numeri complessi ha come modulo il prodotto dei moduli e come argomento la somma degli argomenti*

In particolare per ogni intero non negativo n e per ogni numero complesso $z = \rho(\cos\theta + i \sin\theta)$ si ottiene la cosidetta *formula di De Moivre*:

$$\boxed{z^n = \rho^n [\cos(n\theta) + i \sin(n\theta)]} \tag{1.4}$$

Inoltre è immediato verificare che:

$$|\frac{1}{z}| = \frac{1}{|z|} = \frac{1}{\rho}, \quad \operatorname{Arg}(\frac{1}{z}) = -\operatorname{Arg}(z) = -\theta.$$

Equivalentemente

$$\frac{1}{z} = \frac{1}{\rho}(\cos(-\theta) + i \sin(-\theta)) = \frac{1}{\rho}(\cos\theta - i \sin\theta).$$

Dati

$$z_1 = \rho_1(\cos\theta_1 + i\sin\theta_1) \quad \text{e} \quad z_2 = \rho_2(\cos\theta_2 + i\sin\theta_2)$$

con $z_2 \neq 0$

$$\frac{z_1}{z_2} = \frac{\rho_1}{\rho_2}[\cos(\theta_1 - \theta_2) + i\sin(\theta_1 - \theta_2)]. \qquad (1.5)$$

Proprietà *il quoziente di due numeri complessi ha come modulo il quoziente dei moduli e come argomento la differenza degli argomenti.*

Esempio 6 Calcoliamo $z = (2-2i)^5$. Si ha:

$$z = (2-2i)^5 = 2^5(1-i)^5 = 32(1-i)^5.$$

Per l'esempio (5)

$$1 - i = \sqrt{2}(\cos\frac{7\pi}{4} + i\sin\frac{7\pi}{4})$$

Usando formula di formula di De Moivre (1.4) otteniamo:

$$(1-i)^5 = (\sqrt{2})^5(\cos\frac{7\pi}{4} + i\sin\frac{7\pi}{4})^5 = 4\sqrt{2}(\cos\frac{35\pi}{4} + i\sin\frac{35\pi}{4}).$$

Siccome $\frac{35\pi}{4} = \frac{3\pi}{4} + 8\pi)$, otteniamo

$$(1-i)^5 = 4\sqrt{2}(\cos\frac{3\pi}{4} + i\sin\frac{3\pi}{4}) = 4\sqrt{2}(-\frac{1}{\sqrt{2}} + i\frac{1}{\sqrt{2}}) = 4(-1+i).$$

Quindi $z = 128(-1+i)$.

Esempio 7 Cerchiamo i numeri complessi z che soddisfano l'equazione:

$$z^3 - |z|^2 = 0. \tag{1.6}$$

Primo metodo Scriviamo z in forma algebrica $z = x+iy$. L'equazione (1.6) diventa

$$z^3 = (x+iy)^3 = x^3 - 3xy^2 + i(3x^2y - y^3) = x^2 + y^2,$$

la quale è equivalente al sistema (non lineare!)

$$\begin{cases} x^3 - 3xy^2 & = & x^2 + y^2 \\ 3x^2y - y^3 & = 0 \end{cases} \tag{1.7}$$

nelle incognite reali x e y.

Dalla seconda equazione ricaviamo $y(3x^2 - y^2) = 0$ le cui soluzioni sono $y = 0$ e $y^2 = 3x^2$.

Sostituendo $y = 0$ nella prima equazione otteniamo

$$x^3 = x^2$$

le cui soluzioni sono $x = 0$ e $x = 1$. Quindi le coppie $(x = 0, y = 0)$ e $(x = 1, y = 0)$ sono soluzioni del sistema.

Sostituendo $y^2 = 3x^2$ nella prima equazione si ottiene

$$x^3 - 9x^3 = -8x^3 = x^2 + 3x^2 = 4x^2,$$

o, equivalentemente,

$$4x^2(2x+1)=0,$$

le cui soluzioni sono $x=0$ e $x=-\frac{1}{2}$. La soluzione $x=0$ sostituita in $y^2=3x^2$ ci da $y=0$ mentre $x=-\frac{1}{2}$ sostituita in $y^2=3x^2$ ci da $y=\frac{\sqrt{3}}{2}$ e $y=-\frac{\sqrt{3}}{2}$.
Abbiamo trovato quindi due nuove soluzioni $(x=-\frac{1}{2}, y=\frac{\sqrt{3}}{2})$ e $(x=-\frac{1}{2}, y=\frac{\sqrt{3}}{2})$

In definitiva abbiamo trovato le quattro coppie di soluzioni

$$(0,0),(1,0),(-\frac{1}{2},\frac{\sqrt{3}}{2}),(-\frac{1}{2},-\frac{\sqrt{3}}{2}),$$

e quindi le soluzioni dell'equazione (1.6) sono fornite dai quattro numeri complessi:

$$z_1=0, z_2=1, z_3=-\frac{1}{2}+i\frac{\sqrt{3}}{2}, z_4=-\frac{1}{2}-i\frac{\sqrt{3}}{2}.$$

Secondo metodo Scriviamo z in forma trigonometrica

$$z=\rho(\cos\theta+i\sin\theta).$$

Per la formula di De Moivre, l'equazione (1.6) è equivalente all'equazione

$$\rho^3(\cos(3\theta)+i\sin(3\theta))=\rho^2,$$

la quale è a sua volta equivalente al sistema

$$\begin{cases} \rho^3 = \rho^2 \\ \cos(3\theta) + i\sin(3\theta) = 1 \end{cases} \tag{1.8}$$

nelle incognite $\rho \geq 0$ e $\theta \in [0, 2\pi)$.

Le soluzioni della prima equazione sono $\rho = 0$ e $\rho = 1$. Mentre le soluzioni della seconda equazione sono $3\theta = 2k\pi, k \in \mathbb{Z}$ ossia le tre soluzioni $\theta = 0, \frac{2\pi}{3}, \frac{4\pi}{3}$. In questo modo ritroviamo le soluzioni già trovate in precedenza e cioè:

$$z_1 = 0, z_2 = \cos 0 + i\sin 0 = 1,$$
$$z_3 = \cos\frac{2\pi}{3} + i\sin\frac{2\pi}{3} = -\frac{1}{2} + i\frac{\sqrt{3}}{2}$$
$$z_4 = \cos\frac{4\pi}{3} + i\sin\frac{4\pi}{3} = -\frac{1}{2} - i\frac{\sqrt{3}}{2}.$$

Esempio 8 Calcoliamo il quoziente $\frac{z_1}{z_2}$ dove $z_1 = i$ e $z_2 = 1 - i$ usando le loro forme trigonometriche. Il modulo e l'argomento di $z_1 = i$ sono $\rho_1 = 1$ e $\theta_1 = \frac{\pi}{2}$, mentre il modulo e l'argomento di z_2 sono dati da $\rho_2 = \sqrt{2}$ e $\theta_2 = \frac{7\pi}{4}$. Qunidi per la formula (1.5)

$$\frac{z_1}{z_2} = \frac{1}{\sqrt{2}}\cos(\frac{\pi}{2} - \frac{7\pi}{4}) + i\sin(\frac{\pi}{2} - \frac{7\pi}{4}).$$

Osserviamo che

$$\frac{\pi}{2} - \frac{7\pi}{4} = -\frac{5\pi}{4} = 2\pi - \frac{5\pi}{4} = \frac{3\pi}{4},$$

da cui

$$\frac{z_1}{z_2} = \frac{1}{\sqrt{2}}(\cos\frac{3\pi}{4} + i\sin\frac{3\pi}{4}) = \frac{1}{\sqrt{2}}(-\frac{1}{\sqrt{2}} + i\frac{1}{\sqrt{2}}) = \frac{1}{2}(-1+i)$$

1.3 Radici di un numero complesso

Vogliamo estendere il concetto di radice n-esima valido per i numeri reali positivi a tutti i numeri complessi.

Definizione 9 *Dati un numero naturale $n \geq 1$ e un numero complesso w, diremo che il numero complesso z è una radice n-esima di w, e scriveremo $z = \sqrt[n]{w}$ se $z^n = w$.*

Teorema 10 *Sia $w \in \mathbb{C}$, $w \neq 0$, e n un intero ≥ 1. Esistono esattamente n radici n-esime complesse $z_0, z_1, \ldots, z_{n-1}$ di w. Posto $w = r(cos\,\varphi + i\,sin\,\varphi)$ e $z_k = \rho_k(cos\,\theta_k + i\,sin\,\theta_k)$ abbiamo*

$$\rho_k = \sqrt[n]{r}$$

$$\theta_k = \frac{\varphi + 2k\pi}{n}.$$

Equivalentemente

$$\boxed{z_k = \sqrt[n]{r}(cos(\tfrac{\varphi+2k\pi}{n}) + i\ sin(\tfrac{\varphi+2k\pi}{n})), \quad k = 0, 1, \ldots, n-1} \tag{1.9}$$

Dimostrazione: I numeri z_k sono evidentemente radici n-esime di w, come risulta applicando la formula di De Moivre (1.4). Mostriamo che non ce ne sono altre. Infatti affinchè un numero complesso $R(\cos\psi + i \sin\psi)$ sia radice n-esima di w, dovrebbe risultare:

$$R^n = r \ \text{ e } \ n\psi = \varphi + 2h\pi, h \in \mathbb{Z}$$

e cioè

$$R = \sqrt[n]{r} \ \text{ e } \ \psi = \frac{\varphi}{n} + \frac{2h\pi}{n}.$$

Attribuendo a n i valori $0, 1, \ldots, n-1$ troviamo appunto i numeri z_k. Dando a h un qualsiasi altro valore $\tilde{h}$ diverso dai precedenti, questo può scriversi nella forma $\tilde{h} = k+mn, m \in \mathbb{Z}$ ($m \in \mathbb{Z}$ è il quoziente e k il resto della divisione di $\tilde{h}$ per n) per cui sarebbe

$$\psi = \frac{\varphi}{n} + \frac{2k\pi}{n} + 2m\pi = \theta_k + 2m\pi$$

e ritroveremmo ancora gli stessi z_k precedenti. $\square$

Indichiamo con ϵ_k le radici n-esime del numero 1. Osserviamo che il modulo del numero 1 è uguale a 1 mentre l'argomento è $\theta = 0$.
Dalla formula (1.9) si ottiene allora:

$$\boxed{\epsilon_k = \cos\tfrac{2k\pi}{n} + i\sin\tfrac{2k\pi}{n}, \quad k = 0, 1, \ldots, n-1} \qquad (1.10)$$

