Roberto Lucchetti e Paola Radrizzani

Algebra Lineare e Geometria

Prefazione

Queste note nascono nell'ambito del corso della Laurea on Line in Ingegneria Informatica al Politecnico di Milano. Negli ultimi anni, questa parte di programma è stata dapprima incorporata in un unico corso (Elementi di Analisi Matematica e Geometria) poi, in qualche corso di studio del Politecnico, di nuovo rimessa come esame autonomo. Sta di fatto che oggi un testo tradizionale di geometria ed algebra lineare in genere contiene molto più materiale di quello che può venire erogato in un corso del primo anno di Ingegneria. Allo stesso tempo, i testi di analisi e di calcolo pubblicati negli ultimi tempi, e che svolgono anche una parte di algebra lineare (ma in genere non di geometria), sono costretti a comprimere questa parte, avendo appunto tutta la parte del calcolo da svolgere. Per questo, abbiamo pensato utile avere delle note come queste, che richiamano gli aspetti fondamentali della teoria, fanno dimostrazione di alcuni risultati di base, quando ritenute significative, presentano molti esempi, propongono molti esercizi, sia in forma di test (autovalutazione), sia in forma aperta. Gli autori, a questo proposito, intendono mettere soluzioni e/o risultati in una pagine web all'uopo creata.

Il libro comincia con una breve introduzione relativa ai numeri complessi, di cui descriviamo le principali proprietà, fino ad arrivare a trovare le radici ennesime di un numero complesso. Il capitolo si chiude con l'enunciazione del teorema fondamentale dell'algebra. Passiamo poi a considerare i vettori nel piano, le operazioni con essi, le componenti cartesiane, il prodotto scalare, e le rette nel piano, di cui discutiamo la forma generale e quella parametrica. Introduciamo poi i vettori nello spazio, le equazioni generale e parametrica del piano, la descrizione delle mutue posizioni di rette e piani nello spazio. Si passa poi alle matrici, alle principali operazioni fra matrici, alla definizione di determinante e di inversa, introdotta con lo scopo di mostrare il modo elegante seppur teorico di risolvere i sistemi lineari, per mezzo della regola di Cramer. Nel capitolo successivo introduciamo gli spazi vettoriali, nonché le nozioni di base ad essi legati, come l'indipendenza lineare, la dimensione eccetera. Passiamo poi ad uno studio un poco più approfondito dei sistemi lineari, uscendo dal caso della regola di Cramer, per arrivare al teorema di Rouché-Capelli. Nel capitolo seguente diamo l'idea, abbastanza succintamente, di trasformazione lineare fra spazi euclidei, e delle sue proprietà di base. Passiamo poi al problema della diagonalizzabilità delle matrici, per dedicare il capitolo successivo alle matrici simmetriche, ortogonali ed emisimmetriche. Il libro si chiude con due capitoli, il primo dedicato alle coniche, il secondo, brevissimo, alle forme quadratiche.

Alcuni capitoli contengono degli approfondimenti, che crediamo utili a chi voglia conoscere un po' meglio alcuni aspetti della materia,ma che non sono necessari ad un primo studio della stessa. Tutti i capitoli contengono, alla fine, esercizi proposti, alcuni in forma chiusa, altri in forma aperta.

Nel preparare queste note abbiamo utilizzato materiale elaborato in questi cinque anni di insegnamento alla Laurea On Line, e quindi ringraziamo chi ha collaborato con noi in questo periodo, in particolare G. Arioli, B. Balossi, M. Galeazzi, P.L. Moseneder. Ringraziamo inoltre M. Grasselli, C. Morosi e M. Verri per alcuni utili commenti e per averci dato del materiale che abbiamo incluso nel libro.

1

Numeri complessi

1.1 Insiemi numerici

I primi numeri di cui gli uomini hanno avuto percezione sono ovviamente i numeri naturali, l'insieme dei quali viene indicato con $\mathbb{N}$, e che servono per contare:

$$\mathbb{N} = \{1, 2, \dots, n, \dots\}^1.$$

In seguito, si è ritenuto utile introdurre un insieme che comprendesse anche i numeri negativi, e che viene indicato con$\mathbb{Z}$:

$$\mathbb{Z} = \{0, 1, -1, 2, -2 \dots, n, -n \dots\}.$$

Concettalmente, l'idea di numero relativo è già sofisticata, sebbene abbastanza naturale: tutti noi, anche quelli poco inclini alla matematica, siamo inquieti se sul nostro conto corrente il saldo risulta essere di euro $-1245, 37$. Sostanzialmente, i numeri relativi sono stati introdotti perché il risultato di una differenza tra numeri naturali potrebbe non essere un numero naturale. Una conseguenza interessante dell'aver introdotto i numeri negativi è che possiamo "ignorare" il concetto di differenza[2]. Il passo sucessivo consiste nel considerare l'insieme ($\mathbb{Q}$)delle frazioni:

$$\mathbb{Q} = \{\frac{p}{q} : p \in \mathbb{Z}, q \in \mathbb{N}\}^3.$$

[1] Una questione senza risposta è se zero appartiene all'insieme dei numeri naturali. Per noi, ci appartiene oppure no a seconda di come ci fa comodo, ed il contesto dovrebbe rendere comprensibile che cosa intendiamo

[2] La differenza fra 5 e 12 è la somma fra 5 e -12: questo può sembrare solo formale, ma ha il vantaggio di eliminare la necessità di elencare le proprietà dell'operazione differenza, che sono automaticamente ereditate da quelle di somma

[3] Convenzionalmente, per indicare una frazione negativa, mettiamo il segno meno al numeratore: $\frac{5}{-3} = \frac{-5}{3}$ inoltre, più praticamente, scriveremo $-\frac{5}{3}$

Essenzialmente, i numeri frazionari permettono di considerare parti di oggetti fissati, e quindi di fare misurazioni accurate: una fetta di torta divisa in cinque parti uguali rappresenta $\frac{1}{5}$ della torta, la barriera nel calcio deve stare a 9.15 metri dal pallone, cioè a $\frac{915}{100}$ di metro, anche se quelli della Juventus (e non solo) non rispettano mai la distanza. Osserviamo che, naturalmente, gli insiemi introdotti sono sempre più grandi del precedente, nel senso che lo include ed aggiunge nuovi elementi[4]. Tuttavia i numeri razionali non sono sufficienti a descrivere enti geometrici significativi: ad esempio la diagonale del quadrato con lato di lunghezza 1 non ha lunghezza razionale, e lo stesso succede per la lunghezza della circonferenza di diametro unitario. Dobbiamo allora introdurre un nuovo insieme, indicato con $\mathbb{R}$, chiamato l'insieme dei numeri reali, che rappresenta un completamento di $\mathbb{Q}$. In che senso? un modo per vedere questo è il seguente: dati due insiemi A e B di numeri razionali, con la proprietà che se $a \in A$ e $b \in B$ allora $a \leq b$,[5] non è detto che esista in $\mathbb{Q}$ un numero che sia allo stesso tempo più grande di tutti quelli di A e più piccolo di tutti quelli di b.[6] L'insieme dei numeri reali si può vedere come completamento di $\mathbb{Q}$ nel senso che tra due insiemi separati esiste sempre almeno un separatore[7]. In effetti $\mathbb{R}$ rappresenta l'insieme "giusto" per il calcolo infinitesimale, uno strumento potente che si considera nei corsi di analisi e che porta a risultati molto sofisticati.

Tuttavia, per certe esigenze, nemmeno i numeri reali sono sufficienti. Ad esempio, l'equazione $x^2 + 1 = 0$ non ha soluzioni reali, in quanto nei reali un numero elevato al quadrato è sempre positivo. Tuttavia per molti motivi sarebbe comodo che un'equazione del genere avesse soluzioni (possibilmente due). E dunque, esattamente come si è sentita l'esigenza di introdurre numeri che non fossero rapporto di interi, allo stesso modo è utile *allargare* il sistema dei numeri reali per trovare soluzioni di equazioni non risolvibili nell'insieme dei reali stessi, e questo si fa introducendo l'insieme dei numeri complessi. La cosa interessante è che a volte certi risultati che ci interessano nell'insieme dei reali possono essere ottenuti solo passando all'insieme dei complessi. In altre parole, quel che cerchiamo possiamo trovarlo nell'insieme dove è naturale cercarlo, ma solo se all'inizio facciamo le nostre ricerche in uno spazio più grande. Un esempio sarà la dimostrazione che le matrici simmetriche hanno autovalori reali: che ne abbiano, dipende dal fatto che li cerchiamo nei complessi, poi certe proprietà di quelle matrici ci permettono di concludere che stanno nell'insieme più piccolo dei numeri reali.

Nell'appendice dedicata agli approfondimenti, descriveremo alcune delle proprietà di struttura che caratterizzano gli insiemi precedentemente introdotti. Il prossimo pa-

[4] Se ogni volta che abbiamo costruito un nuovo insieme se ne è costruito uno con un numero maggiore di elementi è uno di quei quesiti che portano a paradossi sull'infinito molto interessanti: la risposta, forse sorprendentemente, è: no!

[5] Si dice che A e B sono insiemi *separati*

[6] Tra l'insieme dei numeri positivi a tali che $a^2 \leq 2$ e quelli positivi b tali che $b^2 >$ non esistono numeri razionali

[7] Nell'esempio precedente si tratta di $\sqrt{2}$, che già gli antichi sapevano essere la diagonale del quadrato di lato unitario, ed un numero non esprimibile come rapporto di interi

ragrafo invece si occupa dell'introduzione dell'insieme dei complessi, e delle principali proprietà dei numeri complessi.

1.2 I numeri complessi

Introduciamo un nuovo numero, chiamato i, con la proprietà che $i^2 = -1$, e che soddisfi le usuali regole delle potenze. Così:

$$i^0 = 1, \quad i^1 = i, \quad i^2 = -1, \quad i^3 = i^2 i = -i, \quad i^4 = 1, \quad \ldots$$ [8]

Definizione 1.2.1 *Si definisce* numero complesso *un'espressione della forma*

$$z = a + ib,$$

con a e b numeri reali e i tale che $i^2 = -1$. i viene detto unità immaginaria*.*

L'insieme dei numeri complessi si indica con $\mathbb{C}$:

$$\mathbb{C} \stackrel{\text{def}}{=} \{z : z = a + ib, \ a, b \in \mathbb{R}\}.$$

Nel numero complesso $z = a + ib$, che si dice espresso *in forma algebrica*, a è chiamata *parte reale di z, b parte immaginaria*. Si scrive anche: $a = \text{Re } z$; $b = \text{Im } z$.

L'insieme dei complessi $\mathbb{C}$ contiene come sottoinsieme l'insieme dei numeri reali (che sono identificati con quei numeri complessi che hanno parte immaginaria nulla: $b = 0$). I numeri complessi $a + ib$ tali che $a = 0$ vengono invece detti *immaginari*.

Osservazione 1.2.1 Due numeri complessi sono uguali se e solo se sono uguali rispettivamente le loro parti reali e le loro parti immaginarie: $a + ib = c + id$ se e solo se $a = c, b = d$.

Per poter operare con i complessi, occorre definire le operazioni sull'insieme $\mathbb{C}$.