Esempio 11 Calcoliamo le radici quarte di 1. In queto caso nella formula (1.10) $k = 0, 1, 2, 3$. Otteniamo quindi:

$$\begin{aligned}
\epsilon_0 &= \cos 0 + i\sin 0 = 1 \\
\epsilon_1 &= \cos\frac{2\pi}{4} + i\sin\frac{2\pi}{4} = \cos\frac{\pi}{2} + i\sin\frac{\pi}{2} = i \\
\epsilon_2 &= \cos\frac{4\pi}{4} + i\sin\frac{4\pi}{4} = \cos\pi + i\sin\pi = -1 \\
\epsilon_3 &= \cos\frac{6\pi}{4} + i\sin\frac{6\pi}{4} = \cos\frac{3\pi}{2} + i\sin\frac{3\pi}{2} = -i.
\end{aligned}$$

Proposizione 12 *Le radici n-esime di un qualunque numero complesso z si possono ottenere moltiplicando una di esse per le n radici n-esime del numero 1.*

Dimostrazione: se z_1 è una radice n-esima di z ed ϵ_k una qualsiasi radice n-esima di 1 si ha:

$$(z_1\epsilon_k)^n = z_1^n(\epsilon_k)^n = z_1^n = z$$

e quindi $z_1\epsilon_k$ è una radice n-esima di z; inoltre, al variare di $\epsilon_k, k = 0, 1, \ldots, n-1$, i numeri $z_1\epsilon_k$ sono tutti distinti. $\square$

Esempio 13 Calcoliamo le radici terze di -27. Sicuramente ci sono tre radici complesse. Una di queste è -3. Per la Proposizione 12 le radici terze di -27 si possono ottenere moltiplicando -3 per le radici terze di 1.

Dalla formula (1.10) queste ultime sono date da:

$$\begin{aligned}
\epsilon_0 &= \cos 0 + i \sin 0 = 1 \\
\epsilon_1 &= \cos \frac{2\pi}{3} + i \sin \frac{2\pi}{3} = -\frac{1}{2} + i\frac{\sqrt{3}}{2} \\
\epsilon_2 &= \cos \frac{4\pi}{3} + i \sin \frac{4\pi}{3} = -\frac{1}{2} - i\frac{\sqrt{3}}{2}
\end{aligned}$$

Quindi le radici terze di -27 sono date $z_0 = -3, z_1 = \frac{3}{2} - i\frac{3\sqrt{3}}{2}, z_2 = \frac{3}{2} + i\frac{3\sqrt{3}}{2}$.

Il Teorema 10 ci dice che un polinomio del tipo $z^n = z_0$ ammette n radici complesse. Vale un risultato più generale noto come il *teorema fondamentale dell'algebra* del quale riportiamo l'enunciato.

Abbiamo bisogno di una definizione

Definizione 14 *Se $P(z)$ è un polinomio in z di grado n e z_0 una sua radice, si dice che z_0 è di molteplicità k (k intero ≥ 1) se vale la formula*

$$P(z) = (z - z_0)^k Q(z),$$

dove Q è un polinomio tale che $Q(z_0) \neq 0$.

Esempio 15 *L'unità immaginaria i è radice di molteplicità due del polinomio $P(z) = z^3 - iz^2 + z - i = z^2(z-i) + (z-i) = (z-i)(z^2+1) = (z-i)^2(z+i)$*

Teorema 16 *(teorema fondamentale dell'algebra) Un'equazione polinomiale*

$$a_0 + a_1 z + \cdots + a_n z^n = 0, \quad a_n \neq 0$$

con coefficienti complessi qualsiasi ammette precisamente n radici in $\mathbb{C}$, se ognuna di esse viene contata con la sua molteplicità.

Capitolo 2

Vettori nel piano e nello spazio

2.1 Introduzione: I vettori in fisica

In questo capitolo andremo ad introdurre un concetto fondamentale in Matematica, Fisica e Informatica, quello di vettore. Vedremo alcune delle principali proprietá di tali oggetti e partendo da un'interpretazione fisica, astrarremo il concetto di vettore sino ad arrivare a parlare di vettore n-dimensionale, che in informatica é noto con il nome di array. Cerchiamo, dunque, di capire cosa é un vettore in fisica, quanto descritto nel seguito vale indifferentemente, a meno di diverso avviso, per vettori nel piano o nello spazio, pertanto non specificheremo dove ci troviamo. Iniziamo, quindi, da una caratterizzazione fisica dei vettori.

Consideriamo una particella P che si muove da un punto A ad un punto B. Indipendentemente dal percorso che essa compie, possiamo identificare lo spostamento da A in B con una freccia uscente da A, che prenderá il nome di punto di applicazione, e terminante in B, detto testa del vettore, tale freccia verrá chiamata vettore e lo indicheremo con $\underline{v}$. Un vettore $\underline{v}$ é completamente descritto da tre caratteristiche: la direzione, il verso e il modulo, o lunghezza del vettore. Spieghiamo cosa sono questi tre oggetti; la direzione di un vettore é individuata dal segmento congiungente A con B, é quindi la retta su cui il vettore giace. Il verso di $\underline{v}$ é invece individuato dal punto di partenza e dal punto d'arrivo della nostra particella, esso dunque, mi dice verso che parte della retta, sulla quale il vettore giace, la particella si sta spostando. Il modulo, o lunghezza del vettore, é dato dalla distanza del punto di partenza e di arrivo della particella. Dunque, due vettori, $\underline{v}$ e $\underline{v'}$, che hanno uguale modulo, uguale direzione e uguale verso sono lo stesso vettore, indipendentemente dal loro punto di applicazione. Ora, che abbiamo un idea di cosa sono i vettori, ci piacerebbe poter lavorare con essi, andiamo a definire quindi un metodo per sommarli, tale metodo é detto ***regola del parallelogramma***, e prende il nome dalla costruzione grafica che tale regola descrive. Supponiamo di avere due vettori, $\underline{u}$ e $\underline{v}$, diremo che un terzo vettore $\underline{w}$ é somma di $\underline{u}$ e $\underline{v}$, ovvero,

$$\underline{w} = \underline{u} + \underline{v},$$

e lo chiameremo vettore risultante, o semplicemente risultante, se lo possiamo ottenere con il seguente procedimento:

1) Usando una scala opportuna si tracci il vettore $\underline{u}$ a partire da un punto P d'applicazione;

2) Con la medesima scala si tracci $\underline{v}$ a partire da P, facendo attenzione che la direzione e il verso di $\underline{v}$ siano quelli corretti;

3) Si tracci ora a partire dalla testa di $\underline{v}$ la retta parallela alla direzione di $\underline{u}$ e si tracci a partire dalla testa di $\underline{u}$ la retta parallela alla direzione di $\underline{v}$. Le rette appena tracciate si incontreranno in un punto Q;

4) Si tracci ora il vettore da P a Q. Tale vettore é il vettore risultante, o la risultante $\underline{w}$.

Abbiamo quindi un modo per sommare, graficamente, due vettori e trovarne la risultante, é abbastanza immediato notare che se al posto del vettore $\underline{v}$, prendo il suo opposto, -$\underline{v}$, ovvero quel vettore che ha stesso modulo e direzione di $\underline{v}$ ma verso opposto, la regola del parallelogramma ci fornisce un metodo per sottrarre due vettori

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

Per l'operazione di somma appena descritta valgono le seguenti proprietá:

1. $\underline{u} + (\underline{v} + \underline{w}) = (\underline{u} + \underline{v}) + \underline{w}$ (Proprietá associativa);
2. $\underline{u} + \underline{v} = \underline{v} + \underline{u}$ (Proprietá commutativa);
3. $\underline{u} + \underline{0} = \underline{u}$ (Esistenza del neutro);
4. $\underline{u} - \underline{u} = \underline{0}$ (Esistenza dell'opposto).

É abbastanza semplice comprendere come questo metodo di lavorare con i vettori risulti essere poco pratico e, spesso, sconveniente, si provi ad immaginare di dover sommare dieci vettori nello spazio, riuscire a tracciare la regola del parallelogramma diventa un'impresa! Dobbiamo cosí cer-

care un modo per far le operazioni con i vettori in maniera rapida e comoda senza peró scordare il loro significato fisico/geometrico.
Prima di fare questo, peró, osserviamo ancora alcuni fatti che riguardano i vettori. Consideriamo un vettore $\underline{v}$, e immaginiamo di poterlo allungare indefinitamente lungo la retta su cui questo giace, questa operazione prende il nome di dilatazione del vettore $\underline{v}$, l'operazione inversa a questa, ovvero, quella di accorciamento del vettore, lungo la retta su cui giace, indefinitamente, senza, peró, poterlo mai rendere un punto, prende il nome di compressione del vettore $\underline{v}$. Nel dilatare o comprimere un vettore, quali, delle tre caratteristiche che individuano il vettore in maniera univoca (modulo, direzione e verso), viene modificata? Sicuramente non il verso ne la direzione. Visto che abbiamo specificamente richiesto che la dilatazione e la compressione avvengano sempre sulla retta su cui il vettore giace, ció che cambia é, il modulo, ovvero, modifichiamo solamente la lunghezza del vettore $\underline{v}$ che stiamo deformando. Matematicamente questa operazione puó essere compiuta moltiplicando il vettore per un numero reale, detto *scalare* e indicato con una lettera

dell'alfabeto greco, che sará maggiore di 1 se stiamo dilatando il vettore, oppure, compreso tra 0 e 1 se lo stiamo comprimendo $\underline{v}$. Osserviamo, ancora, che moltiplicando un vettore per uno scalare $\lambda < 0$ stiamo modificando il verso del vettore, ottenendo cosí, un vettore con verso opposto a quello di partenza, inoltre, se moltiplichiamo un vettore $\underline{v}$ per lo scalare $\lambda = 0$, ció che otteniamo é il vettore nullo $\underline{0}$, giá introdotto come elemento neutro della somma. Per indicare il vettore $\underline{u}$, ottenuto tramite l'operazione di moltiplicazione di un vettore $\underline{v}$ per uno scalare λ, scriveremo semplicemente:

$$\underline{u} = \lambda \underline{v}$$

Diamo alcune delle proprietá di cui il prodotto per uno scalare gode:

1. $\underline{v}\lambda = \lambda\underline{v}$ (Proprietá commutativa);
2. $(\lambda\mu)\underline{v} = \lambda(\mu\underline{v})$ (Proprietá associativa);
3. $1\underline{v} = \underline{v}$ (Esistenza dell'elemento neutro);
4. $(\lambda + \mu)\underline{v} = \lambda\underline{v} + \mu\underline{v}$ (Proprietá distributiva);
5. $\lambda(\underline{v} + \underline{u}) = \lambda\underline{v} + \lambda\underline{u}$ (Proprietá distributiva);

Abbiamo, quindi, un modo per allungare o accorciare i vettori, possiamo, allora, ottenere tutti i vettori, che giacciono su una data direzione, moltiplicando uno qualunque dei vettori, con quella direzione, per uno scalare. Poiché questa scelta é del tutto arbitraria, l'unico vincolo é, infatti, la direzione, possiamo prendere, tra tutti i vettori con una fissata direzione, il vettore che ha lunghezza unitaria, al quale daremo il nome di *versore* e indicheremo con $\underline{\iota}$. Quindi ogni vettore che giace in una determinata direzione puó essere scritto come:

$$\underline{v} = \lambda \underline{\iota}$$

Siamo ora pronti per cercare un metodo agevole che ci consenta di fare le operazioni con i vettori, senza coinvolgere la regola del parallelogramma. Di ció ci occuperemo nei successivi paragrafi, distinguendo il caso del piano da quello dello spazio e, in fine, tratteremo i vettori in $\mathbb{R}^n$.