Per quanto riguarda l'addizione e la moltiplicazione, si procede trattando i numeri come reali ed usando la proprietà $i^2 = -1$: se $z = a + ib$ e $w = c + id$, allora

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

Ecco alcune proprietà dell'addizione e della moltiplicazione: dati $z, v, w \in \mathbb{C}$ vale:

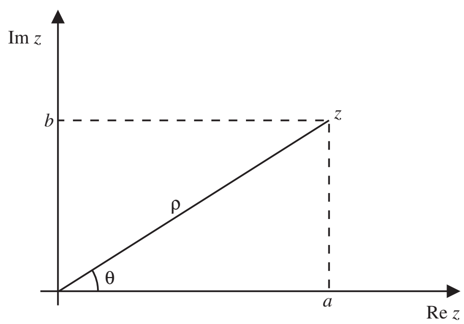
1. $z + w = w + z$, $zw = wz$;
2. $(z + v) + w = z + (v + w)$, $(zv)w = z(vw)$;

[8] Introdurre un numero che abbia queste proprietà, o anche altre, si può sempre fare; si tratta poi di vedere se i risultati che ne seguono sono (non contradditori e) interessanti. In questo caso, nessun dubbio!

3. Esiste l' *opposto* di z, indicato con $-z$, che ha la proprietà che $z + (-z) = 0$, se $z \neq 0$, esiste il *reciproco* di z, indicato con $\frac{1}{z}$, che ha la proprietà che $z\frac{1}{z} = 1$;
4. $z(u + w) = zu + zw$.

Osservazione 1.2.2 Riguardo alle proprietà precedenti, si ha in particolare che $-z = -a - ib$ e $\frac{1}{z} = \frac{1}{a+ib} = \frac{a-ib}{a^2+b^2} = \frac{a}{a^2+b^2} + i\frac{-b}{a^2+b^2}$. (Vedi esercizio 1.4.1)

Un numero complesso $a + ib$ è naturalmente identificato dalla coppia di numeri reali (a, b), e quindi può essere graficamente rappresentato dal punto (a, b) del piano cartesiano. In tal caso il piano viene detto *piano complesso* o piano di *Argand-Gauss*. L'asse x si chiama *asse reale*, l'asse y *asse immaginario*.



$z = a + ib$
Figura 1.1

La rappresentazione di $z = a + ib$ sul piano ci permette di esprimere un numero complesso anche in un'altra forma, che è spesso molto utile.
Indichiamo con θ l'angolo che il semiasse positivo delle x forma con la semiretta passante per l'origine ed il punto (a, b); si ha allora che

$$z = a + ib = \sqrt{a^2 + b^2}(cos\theta + i \sin\theta).$$

Osserviamo che nella formula precedente θ può essere sostituito da un qualsiasi angolo ϕ tale che $\phi = \theta + 2k\pi$ con $k \in \mathbb{Z}$. Quindi ha senso dare la seguente definizione:

Definizione 1.2.2 *Si chiama* argomento *di z (e si scrive* Arg z*) l'insieme:*

$$\{\theta \in \mathbb{R} : \; z = \rho(\cos\theta + i \sin\theta)\}.$$

Si chiama argomento principale *di z l'angolo $\theta \in [0, 2\pi) \cap$* Arg z.[9].

Definizione 1.2.3 *Si chiama* modulo *di z e si indica con $|z|$ il numero reale positivo:*

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

[9] In alcuni testi l'argomento principale viene assunto tra $(-\pi, \pi]$

$|z|$ rappresenta la distanza dell'origine O del piano dal punto $P = (a, b)$. Osservare che se z è reale, allora $|z|$ rappresenta il *valore assoluto* del numero reale stesso.

Quando si scrive $z = \rho(\cos\theta + i \sin\theta)$, con ρ il suo modulo e θ un suo argomento, si dice che il numero complesso z è espresso in *forma polare*.

Esempio 1.2.1 Se $z = 1 + i$, allora $\rho = \sqrt{2}$ e $\theta = \frac{\pi}{4}$.

Vediamo ora le relazioni fra ρ, θ (coordinate polari del piano) ed a, b.
Intanto si ha, come già visto:

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

Inoltre:

$$\begin{cases} a = \rho\cos\theta \\ b = \rho\sin\theta \end{cases}.$$

Si ha quindi $\cos\theta = \frac{a}{\rho}$, $\sin\theta = \frac{b}{\rho}$, e dunque, se $a \neq 0$,

$$\tan\theta = \frac{b}{a}.$$

Questa relazione non ci permette di trovare una formula immediata per θ, in quanto la tangente non è una funzione invertibile. Però se $0 \leq \theta < \pi/2$, si ottiene $\theta = \arctan\frac{b}{a}$ e negli altri quadranti si possono usare opportune simmetrie per ottenere un' espressione algebrica per θ. Inoltre, come vedremo in qualche esempio, il calcolo, in maniera approssimata, di θ, si può fare a partire dalla formula $\tan\theta = \frac{b}{a}$.

Esempio 1.2.2 Se $z = 2(\cos\frac{\pi}{6} + i \sin\frac{\pi}{6})$, allora $z = 2(\frac{\sqrt{3}}{2} + i\frac{1}{2}) = \sqrt{3} + i$.

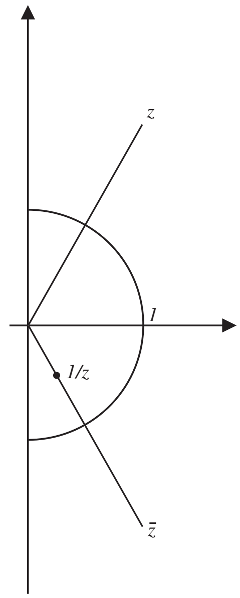
Definizione 1.2.4 *Si chiama* coniugato *di $z = a + ib$ il numero complesso $\bar{z} = a - ib$.*

Geometricamente, $\bar{z}$ rappresenta il punto del piano simmetrico di z rispetto all'asse x.

Proposizione 1.2.1 *Si ha*

$$z + \bar{z} = 2\,\mathrm{Re}\, z, \qquad z\bar{z} = |z|^2, \qquad \frac{1}{z} = \frac{\bar{z}}{|z|^2}.$$

Se $z = \rho(\cos\vartheta + i \sin\vartheta)$, allora $\overline{z} = \rho(\cos\vartheta - i \sin\vartheta)$



z ,$\bar{z}$, $\frac{1}{z}$
Figura 1.2

La seguente proposizione ci dà la formula del prodotto di due numeri complessi espressi in forma polare.

Proposizione 1.2.2 *Se* $z = \rho(\cos\theta + i\sin\theta)$ *e* $w = \tau(\cos\phi + i\sin\phi)$, *allora*

$$z \cdot w = \rho\tau(\cos(\theta + \phi) + i\sin(\theta + \phi)).$$

$z{\cdot}w = \rho(\cos\theta{+}i\sin\theta){\cdot}\tau(\cos\phi{+}i\sin\phi) = \rho\tau((\cos\theta\cos\phi{-}\sin\theta\sin\phi){+}i(\cos\theta\sin\phi{+}\sin\theta\cos\phi))$, da cui si conclude, ricordando due formule di trigonometria (formule di somma).

La *formula di Eulero* pone

$$e^{i\theta} = \cos\theta + i\sin\theta.$$

Naturalmente l'uso della lettera e richiama l'esponenziale, e la notazione è giustificata dal fatto che, come implica la Proposizione 1.2.2, vale la formula:

$$e^{i(\theta+\phi)} = e^{i\theta}e^{i\phi}.$$

Questa notazione permette di scrivere una delle formule che i matematici considerano tra le più belle ed eleganti che siano mai state scritte:

$$e^{i\pi} + 1 = 0.$$

Riassumendo quanto abbiamo visto, un numero complesso z si può scrivere nelle forme seguenti:

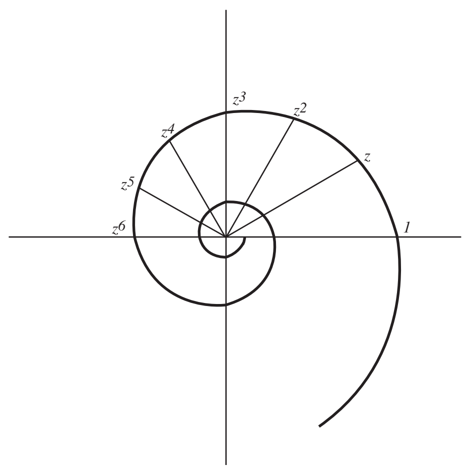
- in forma algebrica: $z = a + ib$;
- in forma polare: $z = \rho(\cos\theta + i \sin\theta)$;
- in forma esponenziale: $z = \rho e^{i\theta}$.

Esempio 1.2.3 $-2i = 2(\cos \frac{3}{2}\pi + i \sin \frac{3}{2}\pi) = 2e^{i\frac{3}{2}\pi}$.

La Proposizione 1.2.2 mostra quanto sia comodo, per fare il prodotto di due numeri complessi, usare la loro forma polare. Estendendo il risultato al quoziente, si ha che il quoziente di due numeri complessi (di cui il secondo diverso da zero) è il numero complesso avente come modulo il quoziente dei moduli e come argomento la differenza dei due argomenti. Inoltre, applicandola in maniera iterata, fornisce un modo molto semplice e rapido per calcolare le potenze di un numero complesso. Si ottiene, con calcolo immediato, la seguente importante *formula di De Moivre*, valida per ogni $n \in \mathbb{N}$:

Proposizione 1.2.3 *Se* $z = \rho(\cos\theta + i\sin\theta)$, *allora:*

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



$z = a + ib$, e le sue potenze, notare $|z| < 1$
Figura 1.3

Definizione 1.2.5 *Dato* $n \in \mathbb{N}$, *si chiama* radice ennesima *di un numero complesso* z *ogni numero complesso* w *tale che* $w^n = z$.

Val la pena osservare che qui il termine radice ennesima non è usato nello stesso senso di quando si considerano funzioni reali di variabile reale (vedi corso di Analisi): in particolare qui la radice ennesima è un insieme di numeri, quindi non è una funzione.

Vediamo ora come risolvere il problema di trovare le soluzioni dell'equazione $w^n = z$, dove z è un numero complesso dato, n un numero naturale positivo, e w indica l'incognita complessa.

Sia $z = \rho(\cos\theta + i\sin\theta)$, cerchiamo ogni $w = r(\cos\phi + i\sin\phi)$ che soddisfi l'equazione $w^n = z$. Si ha

$$r^n(\cos n\phi + i\sin n\phi) = \rho(\cos\theta + i\sin\theta),$$

da cui si ricava, *necessariamente*,

$$r = \sqrt[n]{\rho}, \quad n\phi = \theta + 2k\pi,$$

con k intero. Tenendo conto delle soluzioni che coincidono, si ottiene la seguente proposizione:

Proposizione 1.2.4 *Dato $0 \neq z = \rho(\cos\theta + i\sin\theta)$, le soluzioni dell'equazione $w^n = z$ sono n e date dalla formula:*

$$\sqrt[n]{\rho}\left(\cos\frac{\theta+2k\pi}{n} + i\sin\frac{\theta+2k\pi}{n}\right), \qquad k = 0, 1, \ldots, n-1.$$

Le radici ennesime di un numero complesso giacciono quindi sulla stessa circonferenza, di raggio $\sqrt[n]{\rho}$. Trovata la prima, di argomento $\frac{\theta}{n}$, si ottiene la seconda aggiungendo all'argomento, in senso antiorario, un angolo di ampiezza $\frac{2\pi}{n}$. La terza allo stesso modo a partire dalla seconda, e così via.