2.2 I vettori nel piano

Indichiamo con $\mathcal{V}_2$ l'insieme dei vettori che giacciono sul piano e consideriamo in tale insieme due vettori, $\underline{e}_1$ ed $\underline{e}_2$, che non hanno la medesima direzione, una tale coppia é detta una *base del piano*, e scriveremo $\{\underline{e}_1 , \underline{e}_2\}$. Se, anziché considerare vettori qualunque, prendiamo dei versori ortogonali, ovvero due vettori di lunghezza unitaria le cui direzioni formano un angolo retto, diremo che la *base é ortonormale* e scriveremo $\mathcal{B} = \{i, j\}$. Distinguiamo, inoltre, il caso in cui per portare il versore i sul versore j é necessario effettuare una rotazione in senso antiorario o orario. Nel primo caso, rotazione antioraria, diremo che la base é positiva, nel secondo la diremo negativa.
Fissata una base siamo in grado di identificare $\mathcal{V}_2$ con $\mathbb{R}^2$, ovvero, a ogni vettore del piano possiamo associare una coppia di numeri reali che lo rappresentino rispetto alla base fissata. Fissata, dunque, $\mathcal{B}$ siamo in grado di scrivere ogni vettore $\underline{v}$ di $\mathcal{V}_2$ come:

$$\underline{v} = v_1 i + v_2 j$$

dove v_1 e v_2 sono due scalari opportuni, detti *componenti del vettore nella base* $\mathcal{B}$ e i e j sono i due versori ortonormali che formano la base $\mathcal{B}$. Abbiamo cosí trovato il modo di rappresentare un vettore senza doverlo necessariamente disegnare. Basta infatti conoscere $\mathcal{B}$ e le componenti del vettore rispetto a tale base, per averlo determinato in maniera univoca, in quella base.
Per comoditá, da ora in poi, ometteremo la sottolineatura nell'indicare i vettori, scriveremo, cioé, semplicemente v in luogo di $\underline{v}$.

Sino a ora, si é parlato di lunghezze ed angoli basandoci sulla nozione intuitiva che possediamo di questi due oggetti, abbiamo ora necessitá di darne una definizione rigorosa. Per fare ció introduciamo una nuova operazione, nota con il nome di *prodotto scalare*, da <u>non</u> confondersi con il prodotto *per* uno scalare e la indicheremo con $u \cdot v$, questa operazione prende una coppia di vettori del piano $\mathcal{V}_2$ e restituisce uno scalare ed é definita come segue:

$$u \cdot v = u_1 v_1 + u_2 v_2,\ u = (u_1, u_2), v = (v_1, v_2)$$

Diamo ora le proprietá di questa operazione:

1. $(\lambda u + \mu v) \cdot w = \lambda(u \cdot w) + \mu(v \cdot w)$
   e $u \cdot (\eta v + \xi w) = \eta(u \cdot v) + \xi(u \cdot w)$ (Bilinearitá);

2. $u \cdot v = v \cdot u$ (Simmetria);

3. $u \cdot u \geq 0$ e $u \cdot u = 0 \iff u = 0$ (Definito positivo);

A questo punto siamo in grado di definire la lunghezza di un vettore, o norma del vettore, e di dare un modo per calcolare l'angolo tra due vettori. Iniziamo dalla norma, diciamo *norma* o lunghezza di un vettore u, e lo indicheremo con $\|u\|$, il numero reale dato da:

$$\|u\| = \sqrt{u \cdot u} = \sqrt{u_1^2 + u_2^2}$$

Adesso che sappiamo misurare le lunghezze ci piacerebbe poter misurare anche gli angoli. Diamo per ora la seguente definizione di angolo tra due vettori u e v, nel seguito ne daremo una giustificazione, mostrando che, tale definizione, corrisponde all'idea comune di angolo che tutti possediamo:

$$\cos \hat{uv} = \frac{u \cdot v}{\|u\|\|v\|}, \quad 0 \leq \hat{uv} \leq \pi$$

Da quest'ultima discende che una coppia di vettori é ortogonale se e solamente se il loro prodotto scalare é nullo, ovvero:

$$u \perp v \Longleftrightarrow u \cdot v = 0$$

Abbiamo, dunque, un modo per capire quando un vettore é ortogonale ad un altro. I vettori di una base ortonormale $\mathcal{B}$ dovranno quindi rispettare le condizioni:

$$\|i\| = 1, \|j\| = 1$$

e

$$i \cdot j = 0$$

una coppia di vettori che soddisfano tali condizioni risulta essere:

$$i = (1, 0) \text{ e } j = (0, 1).$$

Diremo la base $\mathcal{B}$, costituita dai versori i e j su detti, *base canonica del piano.*

Definiamo le somme tra vettori e le dilatazioni senza ricorrere alla regola del parallelogramma, o a un metodo grafico.

La somma tra due vettori u e v, si tradurrá, in una base $\mathcal{B} = \{i, j\}$, nella somma delle componenti di u e v, in formule:

Se $u = (u_1, u_2)$ *e* $v = (v_1, v_2)$ *il vettore somma* $w = u + v$ *sará dato da:*

$$w = u + v = (u_1 + v_1, u_2 + v_2).$$

Il prodotto per uno scalare $\lambda \in \mathbb{R}$ *sará, invece, dato da:*

$$\lambda u = (\lambda u_1, \lambda u_2)$$

Essendo, le componenti di un vettore, numeri reali, le operazioni di somma e prodotto per uno scalare, godranno delle usuali proprietá della somma e del prodotto tra numeri reali.

Introduciamo ora il concetto di riferimento ortonormale. Diciamo *riferimento ortonormale* e lo indicheremo con $\mathcal{R}(O, \mathcal{B})$, la coppia costituita dal punto O detto origine del sistema di riferimento, e da una base ortonormale $\mathcal{B}$ di versori del piano. Le rette individuate dai versori i, j, della base canonica del piano, costituiscono rispettivamente l'asse x, o asse delle ascisse, e l'asse y, o asse delle ordinate, di un sistema di riferimento cartesiano del piano. Diremo inoltre *coordinate cartesiane* di un punto del piano, rispetto al riferimento cartesiano, le componenti del vettore $P - O$, cioé:

$$P(x, y) \Longleftrightarrow P - O = xi + yj$$

Abbiamo, dunque, un modo di identificare i punti del piano con i vettori applicati nell'origine.
Siamo ora in grado di dare una spiegazione della definizione di angolo data in precedenza. Consideriamo due vettori u e v e cerchiamo di capire come poter individuare l'angolo $\theta = \hat{uv}$ tra essi, questa misura non Ë immediata, siamo infatti in grado di misurare solo angoli a partire da un asse fissato, usualmente dall'asse x, e non tra due direzioni qualunque, dovremo, quindi, ricorrere alla trigonometria e alle formule del coseno e del seno della somma di due angoli, ricordiamole:

$$\cos(\psi + \theta) = \cos\psi\cos\theta - \sin\psi\sin\theta;$$

$$\sin(\psi + \theta) = \sin\psi\cos\theta + \cos\theta\sin\psi;$$

con questi strumenti iniziamo la misura dell'angolo tra u e v. Fissato un sistema di riferimento cartesiano ortonormale, $\mathcal{R} = (O, \mathcal{B})$, sia ψ l'angolo formato tra il vettore $u = P - O = (x_1, y_1)$ e l'asse delle x, sia invece θ l'angolo tra il vettore u e il vettore $v = Q - O = (x_2, y_2)$, con tale configurazione, l'angolo, che il vettore v forma con l'asse x, sará dato da

$\psi+\theta$, per definizione di coseno e seno di un angolo otteniamo:

$$\begin{cases} \cos\psi = \dfrac{x_1}{\|u\|} \\ \sin\psi = \dfrac{y_1}{\|u\|} \end{cases}$$

$$\begin{cases} \cos(\psi+\theta) = \dfrac{x_2}{\|v\|} \\ \sin(\psi+\theta) = \dfrac{y_2}{\|v\|} \end{cases}$$

Da quest'ultimo sistema, applicando le formule trigonometriche ricordate in precedenza, otteniamo:

$$\begin{cases} \cos\psi\cos\theta - \sin\psi\sin\theta = \dfrac{x_2}{\|v\|} \\ \sin\psi\cos\theta + \cos\psi\sin\theta = \dfrac{y_2}{\|v\|} \end{cases}$$