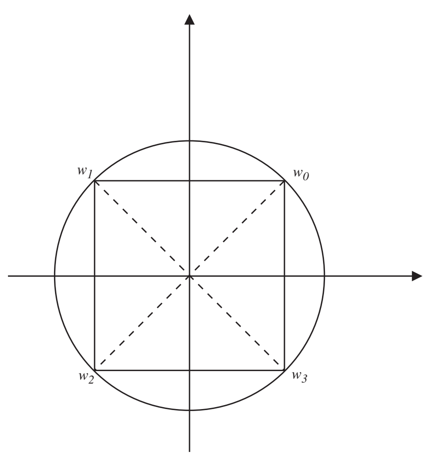
Esempio 1.2.4 Trovare le radici quarte di $z = -4$. Si ha

$$-4 = 4(\cos\pi + i\sin\pi),$$

quindi

$$w_0 = \sqrt{2}(\cos\frac{\pi}{4} + i\sin\frac{\pi}{4}) = 1 + i.$$

Le altre radici si ottengono con rotazioni successive di $\frac{\pi}{2}$. Si ottiene allora $w_1 = -1+i$, $w_2 = -1-i$, $w_3 = 1-i$.



Radici quarte di $z = -4$

Figura 1.4

1.3 Teorema fondamentale dell'algebra

Come accennato nell'introduzione, alla base della teoria dei numeri complessi c'è la necessità di poter risolvere delle equazioni polinomiali che nei reali non hanno soluzione. Questa è una cosa molto utile, anche nel caso in cui ci interessano solo soluzioni reali. In questo paragrafo vediamo i concetti essenziali per arrivare a formulare il teorema fondamentale dell'algebra, uno dei risultati più "belli" di tutta la matematica.

Il polinomio $x^3 - 3x^2 + 2x$ si annulla in $x = 0$, $x = 1$, $x = 2$. Per questo vale l'eguaglianza:

$$x^3 - 3x^2 + 2x = x(x-1)(x-2),$$

(fare la verifica).

In generale, scomporre il polinomio $P(x)$ significa esattamente scriverlo nella forma:

$$P(x) = (x - x_1)(x - x_2)\dots(x - x_n),$$

con $x_1, \dots x_n$ radici del polinomio (cioè soluzioni dell'equazione $P(x) = 0$). Occorre osservare che le radici non sono necessariamente distinte: il polinomio $x^2 - 2x + 1$ ha radice 1, che compare due volte ($x^2 - 2x + 1 = (x-1)(x-1) = (x-1)^2$).

Chiamiamo allora *molteplicità* della radice x_i del polinomio il numero naturale m tale che $P(x)$ è divisibile per $(x - x_i)^m$ ma non per $(x - x_i)^{m+1}$.

Ad esempio: $z^3 - 4z^2 + 5z - 2 = (z-1)^2(z-2)$, e quindi $P(z) = z^3 - 4z^2 + 5z - 2$ ammette 2 come radice semplice, 1 come radice doppia: in tutto 3 radici se 1 viene contato due volte, essendo due la sua molteplicità.

Se di sopra avessimo scritto $P(x)$, invece di $P(z)$, intendendo così considerare il polinomio sui reali, avremmo potuto ugualmente fare la scomposizione. Ma non tutti i polinomi si possono scomporre, se ci si limita a considerare radici reali: abbiamo motivato l'introduzione dei complessi anche per trovare una soluzione dell'equazione $x^2 + 1 = 0$.

Ed ecco l'enunciato del teorema fondamentale dell'algebra.

Teorema 1.3.1 *Un polinomio $P(z)$ nei complessi ammette n soluzioni, se ognuna di esse viene contata con la sua molteplicità.*

In altre parole, dato il polinomio a coefficienti complessi e nella variabile z, $P(z)$, di grado n, esistono $z_1, \dots, z_m \in \mathbb{C}$, tali che $z_i \neq z_j$ per $i \neq j$, e $m_1, \dots, m_r > 0$ tali che

$$P(z) = (z - z_1)^{m_1}(z - z_2)^{m_2}\dots(z - z_r)^{m_r}, \qquad m_1 + m_2 + \dots m_r = n.$$

Risultato straordinario, che da solo spiega l'importanza dell'insieme $\mathbb{C}$.

Esempio 1.3.1 Il polinomio

$$P(z) = z^4 - (2+i)z^3 + (1+2i)z^2 - iz$$

ammette radici $z = 0$, $z = i$, $z = 1$, quest'ultima con molteplicità due. Pertanto

$$P(z) = z(z-i)(z-1)^2.$$

Per concludere, facciamo un primo esempio (altri li vedremo in seguito) di come il teorema fondamentale dell'algebra può fornire risultati importanti anche nel caso in cui si voglia lavorare esclusivamente nel campo dei reali.
Mostriamo che ogni polinomio, di grado dispari e a coefficienti reali, ammette almeno una radice reale[10]. Per fare questo, cominciamo a dimostrare un utile risultato ausiliario.

Proposizione 1.3.1 *Un polinomio a coefficienti reali, se ha la radice complessa z_0, ha anche come radice il coniugato di z_0.*

Dimostrazione. Sia $a_n z^n + \cdots + a_1 z + a_0$ un polinomio con a_i numeri reali. Se

$$a_n z_0^n + \cdots + a_1 z_0 + a_0 = 0,$$

allora

$$\overline{a_n z_0^n + \cdots + a_1 z_0 + a_0} = \overline{0},$$

il che implica

$$a_n \bar{z}_0^n + \cdots + a_1 \bar{z}_0 + a_0 = 0. \quad \blacksquare$$

Ora possiamo facilmente provare quanto prima asserito. Dato un polinomio di grado n dispari e a coefficienti reali, consideriamo le sue n soluzioni nei complessi e, grazie alla Proposizione 1.3.1, accoppiamo a due a due una soluzione e la sua complessa coniugata. Essendo dispari, almeno una *deve* essere accoppiata con se stessa, e quindi deve essere reale, visto che se un numero complesso coincide col suo coniugato allora è necessariamente reale.

1.4 Esercizi

1.4.1 Quesiti a risposta chiusa

Rispondere vero o falso.

1. $1 - 2i$ è nel primo quadrante;

[10] Questo risultato può anche essere ottenuto con tecniche diverse, vedi corso di Analisi, che non richiedono l'uso dei numeri complessi. Ciò non toglie che la seguente semplice dimostrazione, basata sul teorema fondamentale dell'algebra, sia elegante ed interessante

2. $(i^3)^8 = -1$;
3. Sia $w = z^3$. Allora $|z| < |w|$;
4. L'equazione $z = \bar{z}$ ha infinite soluzioni;
5. $4\pi \in Arg(z\bar{z})$;
6. $\left(\frac{\sqrt{2}}{2} + i\frac{\sqrt{2}}{2}\right)^2 = i$;
7. $(\cos\theta - i\sin\theta)^7 = \cos(-7\theta) + i\sin(-7\theta)$;
8. $\text{Im } z - 1 = 0$ è una retta orizzontale del piano complesso;
9. $z \cdot \bar{z} = z$ ha solo soluzioni reali;
10. $|z| - 6 \text{ Re } z + i = 0$ ammette due soluzioni distinte;
11. L'insieme degli z tali che $z^2 \in \mathbb{R}$ è una coppia di rette perpendicolari;
12. Le soluzioni di $z^4 + 8z^2 + 16 = 0$ sono $2i$ e $-2i$;
13. Sia $z = \frac{i}{(1-i)^2}$. Allora: $|z| = \frac{1}{2}$;
14. Sia $z = \frac{i}{(1-i)^2}$. Allora: $\text{Arg } z = \pi$;
15. Sia $z = \frac{i}{(1-i)^2}$. Allora: $\bar{z} = \frac{-i}{(1-i)^2}$;
16. $\text{Re } \frac{1}{(2-3i)^2} = \frac{-5}{169}$.

1.4.2 Esercizi aperti

Esercizio 1.4.1 Verificare le relazioni 1,2,4 di pagina 7

Esercizio 1.4.2 Provare la Proposizione 1.2.1

Esercizio 1.4.3 Verificare le seguenti formule:

- $|\text{ Re } z| \le |z|$, $|\text{ Im } z| \le |z|$;
- $|z + w| \le |z| + |w|$;
- $|zw| = |z||w|$;
- $\left|\frac{1}{z}\right| = \frac{1}{|z|}$.

Esercizio 1.4.4 Verificare che:

- $\overline{z + w} = \bar{z} + \bar{w}$;
- $\overline{z \cdot w} = \bar{z} \cdot \bar{w}$;
- $\overline{\left(\frac{1}{z}\right)} = \frac{1}{\bar{z}}$.

Esercizio 1.4.5 Calcolare $(\frac{\sqrt{3}-i}{2})^6$.

Esercizio 1.4.6 Scrivere in forma algebrica $(1-i)^8$, $(1+i)^{12}$.

Esercizio 1.4.7 Sia z tale che $z \cdot \bar{z} = 1$. Calcolare

$$|1 + z|^2 + |1 - z|^2.$$

Esercizio 1.4.8 Trovare le radici della forma ib del polinomio $P(z) = z^4 - 2z^3 + 7z^2 - 4z + 10$. Trovare tutte le radici di $P(z)$.

Esercizio 1.4.9 Disegnare nel piano l'insieme dei punti z che verificano l'equazione $z^2 + \overline{z}^2 = 0$.

Esercizio 1.4.10 Scomporre $P(z) = z^5 + 6z^4 + 13z^3 + 14z^2 + 12z + 8$, sapendo che -2 è radice di molteplicità tre.

Esercizio 1.4.11 Sia A il seguente sottoinsieme dei numeri interi:

$$A = \{u \in \mathbb{Z} : u = m^2 + n^2,\ m, n \in \mathbb{Z}\}.$$

Dimostrare che se $u, v \in A$ allora $u \cdot v \in A$.
Osservare che $u \in A$ implica che esistono $z \in \mathbb{C}$ tale che $u = |z|^2$, e usare la formula di De Moivre.

Esercizio 1.4.12 Risolvere $z|z| = -\bar{z}$.

Esercizio 1.4.13 Disegnare sul piano l'insieme delle soluzioni del seguente sistema:

$$\begin{cases} \operatorname{Re} z \cdot \operatorname{Im} z < 0 \\ |z-1| = |z+i| \end{cases} .$$

Esercizio 1.4.14 Risolvere $(z-2)^3 = -i$.

Esercizio 1.4.15 Risolvere $2z\bar{z} + 2z^2 = 1 - i$.

Esercizio 1.4.16 Risolvere $z^8 - 3z^4 + 2 = 0$.

Esercizio 1.4.17 Trovare un polinomio a coefficienti reali di quinto grado che abbia fra le sue soluzioni i e $1 - i$.

Esercizio 1.4.18 Trovare un polinomio che abbia come radici soltanto 3 e $1 + i$. Trovarne uno a coefficienti reali, se possibile.

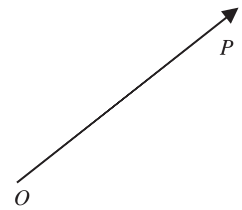
2

Vettori, rette e piani

In questo capitolo introduciamo dapprima i vettori, poi le rette ed i piani. I vettori sono uno strumento utilissimo, che viene usato in questo corso innanzi tutto per questioni legate alla geometria delle rette e dei piani, ma che risulta indispensabile anche in altri campi come, ad esempio, la Fisica. Come spesso accade, purtroppo la loro presentazione formale risulta abbastanza pesante e noiosa; per alleviare le pene, e soprattutto per capire meglio, chi legge dovrebbe continuamente pensare al significato delle proprietà enunciate, che spesso hanno una loro naturale interpretazione. Citiamo, una per tutte, la definizione di somma tra due vettori.

Per introdurre l'idea di vettore, potrebbe essere utile far ricorso all'esempio seguente. Supponiamo di avere un punto che si muova su un piano, e di voler riportare su un foglio, che rappresenta il piano, il punto stesso, "fotografato" in certi istanti. È allora utile avere uno strumento che ci permetta di rappresentare il movimento del punto anche tramite operazioni geometriche. La prima cosa da fare è allora quella di fissare un punto del piano, chiamato convenzionalmente O, e che si potrebbe considerare il punto in cui è posto l'osservatore. Il punto P allora, in un generico istante, rappresenta idealmente l'estremo finale di un segmento il cui punto iniziale è l'osservatore stesso. Per orientare tale segmento si usa spesso una freccia con la punta in P. Questa è l'idea di vettore. Le operazioni che poi definiremo fra vettori ci permetteranno, tra le altre cose, di descrivere il suo movimento.

Definizione 2.0.1 *Un* vettore geometrico *applicato nel punto O del piano (o dello spazio) è un segmento orientato avente come primo estremo O e come punto finale un punto P del piano (o dello spazio).*



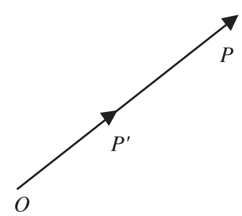
Vettore geometrico
Figura 2.1

Nel caso (degenere, ma importante) in cui $P = O$, si parla di *vettore nullo*.

Notazione Ci sono molte notazioni utilizzate per indicare i vettori. Nel caso si voglia mettere in evidenza che si parla di vettori geometrici, in genere si usano le notazioni $\overrightarrow{OP}$ oppure $P - O$. Ma nel corso di queste note ne saranno utilizzate anche altre.

Un vettore non nullo è caratterizzato da tre proprietà:

1) Intensità: è la lunghezza del segmento orientato; si indica con $|OP|$ oppure $\|OP\|$; si legge *modulo di* OP;
2) Direzione: è la retta su cui è applicato il vettore;
3) Verso: due vettori OP ed OP' hanno lo stesso verso se P, P' giacciono sulla stessa semiretta rispetto ad O, verso opposto se giacciono su due semirette opposte.

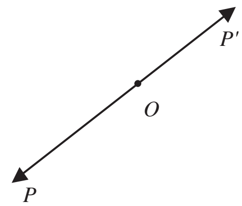


Vettori che hanno stessa direzione e stesso verso
Figura 2.2

Il vettore nullo ha lunghezza zero e, per convenzione (ragionevole), non ha né direzione né verso.

A volte non è importante specificare il punto di applicazione, o si possono considerare vettori con punti di applicazione differenti: si parla in questo caso di *vettore libero*.

Osservazione 2.0.1 Due vettori liberi hanno la stessa direzione quando giacciono sulla stessa retta o su rette parallele non coincidenti.



Vettori che hanno stessa direzione ma verso opposto
Figura 2.3

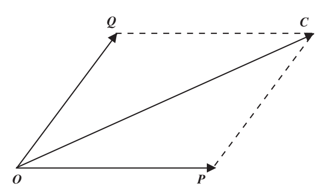
Come esempio di grandezze che sono vettori, possiamo pensare in Fisica alla forza che agisce su un oggetto, e di cui bisogna conoscere il punto di applicazione, l'intensità, in che direzione agisce, per poter conoscere i suoi effetti sull'oggetto stesso.

2.1 Operazioni sui vettori

In questo paragrafo si introducono le principali operazioni fra vettori, e se ne elencano le loro proprietà: si tratta delle operazioni di somma e differenza fra vettori, e del prodotto di un vettore per uno scalare.

2.1.1 Somma

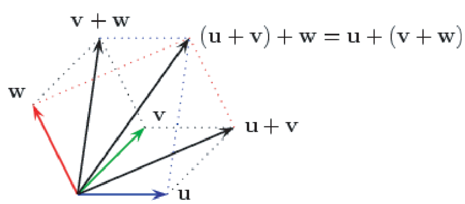
Consideriamo due vettori $\overrightarrow{OP}$ ed $\overrightarrow{OQ}$ non allineati e con lo stesso punto di applicazione O. La somma $\overrightarrow{OC} = \overrightarrow{OP} + \overrightarrow{OQ}$, si definisce mediante la regola del parallelogramma. Da Q si traccia la parallela ad $\overrightarrow{OP}$, da P la parallela ad $\overrightarrow{OQ}$. Queste due rette si intersecano in un punto C che è il quarto vertice del parallelogramma $OQCP$. La diagonale $\overrightarrow{OC}$ rappresenta la somma dei due vettori.



La somma di due vettori
Figura 2.4

Supponiamo di muoverci lungo un piano. Partiamo da un punto O e spostiamoci (in linea retta) fino al punto P, poi da P andiamo in C. Il tratto percorso PC, dal punto di vista dell'osservatore che sta in O, è rappresentato dal vettore, applicato nell'origine, $\overrightarrow{OQ}$. Con la regola di somma ora definita, il punto C di arrivo della nostra passeggiata è descritto da $\overrightarrow{OC} = \overrightarrow{OP} + \overrightarrow{OQ}$. Dunque i vettori, e l'operazione di somma definita, sono uno strumento efficace per descrivere il moto. Tra l'altro, questo spiega perché operativamente utilizzare la scrittura equivalente $\overrightarrow{OP} + \overrightarrow{PC} = (P - O) + (C - P) = C - O = \overrightarrow{OC}$.

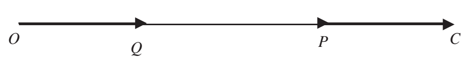
Appare geometricamente evidente che $\overrightarrow{OP} + \overrightarrow{OQ} = \overrightarrow{OQ} + \overrightarrow{OP}$, e cioè che la somma di vettori gode della proprietà commutativa. Inoltre, è chiaro che, una volta definita la somma fra due vettori, si può fare anche la somma di un numero di vettori maggiore di due, sommandone prima due, ed al vettore così ottenuto sommare il terzo. Come appare evidente, si ha che la somma di tre vettori non dipende dall'ordine in cui i vettori si sommano: $(\overrightarrow{OP} + \overrightarrow{OQ}) + \overrightarrow{OR} = \overrightarrow{OP} + (\overrightarrow{OQ} + \overrightarrow{OR})$. Quindi nel seguito le parentesi di sopra non saranno più messe, perché inutili; inoltre, una volta capito questo su tre vettori, è evidente che sappiamo come sommarne un numero finito qualunque.



La legge associativa

Figura 2.5

Nel caso particolare in cui i punti O, P, Q sono allineati ed $\overrightarrow{OP}$ e $\overrightarrow{OQ}$ hanno lo stesso verso, allora $\overrightarrow{OC} = \overrightarrow{OP} + \overrightarrow{OQ}$ ed $\overrightarrow{OC}$ ha la stessa direzione e verso di $\overrightarrow{OP}$ e $\overrightarrow{OQ}$. Il modulo di $\overrightarrow{OC}$ è dato dalla somma dei moduli dei due vettori:



Somma di due vettori che hanno stessa direzione e stesso verso: $\overrightarrow{OP} + \overrightarrow{OQ} = \overrightarrow{OC}$

Figura 2.6

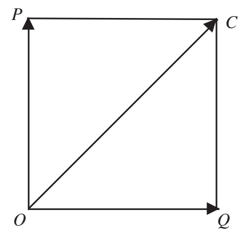
Se i punti O,P,Q sono allineati ed $\overrightarrow{OP}$ e $\overrightarrow{OQ}$ hanno verso opposto, allora $\overrightarrow{OC} = \overrightarrow{OP} + \overrightarrow{OQ}$ ed $\overrightarrow{OC}$ ha la stessa direzione e verso del maggiore tra i due vettori. Il modulo di $\overrightarrow{OC}$ è dato dalla differenza dei moduli dei due vettori :

Somma di due vettori che hanno stessa direzione e verso opposto: $\overrightarrow{OP} + \overrightarrow{OQ} = \overrightarrow{OC}$

Figura 2.7

Osserviamo che in generale $\|\overrightarrow{OC}\| \neq \|\overrightarrow{OP}\| + ||\overrightarrow{OQ}\|$. L'uguaglianza vale solo se i vettori giacciono sulla stessa retta ed hanno lo stesso verso.

Esempio 2.1.1 Consideriamo i vettori: $\overrightarrow{OP}$, $\overrightarrow{OQ}$ tali che $\|\overrightarrow{OP}\| = 4$, $\|\overrightarrow{OQ}\| = 3$ e direzioni perpendicolari tra loro:



La somma di due vettori perpendicolari
Figura 2.8

Si ha $\|OC\| = \sqrt{16+9} = 5$.

2.1.2 Prodotto di un vettore per uno scalare

Definizione 2.1.1 *Sia* $\overrightarrow{OP}$ *un vettore e* t *un numero reale. Si chiama* prodotto del vettore $\overrightarrow{OP}$ per lo scalare t, *il vettore* $t \cdot \overrightarrow{OP}$, *che ha come punto di applicazione* O, *tale che* $\|t \cdot \overrightarrow{OP}\| = |t| \cdot \|\overrightarrow{OP}\|$, *che ha stessa direzione di* $\overrightarrow{OP}$, *e stesso verso se* $t > 0$, *verso opposto se* $t < 0$.

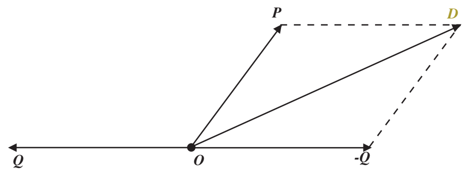
Il caso $t = -1$ è interessante: scriveremo, come è naturale, $(-1)\overrightarrow{OP} = -\overrightarrow{OP}$ e chiameremo $-\overrightarrow{OP}$ l'*opposto* di $\overrightarrow{OP}$. Osserviamo che $-\overrightarrow{OP}$ ha stesso modulo, stessa direzione e verso opposto rispetto ad $\overrightarrow{OP}$. Inoltre $-\overrightarrow{OP} + \overrightarrow{OP} = \overrightarrow{OO}$.