Sostituendo a $\cos\psi$ e $\sin\psi$ i valori giá trovati:

$$\begin{cases} \dfrac{x_1}{\|u\|}\cos\theta - \dfrac{y_1}{\|u\|}\sin\theta = \dfrac{x_2}{\|v\|} \\ \dfrac{y_1}{\|u\|}\cos\theta + \dfrac{x_1}{\|u\|}\sin\theta = \dfrac{y_2}{\|v\|} \end{cases}$$

e risolvendo si ottiene:

$$\cos\theta = \frac{x_1x_2 + y_1y_2}{\|u\|\|v\|}$$

Abbiamo, quindi, trovato la formula data in precedenza per la determinazione dell'angolo $\theta = \hat{uv}$:

$$\cos\theta = \cos(\hat{uv}) = \frac{u \cdot v}{\|u\|\|v\|}$$

Dato, ora, un vettore u, non piú applicato nell'origine $\mathcal{O}$ del sistema del riferimento, ma applicato a un punto P_1, e con la testa in P_2, le componenti di $u = \overrightarrow{P_1P_2} = P_2 - P_1$ saranno date dalla differenza delle coordinate del punto di testa P_2 dal punto di applicazione P_1, ovvero:

$$u = (P_2 - O) - (P_1 - O) = x_2 i + y_2 j - (x_1 i + y_1 j) =$$

$$= (x_2 - x_1)i + (y_2 - y_1)j$$

Possiamo in tal modo descrivere vettori applicati in qualunque punto del piano ed eseguire, su essi, tutte le operazioni descritte sino a ora.
Consideriamo ora il segmento di estremi $P_1(x_1, y_1)$ e $P_2(x_2, y_2)$, il punto medio di tale segmento é quel punto che dista in egual misura dagli estremi del segmento, ovvero, in formule:

$$M - P_1 = P_2 - M$$

Da tale equazione ricaviamo in coordinate:

$$(x_M, y_M) = \frac{(x_2, y_2) + (x_1, y_1)}{2}$$

ovvero:

$$(x_M, y_M) = \left(\frac{x_1 + x_2}{2}, \frac{y_1 + y_2}{2}\right)$$

Definiamo, infine, la distanza tra due punti del piano $P_1(x_1, y_1)$ e $P_2(x_2, y_2)$ come la lunghezza del vettore u, con punto di applicazione in $P_1(x_1, y_1)$ e testa in $P_2(x_2, y_2)$, otteniamo in tal modo:

$$\overline{P_1P_2} = \|P_1P_2\| = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

Facciamo un ultima osservazione su questa formula, supponiamo che $x_1 < x_2$ e $y_1 < y_2$ e poniamo $a = \overline{P_1P_2}$, $b = (x_2 - x_1)$ e $c = (y_2 - y_1)$, otteniamo:

$$a = \sqrt{b^2 + c^2}$$

interpretando a come l'ipotenusa del triangolo formato dai punti $Q(x_2, y_1), P_1(x_1, y_1), P(x_2, y_2)$, e b e c come i cateti del suddetto triangolo si vede subito che la formula scritta per la distanza tra due punti altro non Ë che il teorema di pitagora.

2.3 I vettori nello spazio

Abbiamo visto nel paragrafo precedente come descrivere in termini algebrici, ovvero sfruttando le componenti rispetto a un riferimento fissato, i vettori del piano. Ci piacerebbe ora poter fare lo stesso per i vettori dello spazio, indicheremo con $\mathcal{V}_3$ l'insieme di tali vettori. Per fare quanto fatto in $\mathcal{V}_2$ dobbiamo estendere il concetto di base allo spazio. Nel piano abbiamo definito una base come una coppia di vettori non proporzionali, ovvero, non aventi la medesima direzione. Nello spazio diremo base *una terna di vettori non complanari*, ovvero, una terna di vettori che non giacciono sullo stesso piano. Diremo, dunque, che $\mathcal{B} = \{i, j, k\}$ é una base ortonormale dello spazio, se é una terna di vettori di lunghezza unitaria e tra loro ortogonali; preciseremo, ancora, che la base é positiva se la piú piccola rotazione con cui posso portare il versore i sul versore j é antioraria se vista dal semispazio individuato da k, se la piú piccola rotazione é, invece, oraria, la base sará detta negativa. Fatto ció possiamo identificare $\mathcal{V}_3$ con $\mathbb{R}^3$ associando a ogni vettore $v \in \mathcal{V}_3$ la terna delle sue coordinate nella base $\mathcal{B}$, $(v_1, v_2, v_3) \in \mathbb{R}^3$, dunque:

$$v = v_1 i + v_2 j + v_3 k.$$

Possiamo ora estendere allo spazio quanto fatto per il piano, definendo il prodotto scalare tra due vettori

$$u = u_1 i + u_2 j + u_3 k$$

e

$$v = v_1 i + v_2 j + v_3 k$$

come:

$$u \cdot v = u_1 v_1 + u_2 v_2 + u_3 v_3,$$

la norma, o lunghezza di un vettore, u:

$$\|u\| = \sqrt{u \cdot\ u} = \sqrt{u_1^2 + u_2^2 + u_3^2},$$

e l'angolo tra vettori:

$$\cos \hat{uv} = \frac{u \cdot v}{\|u\| \|v\|}$$

Diremo, ancora, che due vettori dello spazio sono tra loro ortogonali se il loro prodotto scalare é nullo, ovvero:

$$u \perp v \Longleftrightarrow u \cdot v = 0$$

Quindi i vettori di una base ortonormale devono rispettare le seguenti condizioni:

$$\|i\| = 1; \|j\| = 1; \|k\| = 1$$

$$i \cdot j = 0; j \cdot k = 0; k \cdot i = 0$$

Una terna che soddisfa tali condizioni é quella formata dai vettori di componenti:

$$i = (1, 0, 0); j = (0, 1, 0); k = (0, 0, 1),$$

la base formata da tale terna é detta *base canonica dello spazio*.

Data una base dello spazio possiamo, dunque, effettuare l'operazione di somma di vettori e prodotto per uno scalare sfruttando le componenti del vettore, rispetto alla base fissata, nel modo seguente:

$$u + v = (u_1 i + u_2 j + u_3 k) + (v_1 i + v_2 j + v_3 k) =$$

$$(u_1 + v_1)i + (u_2 + v_2)j + (u_3 + v_3)k$$

e

$$\lambda v = \lambda(v_1 i + v_2 j + v_3 k) = (\lambda v_1 i + \lambda v_2 j + \lambda v_3 k)$$

Per tali operazioni sussistono le proprietá della somma e del prodotto tra numeri reali.

Il prodotto vettoriale

Introduciamo una nuova operazione detta *prodotto vettoriale*, che prende due vettori u e v dello spazio e restituisce un'altro vettore w. In formule:

$$w = u \wedge v$$

questa operazione é definita dalle seguenti proprietá:

1. Se u e v sono paralleli $u \wedge v = 0$;
2. Se u e v non sono paralleli:
   - $u \wedge v$ é ortogonale sia ad u che a v, é quindi non complanare con essi;
   - il verso di $u \wedge v$ é tale che la terna $(u, v, u \wedge v)$ é positiva;
   - il modulo é dato da:

$$\|u \wedge v\| = \|u\| \|v\| \sin \hat{uv}$$

Per i vettori della base ortonormale $\mathcal{B} = \{i, j, k\}$ risulta:

$$i \wedge i = j \wedge j = k \wedge k = 0;$$
$$i \wedge j = -j \wedge i = k;$$
$$i \wedge k = -k \wedge i = j;$$
$$j \wedge k = -k \wedge j = i.$$

Il vettore $w = u \wedge v$ ha componenti date dallo sviluppo del determinante:

$$w = u \wedge v = \begin{vmatrix} i & j & k \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix} =$$

$$= (u_2 v_3 - u_3 v_2)i - (u_1 v_3 - u_3 v_1)j + (u_1 v_2 - u_2 v_1)k$$

Osserviamo che il prodotto vettoriale, é un'operazione che prende due vettori dello spazio e ne restituisce un'altro ortogonale ai vettori di partenza.
Il prodotto misto di vettori, ovvero l'operazione $u \wedge v \cdot w$, in cui si da precedenza al prodotto vettoriale e poi al prodotto scalare, é dato da:

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

$$= (u_2v_3 - u_3v_2)w_1 - (u_1v_3 - u_3v_1)w_2 + (u_1v_2 - u_2v_1)w_3$$

Osserviamo che $u \wedge v \cdot w = u \cdot v \wedge w$ (esercizio).

Possiamo dare un'interpretazione geometrica al modulo delle ultime due operazioni introdotte, il modulo del prodotto vettoriale risulta infatti essere l'area del parallelogramma individuato dai vettori u e v posizionati in maniera da avere lo sesso punto di applicazione. Il modulo del prodotto misto, invece, si interpreta come il volume del parallelepipedo individuato dai tre vettori u, v, w posizionati in modo da avere lo stesso punto di applicazione, osserviamo che se u, v e w sono complanari tale modulo risulta essere nullo, in quanto $u \wedge v$ risulta essere ortogonale al piano contenente u , v e w.

Riferimento ortonormale nello spazio

Come per il piano introduciamo ora il concetto di riferimento ortonormale, in modo tale da poter assegnare ai punti dello spazio una terna di numeri reali, detta *le coordinate del*

punto nel riferimento assegnato.
Diciamo quindi riferimento ortonormale, e scriveremo $\mathcal{R}(O, \mathcal{B})$, la coppia costituita da un punto O dello spazio, detto origine del sistema di riferimento, e da una base $\mathcal{B}$ ortonormale dello spazio. Diremo che il riferimento é positivo o negativo se tale é la base $\mathcal{B}$ ortonormale dello spazio. Come nel caso del piano, i versori i,j e k applicati in O individuano tre assi, detti *assi coordinati*, denominati rispettivamente, asse x, asse y e asse z (o delle quote). Presi a coppie questi assi individuano tre piani, detti *piani coordinati* e verranno indicati con $[xy]$, $[yz]$ e $[xz]$. Possiamo ora assegnare a un punto P dello spazio delle coordinate rispetto al riferimento scelto, se il riferimento é ortonormale diremo tali coordinate *coordinate cartesiane del punto* P e saranno individuate dalle coordinate del vettore applicato in O e testa in P, $P - O$, rispetto alla base $\mathcal{B} = (i, j, k)$ scelta nel riferimento, in formule:

$$P(x, y, z) \Longleftrightarrow P - O = xi + yj + zj$$

Un vettore u non applicato nell'origine del sistema di riferimento, ma applicato in un punto P_1 e con testa in un punto P_2, avrá coordinate, rispetto alla base fissata nel sistema di riferimento:

$$P_2 - P_1 = (x_2 - x_1)i + (y_2 - y_1)j + (z_2 - z_1)k$$

Il punto medio di un segmento di estremi i punti $P_1(x_1, y_1, z_1)$ e $P_2(x_2, y_2, z_2)$, analogamente a quanto fatto sul piano, sará l'unico punto tale che:

$$M - P_1 = P_2 - M$$

Da cui in coordinate:

$$(x_M, y_M, z_M) = \frac{(x_2, y_2, z_2) + (x_1, y_1, z_1)}{2}$$

ovvero:

$$(x_M, y_M, z_M) = \left(\frac{x_2 + x_1}{2}, \frac{y_2 + y_1}{2}, \frac{z_2 + z_1}{2}\right)$$

La distanza tra due punti dello spazio $P_1(x_1, y_1, z_1)$ e $P_2(x_2, y_2, z_2)$, denotata con $\overline{P_1P_2}$, é data dalla lunghezza del vettore applicato in P_1 e con testa in P_2, in formule:

$$\overline{P_1P_2} = \|P_2 - P_1\| = \sqrt{(x_2 - x_1)^2 + (z_2 - z_1)^2 + (z_2 - z_1)^2}$$

Risolviamo ora il problema di calcolare le aree e i volumi di triangoli, parallelepipedi e tetraedri, una volta dato un numero sufficiente di punti che li possano individuate.