Al variare di t nei reali, il vettore $t \cdot \overrightarrow{OP}$ descrive una retta: se $t > 0$ descrive la semiretta di origine O, e direzione e verso del vettore $\overrightarrow{OP}$, se $t < 0$ descrive la semiretta opposta.

Il prodotto di un vettore per uno scalare gode di proprietà naturali, che sono riportate nella sezione degli approfondimenti.

2.1.3 Differenza

La differenza fra vettori, come nel caso dei numeri reali, in realtà non è un'operazione nuova, perché si riconduce a quella di somma, tramite l'uso dei vettori opposti. Cioè si definisce: $\overrightarrow{OP} - \overrightarrow{OQ} = \overrightarrow{OP} + (-\overrightarrow{OQ})$. Quindi il vettore $\overrightarrow{OD} = \overrightarrow{OP} - \overrightarrow{OQ}$ si ottiene con la regola del parallelogramma applicata ai due vettori $\overrightarrow{OP} + (-\overrightarrow{OQ})$:

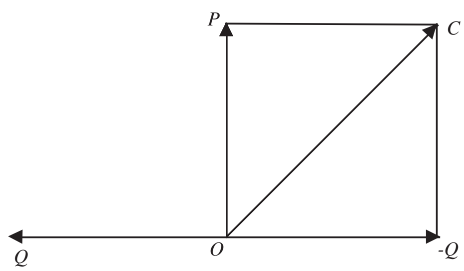


La differenza fra i vettori $\overrightarrow{OP}$ e $\overrightarrow{OQ}$
Figura 2.9

Per l'operazione differenza valgono le seguenti semplici proprietà:

- $\overrightarrow{OP} = \overrightarrow{OQ} + \overrightarrow{OR} \Longleftrightarrow \overrightarrow{OP} - \overrightarrow{OR} = \overrightarrow{OQ}$;
- Legge di cancellazione: $\overrightarrow{OP} + \overrightarrow{OR} = \overrightarrow{OQ} + \overrightarrow{OR} \Longrightarrow \overrightarrow{OP} = \overrightarrow{OQ}$.

Esempio 2.1.2 Consideriamo due vettori $\overrightarrow{OP}$ ed $\overrightarrow{OQ}$ tali che $\|\overrightarrow{OP}\| = 4$, $\|\overrightarrow{OQ}\| = 3$ e direzioni perpendicolari tra loro. La differenza $\overrightarrow{OC} = \overrightarrow{OP} - \overrightarrow{OQ}$, è mostrata nel disegno in figura:



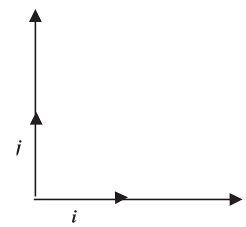
Differenza tra vettori perpendicolari: $\overrightarrow{OP} - \overrightarrow{OQ} = \overrightarrow{OC}$
Figura 2.10

Notiamo che i moduli dei vettori differenza e somma di $\overrightarrow{OP}$ e $\overrightarrow{OQ}$ sono uguali perché rappresentano le lunghezze delle due diagonali del rettangolo $OQCP$.

2.1.4 Componenti cartesiane di un vettore

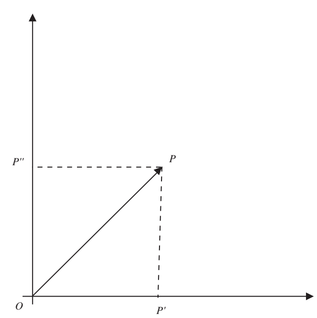
Fissiamo nel piano due rette perpendicolari, una orizzontale una verticale: l'*asse x (o asse delle ascisse)* è la retta orizzontale, e l'*asse y (o asse delle ordinate)* quella verticale; chiamiamo O il punto intersezione e fissiamo sulle rette un' unità di misura ed un verso di percorrenza. Convenzionalmente il verso positivo dell'asse x è verso destra, quello dell'asse y verso l'alto. Questa operazione ha il significato di fissare un sistema di riferimento, detto *cartesiano*, nel piano. Il punto O è quello di applicazione dei vettori, cioè il punto in cui si immagina l'osservatore, ed è quindi naturale porre nello stesso punto l'intersezione fra gli assi.

Definizione 2.1.2 *Si chiamano* versori degli assi *i due vettori (indicati rispettivamente con* $\overrightarrow{i}$ *per l'asse* x*,* $\overrightarrow{j}$ *per l'asse* y*) di modulo unitario, direzione coincidente con quella degli assi cartesiani e verso nel senso positivo degli assi stessi (*base canonica *del piano)*[1].



I versori degli assi
Figura 2.11

Dato il vettore $\overrightarrow{OP}$, chiamiamo $\overrightarrow{OP'}$ e $\overrightarrow{OP''}$ le proiezioni di $\overrightarrow{OP}$ sugli assi. Si ha allora che $\overrightarrow{OP} = \overrightarrow{OP'} + \overrightarrow{OP''}$. $\overrightarrow{OP'}$ è un vettore della forma $x\,\overrightarrow{i}$, $\overrightarrow{OP''}$ è un vettore della forma $y\,\overrightarrow{j}$, quindi $\overrightarrow{OP} = x\,\overrightarrow{i} + y\,\overrightarrow{j}$. A volte scriveremo $\overrightarrow{OP}(x, y)$, o anche $P(x, y)$ e x, y sono dette le *componenti cartesiane* di $\overrightarrow{OP}$.



Le componenti cartesiane del vettore $\overrightarrow{OP}$
Figura 2.12

2.1.5 Operazioni tra vettori utilizzando le componenti scalari

Esprimere un vettore tramite le sue componenti scalari è estremamente utile per fare operazioni fra vettori. Vediamo qualche esempio.

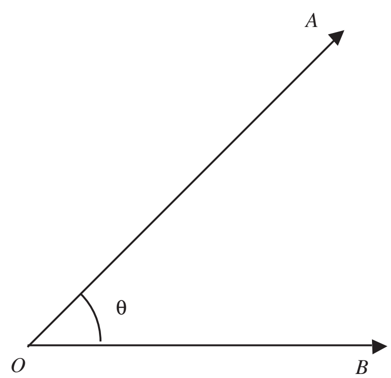
Siano $\overrightarrow{v_1} = \overrightarrow{OP_1} = x_1\,\overrightarrow{i} + y_1\,\overrightarrow{j}$, $\overrightarrow{v_2} = \overrightarrow{OP_2} = x_2\,\overrightarrow{i} + y_2\,\overrightarrow{j}$.

[1] Ecco una nuova notazione per vettori: se non abbiamo particolare interesse a mettere in evidenza il punto di applicazione O, indichiamo un generico vettore semplicemente con una lettera sormontata da una freccia. Useremo anche il carattere in grassetto e dunque, per indicare un vettore, abbiamo come minimo le notazioni $\overrightarrow{OP}, P - O, \overrightarrow{w}, \mathbf{w}$.

1. Il modulo (lunghezza, norma) del vettore $\overrightarrow{v_1}$ è $|\overrightarrow{v_1}| = \sqrt{x_1^2 + y_1^2}$ (segue dal teorema di Pitagora);
2. La somma dei due vettori: $\overrightarrow{v_1}+\overrightarrow{v_2} = (x_1+x_2)\overrightarrow{i}+(y_1+y_2)\overrightarrow{j}$, ha come componenti cartesiane la somma delle componenti dei due vettori dati (la dimostrazione, facile, si basa su uguaglianze fra triangoli: farsi una figura);
3. Il prodotto per uno scalare: $t\overrightarrow{v_1} = tx_1\overrightarrow{i} + ty_1\overrightarrow{j}$ (ovvio);
4. Dato un vettore $\overrightarrow{v} = a\overrightarrow{i} + b\overrightarrow{j}$, tutti i vettori paralleli a $\overrightarrow{v}$ sono del tipo $k\overrightarrow{v} = ka\overrightarrow{i} + kb\overrightarrow{j}$, con $k \neq 0$, ed hanno lo stesso verso di $\overrightarrow{v}$ se $k > 0$.

2.1.6 Prodotto scalare tra vettori

Introduciamo ora una nuova operazione tra vettori: il prodotto scalare. La parola *scalare* sta ad indicare che il risultato di tale operazione è un numero reale: in questo contesto i numeri reali vengono appunto chiamati scalari. Va osservato che questa è un'operazione fra vettori che come risultato *non* produce un vettore: di conseguenza non ha senso, ad esempio, fare il prodotto scalare fra tre o più vettori. Nel seguito introdurremo un altro tipo di prodotto fra vettori, il prodotto vettoriale, così chiamato perché il risultato in tal caso è appunto un vettore.
Siano dunque $\overrightarrow{a} = \overrightarrow{OA}$, $\overrightarrow{b} = \overrightarrow{OB}$ due vettori non nulli applicati in O e sia θ l'angolo $\widehat{AOB}$ formato tra i due vettori:



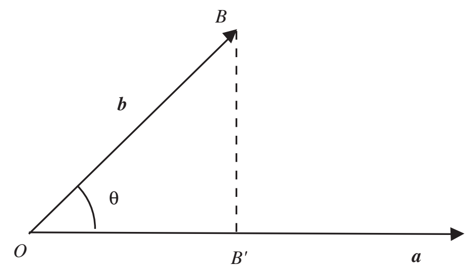
Angolo tra due vettori
Figura 2.13

Definizione 2.1.3 Si definisce *prodotto scalare* dei due vettori $\overrightarrow{a}$, $\overrightarrow{b}$, e si indica con $\overrightarrow{a} \cdot \overrightarrow{b}$, il numero reale

$$\overrightarrow{a} \cdot \overrightarrow{b} = |\overrightarrow{a}| \cdot |\overrightarrow{b}| \cos\theta.$$

Dalla definizione segue che il prodotto scalare è un numero positivo se l'angolo θ è acuto ($\cos\theta > 0$), negativo se θ è ottuso ($\cos\theta < 0$).

Il prodotto scalare, nel caso in cui l'angolo fra i vettori sia minore di un angolo retto, si può anche interpretare come il prodotto della lunghezza di un vettore (ad esempio $\overrightarrow{a}$) per la lunghezza della proiezione del secondo vettore ($\overrightarrow{b}$) sul primo (nel disegno la proiezione è indicata con OB'):



$$|\overrightarrow{OB'}| = |\overrightarrow{b}|\cos\theta$$

Figura 2.14

Nel caso invece in cui l'angolo fra i due vettori è maggiore di un angolo retto, vale la stessa interpretazione, pur di tenere conto del cambiamento di segno.

Il prodotto scalare fra due vettori $\overrightarrow{v_1} = x_1\overrightarrow{i} + y_1\overrightarrow{j}$, $\overrightarrow{v_2} = x_2\overrightarrow{i} + y_2\overrightarrow{j}$ si esprime in maniera semplice, utilizzando le loro componenti scalari. Si ha infatti che:

$$\overrightarrow{v_1} \cdot \overrightarrow{v_2} = x_1x_2 + y_1y_2. \tag{2.1}$$

La dimostrazione di questo fatto si fa così: si considera il triangolo (vedi Figura 2.15) di lati $\overrightarrow{v_1}$, $\overrightarrow{v_2}$ e $\overrightarrow{v_1} - \overrightarrow{v_2}$. Il teorema dei coseni asserisce che

$$|\overrightarrow{v_1} - \overrightarrow{v_2}|^2 = |\overrightarrow{v_1}|^2 + |\overrightarrow{v_2}|^2 - 2|\overrightarrow{v_1}||\overrightarrow{v_2}|\cos\theta = |\overrightarrow{v_1}|^2 + |\overrightarrow{v_2}|^2 - 2\overrightarrow{v_1} \cdot \overrightarrow{v_2}, \tag{2.2}$$

dove al solito θ è l'angolo formato da $\overrightarrow{v_1}$ e $\overrightarrow{v_2}$. D'altra parte si ha anche che

$$|\overrightarrow{v_1} - \overrightarrow{v_2}|^2 = (x_1 - x_2)^2 + (y_1 - y_2)^2 = x_1^2 + x_2^2 + y_1^2 + y_2^2 - 2x_1x_2 - 2y_1y_2. \tag{2.3}$$

Dal confronto fra (2.2) e (2.3) si ottiene allora la (2.1)

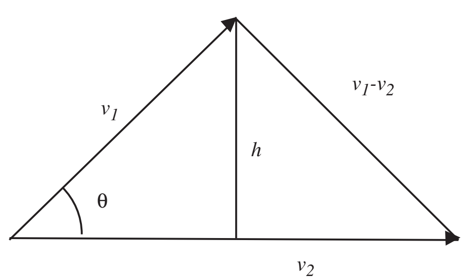


Figura 2.15

Osservazione 2.1.1 Osserviamo che:

1. Se $\overrightarrow{a} \cdot \overrightarrow{b} = 0$, allora o almeno uno dei due vettori è nullo oppure $\cos\theta = 0$. Geometricamente quest'ultimo caso corrisponde al fatto che i due vettori sono *perpendicolari.*
2. I vettori perpendicolari a $\overrightarrow{0} \neq \overrightarrow{v} = a\overrightarrow{i} + b\overrightarrow{j}$ sono i vettori $\overrightarrow{w} = -kb\overrightarrow{i} + ka\overrightarrow{j}$ (infatti in questo caso, e solo in questo caso, il prodotto scalare $\overrightarrow{v} \cdot \overrightarrow{w} = 0$). Il caso particolare $k = 1$ corrisponde al vettore $\overrightarrow{w} = -b\overrightarrow{i} + a\overrightarrow{j}$ che ha lo stesso modulo di $\overrightarrow{v}$ ed è ruotato di $\frac{\pi}{2}$ in verso antiorario;
3. $\overrightarrow{i} \cdot \overrightarrow{i} = 1, \overrightarrow{j} \cdot \overrightarrow{j} = 1$;
4. $\overrightarrow{i} \cdot \overrightarrow{j} = 0$.

Vediamo ora qualche esempio di calcolo con i vettori.

Esempio 2.1.3 Siano $\overrightarrow{a}$ e $\overrightarrow{b}$ due vettori tali che $|\overrightarrow{a}| = \sqrt{2}$ e $|\overrightarrow{b}| = 1$ e tali che $\overrightarrow{a} \cdot \overrightarrow{b} = 1$. Quanto vale l'angolo θ da essi formato?
Dalla definizione di prodotto scalare si ha che $\cos\theta = \frac{\overrightarrow{a} \cdot \overrightarrow{b}}{|\overrightarrow{a}||\overrightarrow{b}|}$, quindi $\cos\theta = \frac{1}{\sqrt{2}} \Rightarrow \theta = \frac{\pi}{4}$(oppure $\theta = -\frac{\pi}{4}$).

Esempio 2.1.4 Consideriamo i vettori AB e CD con $A(0,5)$, $B(-1,3)$, $C(-1,-2)$, $D(0,2)$.

1. Scriviamo i vettori nella forma $\overrightarrow{v} = x\overrightarrow{i} + y\overrightarrow{j}$.
   Le componenti cartesiane di $\overrightarrow{AB}$ sono $(-1,-2)$ ($\overrightarrow{AB} = (x_B - x_A)\overrightarrow{i} + (y_B - y_A)\overrightarrow{j}$), e quelle di $\overrightarrow{CD}$ sono $(1,4)$. Allora possiamo scrivere $\overrightarrow{AB} = -\overrightarrow{i} - 2\overrightarrow{j}$ e $\overrightarrow{CD} = \overrightarrow{i} + 4\overrightarrow{j}$.
2. Calcoliamo il modulo dei vettori.
   $\|\overrightarrow{AB}\| = \sqrt{1+4} = \sqrt{5}, \quad \|\overrightarrow{CD}\| = \sqrt{1+16} = \sqrt{17}$;
3. Calcoliamo l'angolo che essi formano con il semiasse positivo delle x.
   Se chiamiamo α l'angolo e ricordiamo che $\tan\alpha = \frac{y}{x}$, abbiamo: $\tan\alpha = 2$ per il vettore $\overrightarrow{AB}$ e $\tan\alpha = 4$ per il vettore $\overrightarrow{CD}$. Nel primo caso otteniamo l'angolo $\alpha \approx 180 + 63 = 243$ gradi (il vettore ha entrambe le componenti cartesiane negative quindi si trova nel terzo quadrante. L'angolo $\alpha \approx 63$ che si ottiene con una calcolatrice da $\tan\alpha = 2$ va ruotato di 180 gradi). Nel secondo caso otteniamo $\alpha \approx 76$ gradi.

Esempio 2.1.5 Calcoliamo l'ampiezza dell'angolo tra i vettori:

1. $\overrightarrow{v} = \overrightarrow{i} + 3\overrightarrow{j}$, $\overrightarrow{w} = -6\overrightarrow{j}$.
   Calcoliamo il prodotto scalare utilizzando le componenti cartesiane: $\overrightarrow{v} \cdot \overrightarrow{w} = 1 \cdot 0 + (3) \cdot (-6) = -18$. I moduli dei due vettori sono: $\|\overrightarrow{v}\| = \sqrt{10}$, $\|\overrightarrow{w}\| = 6$; utilizzando la formula $\cos\theta = \frac{\overrightarrow{v} \cdot \overrightarrow{w}}{|\overrightarrow{v}||\overrightarrow{w}|} = \frac{-18}{6\sqrt{10}}$ si ottiene $\theta \approx 162$ gradi;

2. $\overrightarrow{v} = \overrightarrow{i} + \overrightarrow{j}$, $\overrightarrow{w} = -5\overrightarrow{i} + 2\overrightarrow{j}$.
   Come nell'esercizio precedente: $\overrightarrow{v} \cdot \overrightarrow{w} = -5 \cdot 1 + 2 \cdot 1 = -3$. $\|\overrightarrow{v}\| = \sqrt{2}$; $\|\overrightarrow{w}\| = \sqrt{29}$; $\cos\theta = \frac{\overrightarrow{v} \cdot \overrightarrow{w}}{|\overrightarrow{v}||\overrightarrow{w}|} = \frac{-3}{\sqrt{2}\sqrt{29}} \Rightarrow \theta \approx 113$ gradi;
3. $\overrightarrow{v} = \overrightarrow{AB}$, con $A(-5,1)$ e $B(3,1)$, $\overrightarrow{w} = \overrightarrow{CD}$, con $C(2,-3)$ e $D(1,-3)$.
   $\overrightarrow{AB} = 8\overrightarrow{i}$, $\overrightarrow{CD} = -\overrightarrow{i}$. I due vettori sono opposti e quindi l'angolo compreso è 180 gradi.

2.2 Rette nel piano

Dato un punto $P_0(x_0, y_0)$ ed un vettore $\overrightarrow{n} = a\overrightarrow{i} + b\overrightarrow{j}$, esiste una ed una sola retta r passante per P_0 e perpendicolare ad $\overrightarrow{n}$. Sia P un punto $P(x,y) \in r$. Poiché il vettore $P - P_0$ è perpendicolare ad $\overrightarrow{n}$, si ha

$$(P - P_0) \cdot \overrightarrow{n} = 0, \tag{2.4}$$

cioè

$$a(x - x_0) + b(y - y_0) = 0. \tag{2.5}$$

Poiché le coordinate di P_0 verificano la formula, la (2.4) ci fornisce allora l'equazione, in forma vettoriale, della retta passante per P_0 e perpendicolare a n, mentre la sua forma scalare è data dalla (2.5).
Ponendo $c = -ax_0 - by_0$, la (2.5) può essere scritta nella forma:

$$ax + by + c = 0, \tag{2.6}$$

che si chiama *forma generale della retta*.

Affermare che la precedente equazione rappresenta una retta, lo ricordiamo, significa due cose:

- tutte le coppie (x, y) che soddisfano tale equazione giacciono su una retta r;
- tutti i punti (x, y) appartenenti alla retta r individuata dall'equazione (2.6) soddisfano l'equazione precedente.

L'equazione precedente rappresenta una retta se e solo se almeno uno fra i parametri a e b è diverso da zero. Se $b \neq 0$ allora la precedente equazione si può scrivere nella forma

$$y = mx + q,$$

(con $m = \frac{-a}{b}, q = \frac{-c}{b}$), che si chiama *forma canonica* della retta. La forma canonica permette di scrivere tutte le rette in forma di funzione del tipo $y = f(x)$. Le rette parallele all'asse y, che non sono funzioni (rispetto alla variabile indipendente x), non hanno rappresentazione in forma canonica, e sono del tipo $x = k$.

In sintesi:

- $ax + by + c = 0$ rappresenta l'equazione generale di una retta, se $a^2 + b^2 > 0$, inoltre (a, b) individua la direzione perpendicolare alla retta;
- $y = mx + q$ rappresenta l'equazione canonica di una retta;
- se $b \neq 0$ si può passare dalla forma generale a quella canonica, il viceversa è sempre possibile.

2.2.1 Equazione del fascio di rette per $P_0(x_0, y_0)$

Al variare di a e b la formula

$$a(x - x_0) + b(y - y_0) = 0$$

fornisce le equazioni di tutte le rette passanti per $P_0(x_0, y_0)$.
Per trovare una retta che passa per $P_0(x_0, y_0)$ e $P_1(x_1, y_1)$ basta considerare la generica retta $a(x - x_0) + b(y - y_0) = 0$ ed imporre il passaggio per P_1. Si ha allora

$$a(x - x_0) = -b(y - y_0) \tag{2.7}$$

e

$$a(x_1 - x_0) = -b(y_1 - y_0). \tag{2.8}$$

Dividendo (2.7) per (2.8), in modo da far sparire a e b, si ottiene la formula:

$$\frac{y - y_0}{y_1 - y_0} = \frac{x - x_0}{x_1 - x_0}.$$

La formula vale naturalmente se $y_1 \neq y_0$ e $x_1 \neq x_0$. Se $x_1 = x_0$ la retta da determinare è verticale e quindi è della forma $x = c$, cioè $x = x_1$. Se $y_1 = y_0$ l'equazione della retta è $y = y_0$.