Iniziamo dal triangolo, dati tre punti dello spazio, $P_1(x_1, y_1, z_1)$, $P_2(x_2, y_2, z_2)$ e $P_3(x_3, y_3, z_3)$, possiamo trovare l'area A del triangolo di cui essi sono i vertici semplicemente calcolando il modulo del prodotto scalare $(P_2 - P_1) \wedge (P_3 - P_1)$, e dividendo il risultato per due, cioé:

$$A = \frac{1}{2}\|(P_2 - P_1) \wedge (P_3 - P_1)\|$$

Dati quattro punti $P_1(x_1, y_1, z_1)$, $P_2(x_2, y_2, z_2)$, $P_3(x_3, y_3, z_3)$ e $P_4(x_4, y_4, z_4)$, invece, possiamo trovare il volume V del parallelepipedo, individuato dai vettori $P_2 - P_1$, $P_3 - P_1$ e $P_2 - P_1$, calcolando il modulo del prodotto misto $(P_2 - P_1) \wedge (P_3 - P_1) \cdot (P_4 - P_1)$, ovvero:

$$V = \|(P_2 - P_1) \wedge (P_3 - P_1) \cdot (P_4 - P_1)\|$$

Infine, il volume del tetraedro di vertici $P_1(x_1, y_1, z_1)$, $P_2(x_2, y_2, z_2)$, $P_3(x_3, y_3, z_3)$ e $P_4(x_4, y_4, z_4)$ é dato da un sesto del volume del parallelepipedo individuato dai vettori $P_2 - P_1$, $P_3 - P_1$ e $P_2 - P_1$, ovvero:

$$V = \frac{1}{6}\|(P_2 - P_1) \wedge (P_3 - P_1) \cdot (P_4 - P_1)\|$$

2.4 Lo spazio vettoriale $\mathbb{R}^n$

Sia $\mathbb{R}^n$ l'insieme costituito dalle n-uple ordinate di numeri reali. Un elemento $v \in \mathbb{R}^n$ sará, quindi, della forma:

$$v = (v_1, v_2, \ldots, v_n), \; v_j \in \mathbb{R}^n, \; j = 1, 2, \ldots n.$$

Chiamiamo gli elementi di $\mathbb{R}^n$ *vettori* e analogamente a ciá che abbiamo fatto nel piano e nello spazio definiamo la *somma* tra $u = (u_1, u_2, \ldots, u_n)$ e $v = (v_1, v_2, \ldots, v_n)$ in $\mathbb{R}^n$ come il vettore

$$u + v = (u_1 + v_1, u_2 + v_2, \ldots, u_n + v_n)$$

e la *moltiplicazione* di $\lambda \in \mathbb{R}$ per il vettore $u \in \mathbb{R}^n$ come

$$\lambda u = (\lambda u_1, \lambda u_2, \ldots, \lambda u_n).$$

Diremo inoltre che due vettori u e v in $\mathbb{R}^n$ sono *proporzionali o paralleli* se: $u = \lambda v$.

Valgono le seguenti proprietà sono di facile verifica.

1. $u + (v + w) = (u + v) + w, \forall u, v, w \in \mathbb{R}^n$
   (Proprietá associativa per la somma);

2. $u + v = v + u, \forall u, v \in \mathbb{R}^n$
   (Proprietá commutativa per la somma);

3. Esiste un elemento 0 tale che $u + 0 = u, \forall u \in \mathbb{R}^n$
(Esistenza del neutro per la somma);

4. Esiste un elemento $-u$ tale che $u + (-u) = 0, \forall u \in \mathbb{R}^n$
(Esistenza dell'opposto per la somma);

5. $v\lambda = \lambda v, \forall v \in \mathbb{R}^n, \forall \lambda \in \mathbb{R}$
(Proprietá commutativa del prodotto);

6. $(\lambda\mu)v = \lambda(\mu v), \forall v \in \mathbb{R}^n, \forall \lambda, \mu \in \mathbb{R}$
(Proprietá associativa del prodotto);

7. $1v = v, \forall v \in \mathbb{R}^n$
(Esistenza dell'elemento neutro per il prodotto);

8. $(\lambda + \mu)v = \lambda v + \mu v, \forall v \in \mathbb{R}^n, \forall \lambda, \mu \in \mathbb{R}$
(Proprietá distributiva per i vettori);

9. $\lambda(v + u) = \lambda v + \lambda u, \forall u, v \in \mathbb{R}^n, \forall \lambda \in \mathbb{R}$
(Proprietá distributiva per gli scalari);

Occupiamoci di risolvere il problema di misurare lunghezze e angoli in $\mathbb{R}^n$. Ora infatti non abbiamo idea di cosa possa essere una lunghezza o un angolo, viene meno l'intuizione e il riscontro di tale intuizione con la realtá, poiché non siamo in grado ne vedere ne di immaginare uno spazio con piú di tre dimensioni, ció non significa che questi non esistano. Per misurare in uno spazio con piú di tre dimensioni dobbiamo, quindi, definire cosa significa lunghezza di un vettore e cosa significa angolo tra due vettori.Definiamo prima il prodotto scalare in $\mathbb{R}^n$ come:

$$u \cdot v = u_1 v_1 + u_2 v_2 + \ldots + u_n v_n$$

Sussistono ancora tutte le proprietá che il prodotto scalare possedeva nel piano e nello spazio.
Siamo adesso in grado di definire la lunghezza o norma di un vettore in $\mathbb{R}^n$. Definiamo *norma o lunghezza* di un vettore $v = (v_1, v_2, \ldots, v_n) \in \mathbb{R}^n$, e la indicheremo con $\|v\|$, il numero reale dato da:

$$\|v\| = \sqrt{v_1^2 + \ldots + v_n^2}$$

Vale il seguente teorema:

Teorema 17 *Valgono le seguenti:*

1. $\|v\| = 0$ *e solo se* $v = 0$*;*
2. $\|v\| > 0$ *per ogni* $v \neq 0$*;*
3. $\|\lambda v\| = |\lambda| \|v\|$, $\forall \lambda \in \mathbb{R}^n$ *e* $\forall v \in \mathbb{R}^n$*;*

4. $\|u+v\|^2 = \|u\|^2 + 2u \cdot v + \|v\|^2$ *per ogni* $u, v \in \mathbb{R}^n$*;*

5. *Per ogni* $u, v \in \mathbb{R}^n$ *si ha :*

$$|u \cdot v| \leq \|u\|\|v\|$$

(Disuguaglianza di Cauchy-Schwartz). Inoltre vale l'uguaglianza se e solo se u *e* v *sono paralleli;*

6. *Per ogni* $u, v \in \mathbb{R}^n$ *si ha :*

$$|\|u\| - \|v\|| \leq \|u+v\| \leq \|u\| + \|v\|$$

(Disuguaglianza triangolare);

7. $u \cdot v = \frac{1}{4}(\|u+v\|^2 - \|u-v\|^2)$

Dimostrazione: Le affermazioni 1. e 2. discendono direttamente dalla definizione di prodotto scalare definito positivo.

3. $\|\lambda v\| = \sqrt{\lambda v \cdot \lambda v} = \sqrt{|\lambda|^2 v \cdot v} = |\lambda|\|v\|$;

4.

$$\|u+v\| = (u+v)\cdot(u+v) = u\cdot u + u\cdot v + v\cdot u + v\cdot v =$$

$$\|u\|^2 + 2u \cdot v + \|v\|^2$$

5. Se $u = 0$ o $v = 0$ non c'é nulla da dimostrare; supponiamoli dunque entrambe non nulli. Poiché il prodotto scalare é definito positivo, si ha:

$$0 \leq \|au + bv\|^2 = a^2\|u\|^2 + 2abu \cdot v + b^2\|v\|^2$$

per ogni coppia di scalari a,b. In particolare posti $a = \|v\|^2$ e $b = -u \cdot v$ otteniamo:

$$0 \leq \|v\|^2(\|u\|^2\|v\|^2 - (u \cdot v)^2)$$

dividendo ora per $\|v\|^2$ si ottiene l'asserto. L'uguaglianza vale se e solo se u e v sono paralleli.

6. Dalla 5. si ha $|u \cdot v| \leq \|u\|\|v\|$, da cui:
$(\|u\| - \|v\|)^2 = \|u\|^2 + \|v\|^2 - 2\|u\|\|v\| \leq$
$\leq \|u\|^2 + \|v\|^2 + 2u \cdot v = \|u + v\|^2 \leq$
$\leq \|u\|^2 + \|v\|^2 + 2\|u\|\|v\| = (\|u\| + \|v\|)^2 \qquad \square$;

7. lasciato per esercizio.

Definiamo, ora, l'angolo tra due vettori u e v in $\mathbb{R}^n$, come quel numero $\theta \in [0, \pi]$ tale che:

$$\cos(\theta) = \frac{u \cdot v}{\|u\|\|v\|}$$

Grazie alla disuguaglianza di Cauchy-Schwartz questo é ben definito, ovvero:

$$|\cos(\theta)| \leq 1$$

Inoltre, essendo il coseno una funzione iniettiva tra $[0, \pi]$ la definizione data individua in modo univoco l'angolo tra u e v.
Abbiamo cosí un modo per calcolare lunghezze e misurare angoli anche in spazi di dimensione maggiore di tre.

Combinazioni lineari

Sul piano e nello spazio eravamo in grado di scrivere i vettori a partire da dei vettori fissati, ci piacerebbe poter fare qualcosa di simile anche in $\mathbb{R}^n$, per far ció dobbiamo introdurre dei concetti nuovi, che sono, quello di combinazione lineare, spazio generato e indipendenza lineare. Una *combinazione lineare* di $v_1, v_2, \ldots v_k$ vettori in $\mathbb{R}^n$ è un'espressione del tipo;

$$a_1 v_1 + a_2 v_2 + \ldots + a_k v_k$$

con $a_1, a_2, \ldots, a_k \in \mathbb{R}$. L'<u>insieme</u> di tutti i vettori scrivibili come combinazione lineare dei vettori $v_1, \ldots, v_k$ è detto *spazio generato* dai vettori $v_1, \ldots, v_k$ e lo indicheremo con

$\mathcal{L} = \mathcal{L}(v_1, v_2, \ldots, v_k)$, in formule:

$\mathcal{L} = \{v \in \mathbb{R}^n |\ v = a_1v_1 + a_2v_2 + \ldots + a_kv_k;\ a_1, \ldots, a_k \in \mathbb{R}\}$

1. Consideriamo l'insieme

$$\mathcal{L}(v_1) = \{v \in \mathbb{R}^3 |\ v = v_1t;\ \text{con } t \in \mathbb{R}\}.$$

Se $v_1 \neq 0$ allora $\mathcal{L}(v_1)$ sará una retta di $\mathbb{R}^3$ passante per 0, infatti ció che stiamo facendo moltiplicando il vettore v_1 per lo scalare t altro non é che dilatare o comprimere v_1 lungo la sua direzione, ció descrive, al variare di $t \in \mathbb{R}$ tutta la retta su cui v_1 giace. Se invece $v_1 = 0$, allora $\mathcal{L}(v_1)$ corrisponderá con il punto 0, in quanto il prodotto di v_1 per $t \in \mathbb{R}$ restituirá sempre 0.

2. Consideriamo ora l'insieme:

$$\mathcal{L}(v_1, v_2) = \{v \in \mathbb{R}^3 | v = a_1v_1 + a_2v_2;\ \text{con } a_1, a_2 \in \mathbb{R}\}.$$

Se v_1 e v_2 sono non proporzionali, allora questi formeranno una base del piano $\mathcal{V}_2$, e quindi $\mathcal{L}(v_1, v_2)$ sará un piano contenente 0; se v_1 e v_2 sono proporzionali allora: $v_2 = tv_1$ e

$$\mathcal{L}(v_1, v_2) = \{v \in \mathbb{R}^3 |\ v = (a_1 + ta_2)v_1;\ \text{con } a_1, a_2, t \in \mathbb{R}\}$$

e siamo nel caso dell'esempio precedente; se, invece $v_1 = v_2 = 0$ allora $\mathcal{L}(v_1, v_2) = \big\{0\big\}$ abbiamo il punto 0.

3. L'insieme:

$$\mathcal{L}(v_1, v_2, v_3) =$$

$$= \big\{v \in \mathbb{R}^3 |\ v = a_1v_1 + a_2v_2 + a_3v_3;\ \text{con } a_1, a_2, a_3 \in \mathbb{R}\big\}$$

Genererá tutto lo spazio se i tre vettori non sono complanari, se invece sono complanari genereranno un piano, mentre se sono proporzionali genereranno una retta, se infine sono tutti e tre nulli allora $\mathcal{L}(v_1, v_2, v_3)$ genererá il punto 0.

Diciamo ora che $\{v_1, \ldots, v_k\}$ vettori in $\mathbb{R}^n$ sono *linearmente indipendenti* quando:

$$a_1v_1 + a_2v_2 + \ldots + a_kv_k = 0$$

se e solamente se $a_1 = a_2 = \ldots = a_k = 0$, dove $a_1, a_2, \ldots, a_k$ sono scalari. Diremo che sono *linearmente dipendenti* se non sono linearmente indipendenti. Facciamo alcune osservazioni sull'indipendenza lineare di k vettori $v_1, \ldots, v_k$:

1. Se in $v_1, \ldots, v_k$, vi é un vettore nullo, allora i k vettori sono dipendenti;

2. Un vettore v_1 é linearmente indipendente, se e solo se é non nullo;

3. Due vettori sono indipendenti se e solo se sono non proporzionali;

4. Se due vettori v_1 e v_2 sono linearmente indipendenti, ma v_1, v_2, v_3, sono linearmente dipendenti, allora $v_3 \in \mathcal{L}(v_1, v_2)$, ovvero v_3 é combinazione lineare di v_1 e v_2. Viceversa, se $v_3 \in \mathcal{L}(v_1, v_2)$ allora v_1, v_2, e v_3 sono linearmente dipendenti;

5. Vettori di $\mathbb{R}^n$ tra loro ortogonali sono sempre linearmente indipendenti.

Definiamo la *dimensione* dello spazio $\mathcal{L}(v_1, v_2, \dots, v_k)$ generato da k vettori in $\mathbb{R}^n$ come il massimo numero di vettori linearmente indipendenti di $\mathcal{L}$.

Le tecniche per affrontare il problema di capire quando k vettori in $\mathbb{R}^n$ sono linearmente indipendenti e calcolare la dimensione dello spazio generato che essi generano verranno sviluppate alla fine del prossimo capitolo.

Capitolo 3

Matrici e sistemi lineari

3.1 Matrici

Una tabella

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & & \vdots & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \tag{3.1}$$

si chiama *matrice m per n*.

Le m n-uple orizzontali

$$(a_{11},\ a_{12},\ \ldots,\ a_{1n})\ ,\ (a_{21},\ a_{22},\ \ldots,\ a_{2n})\ ,\ldots\ (a_{m1},\ a_{m2},\ \ldots,\ a_{mn})$$

si chiamano *righe* della matrice e le n m-uple

$$\begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}, \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix}, \ldots \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix}$$

verticali sono chiamate *le colonne* di A.

Indicheremo una matrice con il simbolo

$$A = (a_{ij}), i = 1, \ldots, m, j = 1, \ldots, n. \tag{3.2}$$

Siano A e B due matrici $m \times n$:

$$\begin{pmatrix} a_{11} & a_{12} & \vdots & a_{1n} \\ a_{21} & a_{22} & \vdots & a_{2n} \\ & & \vdots & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \text{ e } \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ & & \vdots & \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{pmatrix}$$

La *somma* di A e B si scrive $A+B$ ed è la matrice $m \times n$ ottenuta come segue:

$$\begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \cdots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \cdots & a_{2n}+b_{2n} \\ & & \vdots & \\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \cdots & a_{mn}+b_{mn} \end{pmatrix} \tag{3.3}$$

Usando la notazione (3.2)

$$(A+B)_{ij} = a_{ij} + b_{ij}$$

Osservazione 18 *Osserviamo che la somma tra due matrici è possibile solo quando le due matrici hanno lo stesso numero di righe e di colonne.*

Il *prodotto* di una matrice A e uno scalare λ si scrive λA ed è la matrice $m \times n$ che si ottiene moltiplicando per λ ogni elemento di A:

$$\begin{pmatrix} \lambda a_{11} & \lambda a_{12} & \cdots & \lambda a_{1n} \\ \lambda a_{21} & \lambda a_{22} & \cdots & \lambda a_{2n} \\ & & \vdots & \\ \lambda a_{m1} & \lambda a_{m2} & \cdots & \lambda a_{mn} \end{pmatrix} \tag{3.4}$$

L'insieme delle matrici $m \times n$ verrà denotato con $\mathrm{M}_{m,n}(\mathbb{R})$ o più semplicemente $\mathrm{M}_{m,n}$ quando risulterà chiaro dal contesto che stiamo considerando il campo dei numeri reali.

Vale il seguente:

Teorema 19 *L'insieme $M_{m,n}$ è uno spazio vettoriale rispetto alle operazione di somma e moltiplicazione per uno scalare definite da (3.3) e (3.4) rispettivamente. Il vettore nullo è la matrice nulla $m \times n$ mentre l'opposto della matrice $A = (a_{ij})$ è la matrice $-A = (-a_{ij})$.*

Siano $A = (a_{ij})$ e $B = (b_{jk})$ due matrici $m \times p$ e $p \times n$ rispettivamente (il numero di colonne di A è uguale al numero di righe di B). Allora il prodotto AB è la matrice $m \times n$ il cui elemento ik è ottenuto moltiplicando la riga i-ma A_i di A per la riga j-ma B_j di B (qui si intende il prodotto scalare della riga e la colonna considerati come vettori in $\mathbb{R}^p$). In altre parole se $C = (c_{ik}), i = i, 2, \ldots m, j = 1, 2 \ldots n$ è il prodotto tra A e B allora:

$$c_{ik} = a_{i1}b_{1k} + \ a_{i2}b_{2k} + \cdots + a_{ip}b_{pk} = \sum_{j=1}^{p} a_{ij}b_{jk}.$$

Osserviamo che che il prodotto AB tra due matrici A e B non è definito se il numero di colonne di A è diverso dal numero di righe di B. Osserviamo inoltre che anche se definito il prodotto di due matrici è , in generale, non commutativo (costruire un esempio).

Valgono le seguenti relazioni:

(i) $(AB)C = A(BC), \forall A \in \mathrm{M}_{m,p}, B \in \mathrm{M}_{p,q}, C \in \mathrm{M}_{q,n}$;

(ii) $A(B + C) = AB + AC, \forall A \in \mathrm{M}_{m,p}, B, C \in \mathrm{M}_{p,n}$;

(iii) $(B + C)A = BA + CA, \forall B, C \in \mathrm{M}_{m,p}, A \in \mathrm{M}_{p,n}$;

(iv) $\lambda(AB) = (\lambda A)B = A(\lambda B), \forall A \in \mathrm{M}_{m,p}, B \in \mathrm{M}_{p,n}, \lambda \in \mathbb{R}$;

(v) $O_{m,p}A = O_{m,n}, \forall A \in \mathrm{M}_{p,n}$ dove $O_{m,p}$ (risp. $O_{m,n}$) denota la matrice nulla $m \times p$ (risp. $m \times n$);

(vi) $BO_{p,n} = O_{m,n}, \forall B \in \mathrm{M}_{m,p}$, dove $O_{p,n}$ denota la matrice nulla $p \times n$.