2.2.2 Rette in forma parametrica

Abbiamo visto, nel paragrafo precedente, come determinare l'equazione della retta passante per un punto e perpendicolare ad una direzione data. Un problema analogo è quello di scrivere l'equazione di una retta passante per un punto e parallela alla direzione di un vettore dato[2]. Se $P(x, y)$ rappresenta un generico punto di tale retta, e $P_0(x_0, y_0)$, $P_1(x_1, y_1)$ ne sono due punti fissati, che quindi ne determinano la direzione, allora $P - P_0$ e $P_1 - P_0$ devono avere la stessa direzione, per cui deve esistere t tale che $P - P_0 = t(P_1 - P_0)$.
Indicando ora con $Q(a, b)$ una generica direzione della retta, abbiamo allora la rappresentazione, in forma vettoriale, della retta passante per P_0 e con direzione Q:

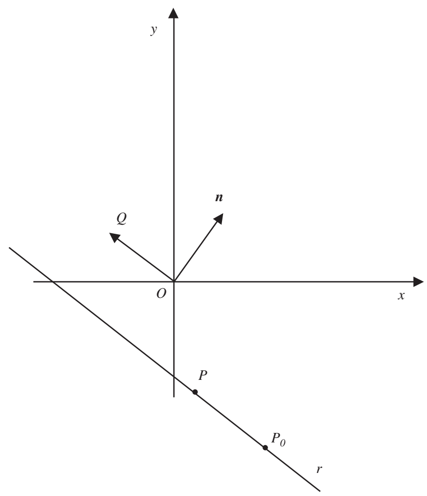
$$P = P_0 + tQ, \tag{2.9}$$

[2] Chiameremo anche direzione di una retta r un vettore qualsiasi che abbia direzione parallela alla retta r

che si scrive in forma scalare nel modo seguente:

$$\begin{cases} x = x_0 + at \\ y = y_0 + bt \end{cases} . \tag{2.10}$$

Queste sono dette le *rappresentazioni parametriche* della retta: il nome si giustifica con la presenza del parametro t nella sua definizione.



Direzione della retta e direzione ortogonale
Figura 2.16

Abbiamo dunque visto che una retta qualunque può essere rappresentata sia in forma parametrica sia in forma generale: vediamo ora in alcuni esercizi come passare dall'una all'altra.

Esempio 2.2.1 Scrivere in forma parametrica ed in forma generale l'equazione della retta passante per $P_0(2,1)$ ed avente come vettore direzione il vettore $\overrightarrow{v} = (-2,3)$. L'equazione in forma parametrica è:

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

Per passare alla forma generale possiamo procedere in due modi:

1. Ricaviamo t dalla prima equazione e sostituiamolo nella seconda:

$$\begin{cases} t = \frac{2-x}{2} \\ y = 1 + 3(\frac{2-x}{2}) \end{cases} ,$$

dalla quale si ottiene $2y + 3x - 8 = 0$.

2. Calcoliamo le componenti cartesiane di un vettore $\overrightarrow{w}$ perpendicolare a $\overrightarrow{v} = (-2, 3)$. Ricordiamo che se le componenti cartesiane di $\overrightarrow{v}$ sono (a, b) quelle di $\overrightarrow{w}$ sono ad esempio $(-b, a)$, quindi nel nostro caso $\overrightarrow{w} = (-3, -2)$. Il fascio di rette perpendicolari a $\overrightarrow{v}$ ha equazione $-3x-2y+c = 0$; si tratta adesso di determinare c imponendo l'appartenenza di P_0 alla retta. Sostituiamo le coordinate del punto nel fascio di rette:$-3\cdot 2-2\cdot 1+c = 0$ da cui $c = 8$ e quindi l'equazione della retta è: $-3x - 2y + 8 = 0$ che, fortunatamente, è lo stesso risultato ricavato prima. In alternativa, trovata la direzione $\overrightarrow{w} = (-3, -2)$ si impone il passaggio per il punto $P(2, 1)$:$-3(x - 2) - 2(y - 1) = 0$ che fornisce sempre lo stesso risultato.

Il prossimo esempio mostra come passare dalla forma generale a quella parametrica.

Esempio 2.2.2 Scriviamo l'equazione della retta r: $2x - y + 4 = 0$ in forma parametrica. Anche in questo caso possiamo procedere in due modi:

1. Prendiamo un punto sulla retta, ad esempio $P_0(0, 4)$, ed un vettore perpendicolare ad r, ad esempio $\overrightarrow{w} = (2, -1)$. Ricaviamo adesso un vettore direzione per r: $\overrightarrow{v} = (1, 2)$, e scriviamo l'equazione parametrica di r:
$$\begin{cases} x = 0 + t = t \\ y = 4 + 2t \end{cases} .$$
2. Poniamo $x = t$ e ricaviamo y in funzione di questo parametro sostituendo nell'equazione generale della retta:
$$\begin{cases} x = t \\ y = 4 + 2t \end{cases} .$$

Esempio 2.2.3 Calcoliamo, se possibile, il punto P di intersezione delle rette dei due esempi precedenti:
$$r : \begin{cases} x = 2 - 2t \\ y = 1 + 3t \end{cases} ,$$
$$s : \begin{cases} x = k \\ y = 4 + 2k \end{cases} .$$
(Notare che è obbligatorio utilizzare lettere diverse per il parametro nelle due equazioni, per evitare grosse confusioni). Perché esista uno (ed un solo) punto intersezione, bisogna che le due rette non siano parallele (eventualmente coincidenti). Due rette sono parallele se e solo se hanno stessa direzione; questo significa che, presi due vettori direzione qualunque delle due rette, essi devono essere proporzionali: un vettore direzione di r è $\overrightarrow{v_r} = (-2, 3)$, uno di s è $\overrightarrow{v_s} = (1, 2)$. Se le componenti fossero proporzionali dovremmo trovare h tale che:

$$\begin{cases} -2h = 1 \\ 3h = 2 \end{cases},$$

il che è impossibile. Dunque le due rette non sono parallele. Per calcolare P risolviamo allora il sistema

$$\begin{cases} k = 2 - 2t \\ 4 + 2k = 1 + 3t \end{cases},$$

cioè

$$\begin{cases} k = 2 - 2t \\ 4 + 4 - 4t = 1 + 3t \end{cases},$$

$$\begin{cases} t = 1 \\ k = 0 \end{cases},$$

che fornisce il punto $P(0, 4)$.

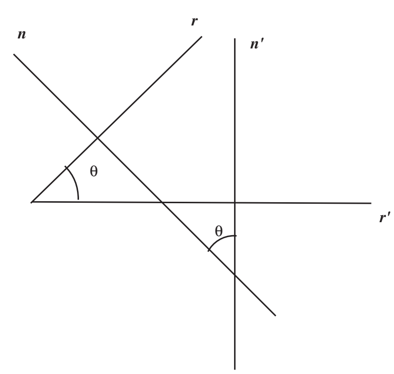
2.2.3 Angolo tra due rette

Se $\overrightarrow{v_1}$ e $\overrightarrow{v_2}$ sono due vettori direzione delle rette r,r', per calcolare l'angolo da esse formato utilizziamo il prodotto scalare:

$$\cos\theta = \frac{\overrightarrow{v_1} \cdot \overrightarrow{v_2}}{|\overrightarrow{v_1}||\overrightarrow{v_2}|}.$$

Se invece $\overrightarrow{n_1}$ e $\overrightarrow{n_2}$ sono due vettori normali alle rette:

$$\cos\theta = \frac{\overrightarrow{n_1} \cdot \overrightarrow{n_2}}{|\overrightarrow{n_1}||\overrightarrow{n_2}|}.$$



Angolo tra due rette
Figura 2.17

I due metodi sono equivalenti, ma è preferibile usare l'uno o l'altro a seconda di come la retta è espressa: se la retta è in forma parametrica abbiamo automaticamente vettori direzione quindi usiamo la prima formula, se è in forma generale abbiamo le normali e quindi usiamo la seconda.

Esempio 2.2.4 Calcoliamo l'angolo formato dalle due rette

$$r : \begin{cases} x = 2 - 2t \\ y = 1 + 3t \end{cases} ,$$

$$s : \begin{cases} x = k \\ y = 4 + 2k \end{cases} .$$

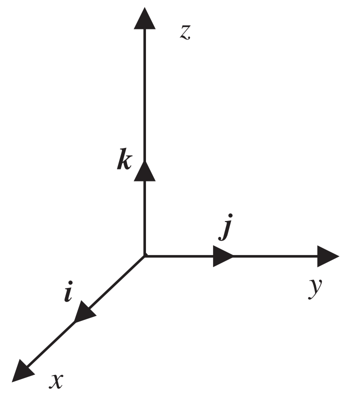
Dato che sono in forma parametrica usiamo la prima formula: $\cos\theta = \frac{\overrightarrow{v_1} \cdot \overrightarrow{v_2}}{|\overrightarrow{v_1}||\overrightarrow{v_2}|}$ con $\overrightarrow{v_1} = (-2, 3)$ vettore direzione della retta r e $\overrightarrow{v_2} = (1, 2)$ vettore direzione della retta s. Il prodotto scalare dei due vettori direzione è : $\overrightarrow{v_1} \cdot \overrightarrow{v_2} = 4$, i moduli dei due vettori sono: $|\overrightarrow{v_1}| = \sqrt{13}$, $|\overrightarrow{v_2}| = \sqrt{5}$ e quindi $\cos\theta = \frac{4}{\sqrt{65}}$ cioè $\theta \approx 60$ gradi.
Le rette r ed s in forma generale sono rappresentate dalle equazioni: r: $2y+3x-8 = 0$, s: $y - 2x - 4 = 0$ e due vettori normali alle rette sono: $\overrightarrow{n_1} = (-3, -2)$, $\overrightarrow{n_2} = (-2, 1)$, e quindi, procedendo come nel caso precedente: $\overrightarrow{n_1} \cdot \overrightarrow{n_2} = 4$, $|\overrightarrow{n_1}| = \sqrt{13}$, $|\overrightarrow{n_2}| = \sqrt{5}$ $\cos\theta = \frac{4}{\sqrt{65}}$, che fortunatamente fornisce lo stesso valore dell'angolo θ calcolato con l'altra formula.

2.3 Rette e piani nello spazio

2.3.1 Vettori nello spazio

Per non appesantire ulteriormente la trattazione,abbiamo considerato nei paragrafi precedenti i vettori come oggetti del piano. Tuttavia quanto fatto si trasporta facilmente allo stesso tipo di oggetti nello spazio. Vediamolo rapidamente.

Definizione 2.3.1 Si chiamano *versori degli assi* tre vettori ($\overrightarrow{i}$ per l'asse x, $\overrightarrow{j}$ per l'asse y, $\overrightarrow{k}$ per l'asse z) di modulo unitario, direzione coincidente con quella degli assi cartesiani e verso nel senso positivo degli assi stessi (*base canonica dello spazio*).