La *trasposta* A^T di una matrice $A \in \mathrm{M}_{m,n}$ è la matrice in $\mathrm{M}_{n,m}$ ottenuta da A scambiando le sue righe con le sue colonne.

Una matrice $n \times n$ cioè che presenta lo stesso numero di righe e colonne si dice *quadrata*. Il numero di righe (e quello delle colonne) si chiama *ordine* della matrice. L'insieme delle matrici quadrate $n \times n$ verrà denotato con $\mathrm{M}_n(\mathbb{R})$ o più semplicemente M_n quando risulterà chiaro dal contesto che stiamo considerando il campo dei numeri reali. Una matrice quadrata $A \in M_n$ si dice *simmetrica* (risp. antisimmetrica) se $A^T = A$ (risp. $A^T = -A$).

La *diagonale* principale di $A = (a_{ij}) \in M_n$ è costituita dagli elementi $a_{11}, a_{22}, \ldots, a_{nn}$.
Data una matrice quadrata $A \in M_n$ la *traccia* di A è il numero reale che si ottiene sommando gli elementi della sua diagonale principale. e si indica con $\operatorname{tr} A$, cioè

$$\operatorname{tr} A = \sum_{j=1}^{n} a_{jj}.$$

Osserviamo che se A e B sono due matrici quadrate allora ha senso considerare sia AB che BA. Se $A = (a_{ij})$ è una matrice quadrata. La *diagonale* di A è costituita dagli elementi $a_{11}, a_{22}, \ldots, a_{nn}$.
La matrice diagonale che ha tutti gli elementi 1 nella diagonale e 0 altrove si chiama la matrice *identica* e si indica con I_n

Osserviamo che la matrice identica $I_n = (\delta_{ij})$ è tale che

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

Essa soddisfa: $AI_n = I_nA$ per ogni $A \in M_n$. Uno matrice $A = (a_{ij})$ è detta *triangolare superiore* se $a_{ij} = 0, i \leq j$ mentre è *triangolare inferiore* se $a_{ij} = 0, i \geq j$.
Una matrice quadrata A di ordine n si dice *invertibile* se esiste una matrice B tale che:

$$AB = BA = I_n. \tag{3.5}$$

La matrice B che soddisfa la (3.5) è unica:

$AB_1 = B_1A = I_n$ e $AB_2 = B_2A = I_n$ implicano

$$B_1 = B_1I_n = B_1(AB_2) = (B_1A)B_2 = I_nB_2 = B_2.$$

La matrice B che soddisfa (3.5) si chiama la matrice *inversa* di A e si indica con A^{-1}.

Osserviamo che che la moltiplicazione di due matrici di ordine n A e B da origine ad una matrice invertibile AB la cui inversa è $B^{-1}A^{-1}$. Più in generale se $A_1, A_2, \ldots A_p$ sono matrici invertibili allora $A_1A_2 \ldots A_p$ è invertibile con inversa $A_p^{-1} \ldots A_2^{-1}A_1^{-1}$.

Il primo caso non banale di inversione di una matrice quadrata è il caso 2×2 (se a è una matrice non nulla 1×1, cioè un numero non nullo allora il suo inverso è dato da $\frac{1}{a}$).

Sia allora

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

una matrice 2×2. Cerchiamo degli scalari x, y, z, t tali che

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x & y \\ z & t \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

o equivalentemente

$$\begin{aligned} ax + bz &= 1, \quad & ay + bt &= 0 \\ cx + dz &= 0, \quad & cy + dt &= 1 \end{aligned}$$

I due sistemi ammettono soluzione se e solo se $\det A = ad - bc \neq 0$ e si ottiene

$$A^{-1} = \frac{1}{\det A} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Per trovare l'inversa di una matrice di ordine n qualunque o per capire se una tale inversa esiste abbiamo bisogno di alcune definizioni.

Determinante di una matrice quadrata

Il *determinante* di una matrice quadrata $n \times n$ si definisce come (*sviluppo di Laplace rispetto alla prima riga*)

$$\det A = \sum_{j=1}^{n} (-1)^{1+j} a_{1j} A_{1j},$$

dove A_{1j} è la matrice $(n-1) \times (n-1)$ ottenuta eleminando la prima riga e la colonna j-sima dalla matrice A.

Si può dimostrare che il determinante può essere sviluppato anche secondo una riga qualunque tenendo conto opportunamente dei segni. Più precisamente vale la seguente formula, che rappresenta lo sviluppo rispetto alla i-sima riga.

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

Definiamo ora le operazioni elementari sulle righe di una matrice $A \in M_{m,n}$.

Sia A_i la riga i-esima di A.

Allora
$$A = \begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_m \end{pmatrix}.$$
Consideriamo le seguenti operazioni elementari:

1. Scambiare la riga i-esima con la riga j-esima $A_i \leftrightarrow A_j$;
2. moltiplicare la riga i-esima per uno scalare non nullo λ, $A_i \rightarrow \lambda A_i$;
3. sostituire alla riga i-esima λ volte la j-esima sommata alla riga i-esima stessa, $A_i \rightarrow \lambda A_j + A_i$.

Valgono le seguenti proprietà:

- il determinante cambia di segno se si scambiano due righe tra loro;
- il determinante risulta moltiplicato per λ se si moltiplica una riga per λ;
- il determinante non cambia se si applica un' operazione elementare di tipo 3;

- il determinante di una matrice con due righe uguali è nullo;
- il determinante di una matrice con due righe proporzionali è nullo;
- è nullo il determinante di una matrice che ha una riga che è combinazione lineare delle altre righe;
- $\det(A) = \det(A^T)$;
- $\det(AB) = \det A \det B$

Qualche criterio

Un criterio per capire se una matrice è invertibile è il seguente: *una matrice invertibile se e solo se il suo determinante è diverso da zero.*

Un criterio per capire se n vettori in $\mathbb{R}^n$ formano una base è il seguente: *n-vettori in $\mathbb{R}^n$ formano una base se e solo se la matrice $n \times n$ che ha come righe i vettori in questione ha determinante diverso da zero.*

Algoritmo per trovare l'inversa di una matrice

L' algoritmo per trovare l'inversa di una matrice quadrata A di ordine n, o per determinare se A non è invertibile è il seguente:

Passo 1. formare la matrice $n \times 2n$ $M = (A|I_n)$: A nella metà sinistra I_n nella metà destra di M.

Passo 2. usando operazioni elementari sulle righe ridurre M nella forma $(I_n|B)$, dove I_n è la matrice identità $n \times n$

Passo 3. porre $A^{-1} = B$.

Esempio 20 Trovare l'inversa della matrice

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

Scriviamo la matrice 4×8

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

Applicando le operazioni elementari:

$A_3 \to -A_1 + A_3, A_4 \to -A_1 + A_4$

si ottiene:

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

Applicando le operazioni elementari:

$A_1 \to -2A_2 + A_1, A_3 \to -A_2 + A_3, A_4 \to -2A_2 + A_4$

si ottiene:

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

Applicando le operazioni elementari:

$A_1 \to -3A_3 + A_1, A_2 \to A_3 + A_2, A_4 \to A_3 + A_4$

si ottiene:

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

Applicando l' operazione elementare: $A_4 \to -A_4$ si ottiene:

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

Infine, applicando le operazioni elementari:

$A_3 \to 3A_4 + A_3, A_2 \to 2A_4 + A_2, A_1 \to -7A_4 + A_1$

si ottiene:

$$\begin{pmatrix} 1 & 0 & 0 & 0 & -10 & -20 & 4 & 7 \\ 0 & 1 & 0 & 0 & 3 & 6 & -1 & -2 \\ 0 & 0 & 1 & 0 & 5 & 8 & -2 & -3 \\ 0 & 0 & 0 & 1 & 2 & 3 & -1 & -1 \end{pmatrix}.$$

Di conseguenza:

$$A^{-1} = \begin{pmatrix} -10 & -20 & 4 & 7 \\ 3 & 6 & -1 & -2 \\ 5 & 8 & -2 & -3 \\ 2 & 3 & -1 & -1 \end{pmatrix}.$$

Esercizio: verificare che la matrice scritta sopra è effettivamente l'inversa della matrice A.

Osservazione 21 Se applicando l'algoritmo precedente si viene a creare nella metà sinistra di M una riga o una colonna nulla questo significa che la matrice non è invertibile.

Rango di una matrice

Sia $A \in M_{m,n}$. Un *minore di ordine* p di A è una matrice $p \times p$ che si ottiene da A cancellando $m - p$ righe e $n - p$ colonne. Diremo che il *rango* di A è r, e scriveremo $\mathrm{rg}(A) = r$, se A ha un minore M di ordine di r con $\det M \neq 0$ e tutti i minori di ordine maggiore di r hanno determinante nullo.

Calcolo del rango di una matrice Il rango di una matrice è invariante per operazioni elementari. Questo fornisce anche un metodo per il calcolo del rango di una matrice. Un altro metodo per calcolare il rango di una matrice è il così-detto teorema degli orlati. Prima dio enunciarlo diamo una definizione.

Definizione Sia $A \in M_{m,n}$ e sia B un minore di ordine p di A. Se alla matrice B aggiungiamo gli elementi di un'altra riga o un'altra colonna di A diremo che stiamo orlando la matrice B.

Teorema (degli orlati) Sia $A \in M_{m,n}$. Allora il rango di A è uguale a r se e solo se esiste un minore B di ordine r di A invertibile ($\det B \neq 0$), e tutte le sottomatrici di ordine $r+1$ di A ottenute orlando B hanno determinante nullo.

A lezione verranno svolti vari esempi di calcolo del rango di una matrice usando le operazione elementari e il teorema degli orlati (vedi anche l'ultimo paragrafo).

3.2 Sistemi di equazioni lineari

Consideriamo un sistema di m equazioni lineari, diciamo $L_1, L_2, \ldots, L_m$ in n incognite $x_1, x_2, \ldots, x_n$ nella forma

$$
\begin{cases}
a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1 \\
a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & b_2 \\
 & & & & \vdots & & & & \\
a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & b_m
\end{cases}
\tag{3.6}
$$

dove le a_{ij} e le b_i sono costanti. Una *soluzione* (o *soluzione particolare*) di detto sistema è un n-upla $u = (\lambda_1, \lambda_2, \dots \lambda_n)$ di costanti soluzione di ogni equazione del sistema. L'insieme di tutte queste soluzioni si chiama *soluzione generale* del sistema. Sistemi di equazioni lineari nelle stesse incognite si dicono *equivalenti* quando hanno la stessa soluzione generale. Un modo per ottenere un sistema equivalente ad uno assegnato è di applicare una sequenza delle seguenti operazioni elementari:

- Scambiare l'equazione i-ma con la j-ma;
- moltiplicare l'equazione i-ma per uno scalare non nullo λ;
- sostituire l'equazione i-ma con λ volte la j-ma sommata alla i-ma.

Un metodo per risolvere un sistema di equazioni lineari è quello di *eliminazione di Gauss*:

Passo 1. Usando le operazioni elementari ridurre il sistema ad uno equivalente più semplice (in forma triangolare o a gradini, vedi oltre).

Passo 2. Trovare le soluzioni del sistema semplificato

Supponiamo che durante l'applicazione del passo 1 si ottenga l'equazione

$$L : 0x_1 + 0x_2 + \cdots + 0x_n = b$$

allora:

- se $b = 0$, L si può cancellare senza che l'insieme delle soluzioni cambi;
- se $b \neq 0$ allora il sistema non ha soluzioni.

Sistemi in forma triangolare

Un sistema di equazioni lineari è in *forma triangolare* se il numero di equazioni è uguale a quello delle incognite, e se x_k è incognita iniziale dell'equazione k-ma. Perciò un sistema triangolare di equazioni lineari avrà la forma:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1,n-1}x_{n-1} + a_{1n}x_n = b_1 \\ \qquad\quad a_{22}x_2 + \cdots + a_{2,n-1}x_{n-1} + a_{2,n-1}x_{n-1} = b_2 \\ \qquad\qquad\qquad \vdots \\ \qquad\qquad a_{n-1,n-1}x_{n-1} + a_{n-1,n}x_n = b_{n-1} \\ \qquad\qquad\qquad a_{nn}x_n = b_n \end{cases}$$

dove $a_{11} \neq 0, a_{22} \neq 0 \ldots a_{nn} \neq 0$.

Il suddetto sistema ha una soluzione unica, che si può ottenere con un procedimento di sostituzione a posteriori: si ricava x_n dall'ultima equazione lo si sostituisce nella penultima che si risolve rispetto all'incognita x_{n-1} etc. Il procedimento ha fine quando si trova la prima incognita x_1.

Sistema a gradini

Un sistema di equazioni lineari è in *forma a gradini* se è della forma seguente:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ \qquad\quad a_{2j_2}x_{j_2} + a_{2,j_2+1}x_{j_2+1} + \cdots + a_{2n}x_n = b_2 \\ \qquad\quad \vdots \\ \qquad\quad a_{rj_r}x_{j_r} + a_{r,j_r+1}x_{j_r+1} + \cdots + a_{rn}x_n = b_r \end{cases}$$

dove $1 < j_2 < \cdots < j_r$ e $a_{11} \neq 0, a_{2j_2} \neq 0, \ldots, a_{rj_r} \neq 0$.

Un'incognita x_k in un sistema a gradini si chiama *variabile libera* se essa non è incognita iniziale in alcuna equazione, ovvero $x_k \neq x_1, x_k \neq x_{j_2}, \ldots, x_k \neq x_{j_r}$.

Osserviamo che $r \leq n$. Se $r = n$ il sistema è allora in forma triangolare e quindi ammette un'unica soluzione; se $r < n$ ci sono meno equazioni che incognite. Si possono allora assegnare valori arbitrari alle $n - r$ variabili libere e ottenere una soluzione del sistema. Al variare di questi valori arbitrari si ottengono soluzioni diverse, quindi il sistema ha infinite soluzioni.

Algoritmo di Gauss–Jordan

Quest'algoritmo permette di ridurre un sistema di m equazioni in n incognite ad un sistema a gradini (o in forma) triangolare ovvero determina se il sistema non ha soluzioni.

Passo 1. Scambiare le righe in modo che la prima incognita x_1 appaia con un coefficiente non nullo nella prima equazione: $a_{11} \neq 0$;

Passo 2. Usare a_{11} come perno per eliminare x_1 da tutte le equazioni meno che la prima (sostituendo la i-ma equazione $i > 1$ con la prima equazione moltiplicata per $-\frac{a_{i1}}{a_{11}}$

sommata alla i-ma equazione);

Passo 3. Esaminare ogni nuova equazione L:

(a) se L ha la forma

$$L : 0x_1 + 0x_2 + \cdots + 0x_n = 0$$

o se L è multipla di un'altra equazione, cancellarla dal sistema;

(b) se L ha la forma

$$L : 0x_1 + 0x_2 + \cdots + 0x_n = b,$$

$b \neq 0$, uscire dall'algoritmo: il sistema non ha soluzioni.

Passo 4. Ripetere i passi $1, 2, 3$ sul sottosistema formato da tutte le equazioni meno la prima;

Passo 5. Continuare il procedimento finchè il sistema non si presenta in forma a gradini, o finche nel caso (b) non si ottiene un'equazione degenere.

Sistemi omogenei

Una classe importante di sistemi sono quelli *omogenei*

Un sistema omogeneo di m equazioni in n incognite ha la forma:

$$\begin{cases} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & 0 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & 0 \\ & & & & \vdots & & & & \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & 0 \end{cases} \tag{3.7}$$

Un tale sistema ha sempre una soluzione, la soluzione nulla consistente della n-upla $(0, 0, \ldots, 0)$, detta anche *soluzione banale*. Ogni altra soluzione, se esiste, si chiama *soluzione non banale*.

Come conseguenza un sistema omogeneo potrà essere sempre ricondotto ad un sistema omogeno in forma a gradini:

$$\begin{cases} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & & & + & a_{1n}x_n & = & 0 \\ & & a_{2j_2}x_{j_2} & + & a_{2,j_2+1}x_{j_2+1} & + & \cdots & + & a_{2n}x_n & = & 0 \\ & & & \vdots & & & & & & & \\ & & a_{rj_r}x_{j_r} & + & a_{r,j_r+1}x_{j_r+1} & + & \cdots & + & a_{rn}x_n & = & 0 \end{cases}$$

dove $1 < j_2 < \cdots < j_r$ e $a_{11} \neq 0, a_{2j_2} \neq 0, \ldots, a_{rj_r} \neq 0$.

Ci sono quindi due possibilità :

(i) $r = n$, il sistema è triangolare e ammette quindi la sola soluzione nulla.

(ii) $r < n$, il sistema ha una soluzione non zero.

Dai risultati precedenti si deduce il seguente teorema e il suo corollario.

Teorema 22 *(Rouché–Capelli) Sia A una matrice $m \times n$. Allora il sistema $Ax = b$ è compatibile (ha almeno una soluzione) se e solo se $\mathrm{rg}\, A = \mathrm{rg}(A|b) = r$. Quando il sistema è compatibile, le soluzioni dipendono da $n - r$ parametri (variabili libere). In particolare la soluzione è unica se $\mathrm{rg}\, A = n$.*

Nel caso di sistemi compatibili diremo anche che il sistema ha ∞^{n-r} soluzioni. Il numero $n - r$ è detto la *dimensione* dell'insieme delle soluzioni.

Corollario 23 *Sia A una matrice $m \times n$. Allora il sistema omogeneo $Ax = 0$ ha una soluzione non banale $x \neq 0$ se e solo se $\mathrm{rg}\, A < n$. Perciò se $m < n$, $Ax = 0$ ha sempre una soluzione non banale.*

Nel caso $m = n$ (sistemi quadrati ossia $A \in M_n(\mathbb{R})$):

- Ax=0 ha una soluzione non banale se e solo se A è **non** invertibile ($\det A = 0$),
- $Ax = b$ ha esattamente una soluzione se e solo se A è invertibile ($\det A \neq 0$),

Matrici, sistemi e combinazioni lineari

I vettori $\mathbf{e_i}$ sono i vettori di $\mathbb{R}^n$ con 1 nell' i-esima posizione e 0 altrove.

Dati $\mathbf{v_1}, \ldots, \mathbf{v_k} \in \mathbb{R}^n$, possiamo formare la matrice $n \times k$, $A = (\mathbf{v_1}, \ldots, \mathbf{v_k})$ che ha per colonne i vettori dati. Allora ogni combinazione lineare dei vettori $\mathbf{v_1}, \ldots, \mathbf{v_k}$ si può esprimere come un prodotto di matrici:

$$x_1\mathbf{v_1} + x_2\mathbf{v_2} + \cdots + x_k\mathbf{v_k} = Ax,$$

dove x è il vettore colonna di componenti $(x_1, x_2, \ldots, x_k)$.

Ciò consente di riformulare le nozioni di dipendenza lineare, indipendenza lineare e spazio generato nella terminologia dei sistemi lineari:

1. I vettori $\mathbf{v_1}, \ldots, \mathbf{v_k} \in \mathbb{R}^n$ sono linearmente dipendenti se e solo se esiste una soluzione $x \neq 0$ del sistema lineare omogeneo $Ax = 0$;

2. I vettori $\mathbf{v_1}, \ldots, \mathbf{v_k} \in \mathbb{R}^n$ sono linearmente indipendenti se e solo se l'unica soluzione del sistema lineare omogeneo $Ax = 0$ è quella banale $x = 0$;

3. Un vettore $\mathbf{b}$ appartiene a $\mathcal{L}(\mathbf{v_1}, \ldots, \mathbf{v_k})$ se e solo se il sistema lineare $Ax = \mathbf{b}$ è compatibile.

Combinando le affermazioni precedenti con i teoremi relativi ai sistemi lineari, otteniamo:

4. k vettori di $\mathbb{R}^n$ sono linearmente indipendenti se e solo se la corrispondente matrice $n \times k$ ha rango k (quindi $k \leq n$);

5. se $k > n$, k vettori di $\mathbb{R}^n$ sono linearmente dipendenti (Attenzione: il viceversa non è vero. Fare un esempio);

6. k vettori di $\mathbb{R}^n$ generano $\mathbb{R}^n$ se e solo se la matrice A, $n \times k$ ha rango n. Questo richiede che $k \geq n$. (Attenzione non è vero che un qualsiasi insieme di n o più vettori di di $\mathbb{R}^n$ genera $\mathbb{R}^n$. Fare un esempio).

7. La dimensione dello spazio $\mathcal{L}(\mathbf{v_1}, \ldots, \mathbf{v_k})$ generato da k vettori è il rango della matrice A.