Versori nello spazio
Figura 2.18

- Un vettore $\overrightarrow{OP}$ con $P(x, y, z)$ si può scrivere: $\overrightarrow{OP} = x\,\overrightarrow{i} + y\,\overrightarrow{j} + z\,\overrightarrow{k}$;
- Il modulo di $\overrightarrow{OP}$ è $|\overrightarrow{OP}| = \sqrt{x^2 + y^2 + z^2}$;
- Il prodotto scalare tra due vettori $\overrightarrow{v_1} = x_1\,\overrightarrow{i} + y_1\,\overrightarrow{j} + z_1\,\overrightarrow{k}$ e $\overrightarrow{v_2} = x_2\,\overrightarrow{i} + y_2\,\overrightarrow{j} + z_2\,\overrightarrow{k}$ è dato da $\overrightarrow{v_1} \cdot \overrightarrow{v_2} = x_1x_2 + y_1y_2 + z_1z_2$;
- L'angolo tra i due vettori $\overrightarrow{v_1}$ e $\overrightarrow{v_2}$ è dato da $\cos\theta = \frac{\overrightarrow{v_1} \cdot \overrightarrow{v_2}}{|\overrightarrow{v_1}||\overrightarrow{v_2}|}$;

Il resto del capitolo è dedicato a mostrare come rette e piani nello spazio, e problemi relativi, possono efficacemente essere trattati con il linguaggio dei vettori.

2.3.2 Equazione del piano

Per trovare l'equazione generale di una retta r, abbiamo sfruttato l'idea che, dato un punto P_0 fissato della retta r ed un generico punto P di r allora, al variare di P, tutti i vettori di $P - P_0$ sono perpendicolari ad una direzione fissata. Sfruttiamo ora esattamente la stessa idea per arrivare all'equazione generale del piano.
Dati un vettore $\overrightarrow{n}$ ed un punto P_0,

$$\overrightarrow{n} = a\,\overrightarrow{i} + b\,\overrightarrow{j} + c\,\overrightarrow{k}, \quad P_0(x_0, y_0, z_0),$$

l'equazione del piano passante per P_0 e perpendicolare ad $\overrightarrow{n}$ è, in forma vettoriale:

$$(P - P_0) \cdot \overrightarrow{n} = 0 \tag{2.11}$$

e cioè, in forma scalare:

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0, \tag{2.12}$$

che può essere riscritta nella forma

$$ax + by + cz + d = 0, \tag{2.13}$$

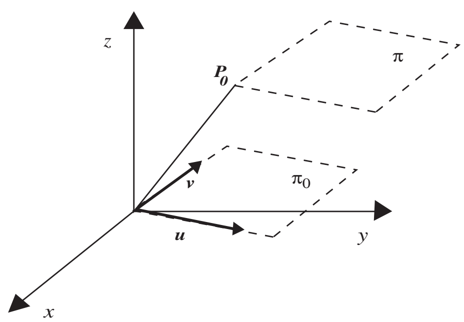
che dunque rappresenta l'*equazione generale del piano*.

Abbiamo visto come una retta possa essere individuata dandone un punto appartenente alla retta stessa e la direzione: questo porta alla rappresentazione parametrica della retta. Analogamente, un piano può essere individuato imponendo che contenga un punto dato e che sia parallelo al piano π_0 contente due vettori dati (non proporzionali). Quindi, i punti P del piano π possono essere rappresentati nella forma

$$\overrightarrow{P} = \overrightarrow{P_0} + t\,\overrightarrow{u} + s\,\overrightarrow{v},$$

che porta alla *forma parametrica del piano*:

$$\begin{cases} x = x_0 + tu_1 + sv_1 \\ y = y_0 + tu_2 + sv_2 \\ z = z_0 + tu_3 + sv_3, \end{cases}$$



Il piano π
Figura 2.19

Esempio 2.3.1 1. Il piano xy ha equazione $z = 0$ (passa per l'origine ed è perpendicolare all'asse z: $\overrightarrow{n} = \overrightarrow{k}$);
2. Il piano xz ha equazione $y = 0$;
3. Il piano yz ha equazione $x = 0$;
4. Un piano parallelo all'asse x ha come vettore normale un vettore $\overrightarrow{n}$ parallelo al piano yz, quindi un vettore della forma $\overrightarrow{n} = b\,\overrightarrow{j} + c\,\overrightarrow{k}$; dunque tale piano ha equazione $b(y - y_0) + c(z - z_0) = 0$, cioè un' equazione in cui manca il termine in x;
5. Un piano parallelo all'asse y ha equazione $a(x - x_0) + c(z - z_0) = 0$, un piano parallelo all'asse z ha equazione $a(x - x_0) + b(y - y_0) = 0$.

Esempio 2.3.2 Troviamo la forma parametrica e generale del piano passante per $(1,0,1)$ e parallelo al piano generato dai vettori $\overrightarrow{i+j+k}$ e $\overrightarrow{i}$. Da quanto visto prima, si ha che la forma parametrica del piano è:

$$\begin{cases} x = 1 + s + t \\ y = s \\ z = 1 + s \end{cases} ,$$

mentre la sua forma generale è $-y + z = 1$.

Esercizio 2.3.1 Vediamo alcuni semplici esercizi sui piani.

1. Scriviamo l'equazione del piano passante per $P_0(2,-1,0)$ e perpendicolare al vettore $\overrightarrow{v} = (2,3,5)$. Indicando con $P(x,y,z)$ un generico punto di $\mathbb{R}^3$, l'equazione del piano in forma generale è data da $\overrightarrow{PP_0} \cdot \overrightarrow{v} = 0$ e quindi

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

   L'equazione del piano è dunque: $2x + 3y + 5z - 1 = 0$;
2. Determiniamo l'equazione del piano passante per $P(2,1,3)$, $A(-1,2,0)$, $B(3,-4,2)$, dopo aver verificato che i tre punti non sono allineati.
   Per stabilire che i tre punti non sono allineati, basta verificare che i vettori $\overrightarrow{AP}$ e $\overrightarrow{AB}$ non sono paralleli. $\overrightarrow{AP} = -3\overrightarrow{i} + \overrightarrow{j} - 3\overrightarrow{k}$, $\overrightarrow{AB} = 4\overrightarrow{i} - 6\overrightarrow{j} + 2\overrightarrow{k}$, le loro componenti non sono proporzionali, infatti non esiste alcun valore di h che soddisfa il sistema:

   $$\begin{cases} -3h = 4 \\ h = -6 \\ -3h = 2 \end{cases} .$$

   Per determinare l'equazione del piano richiesto, consideriamo un qualsiasi vettore perpendicolare a tale piano, cioè un vettore $\overrightarrow{v}$ tale che $\overrightarrow{AP} \cdot \overrightarrow{v} = 0$, $\overrightarrow{AB} \cdot \overrightarrow{v} = 0$. Quindi, se $\overrightarrow{v} = (a,b,c)$, allora

   $$\begin{cases} (a,b,c) \cdot (-3,1,-3) = 0 \\ (a,b,c) \cdot (4,-6,2) = 0 \end{cases} ,$$

   $$\begin{cases} -3a + b - 3c = 0 \\ 4a - 6b + 2c = 0 \end{cases} ,$$

   $$\begin{cases} c = \frac{-7}{8}a \\ b = \frac{3}{8}a \end{cases} .$$

   Troviamo allora un vettore perpendicolare al piano scegliendo opportunamente a. Una buona scelta fornisce $\overrightarrow{v} = (8,3,-7)$. Sostituendo il vettore trovato nell'equazione iniziale, abbiamo il piano cercato:
   $((x-2),(y-1),(z-3)) \cdot (8,3,-7) = 0$, cioè $8x + 3y - 7z + 2 = 0$.

2.3.3 Equazione della retta nello spazio

Dati due piani, $ax + by + cz + d = 0$, $a'x + b'y + c'z + d' = 0$, che non siano né coincidenti né paralleli, la loro intersezione è una retta. Il sistema costituito dalle due equazioni rappresenta allora la *forma generale* della retta nello spazio.

$$r : \begin{cases} ax + by + cz + d = 0 \\ a'x + b'y + c'z + d' = 0 \end{cases} .$$

Per verificare che i due piani non siano né coincidenti né paralleli, basta controllare che non abbiano la stessa direzione normale e poiché $\overrightarrow{n} = (a, b, c)$ e $\overrightarrow{n'} = (a', b', c')$ individuano le direzioni normali, è sufficiente controllare che i due vettori non siano proporzionali.
La stessa retta può essere rappresentata in

- *Forma parametrica*:
$$r : \begin{cases} x = x_0 + lt \\ y = y_0 + mt \\ z = z_0 + nt \end{cases} ,$$
con $P_0(x_0, y_0, z_0) \in \mathbb{R}^3$ e $\overrightarrow{v} = (l, m, n)$ vettore direzione di r. La stessa forma parametrica si può scrivere anche: $P - P_0 = t\overrightarrow{v}$ con $\overrightarrow{v} = l\,\overrightarrow{i} + m\,\overrightarrow{j} + n\,\overrightarrow{k} \neq \overrightarrow{0}$;
- *Forma normale:*
Se $l \neq 0, m \neq 0, n \neq 0$,
$$r : \frac{x - x_0}{l} = \frac{y - y_0}{m} = \frac{z - z_0}{n}$$

2.3.4 Intersezione retta-piano

Consideriamo il problema di determinare l'intersezione fra un piano ed una retta.

Se la retta è in forma generale

$$r : \begin{cases} ax + by + cz + d = 0 \\ a'x + b'y + c'z + d' = 0 \end{cases} ,$$

ed il piano ha equazione $a''x + b''y + c''z = 0$, otteniamo un sistema di tre equazioni in tre incognite che ha in generale (salvo casi particolari, pensate quali) una ed una sola soluzione. Supponiamo invece che la retta sia in forma parametrica:

$$r : \begin{cases} x = x_0 + lt \\ y = y_0 + mt \\ z = z_0 + nt \end{cases}$$

ed il piano abbia equazione $a''x + b''y + c''z = 0$; sostituendo nell'equazione del piano le coordinate x, y, z della retta si ottiene un'equazione con un incognita (che lettera è l'incognita?), che ha in generale una ed una sola soluzione.

Esempio 2.3.3 Scriviamo l'equazione del piano α passante per $P(1, 2, 1)$ e perpendicolare alla retta r di equazione parametrica

$$r : \begin{cases} x = 2 - 2t \\ y = -3 - t \\ z = 1 + t \end{cases} ,$$

e troviamo il punto A di intersezione tra α ed r.
Un vettore direzione di r è $\overrightarrow{v} = (-2, -1, 1)$ e poiché r deve essere perpendicolare al piano, dovrà essere nullo il prodotto scalare tra i vettori $\overrightarrow{PX}$ (dove $X(x, y, z)$ è un punto generico del piano α) e $\overrightarrow{v}$:
$\overrightarrow{PX} \cdot \overrightarrow{v} = 0 \Longrightarrow -2(x-1) - (y-2) + (z-1) = 0$, cioè $2x + y - z - 3 = 0$.
Per calcolare le coordinate del punto A di intersezione tra la retta r ed il piano trovato, sostituiamo l'equazione della retta in forma parametrica in quella del piano ed otteniamo un'equazione in un'incognita:
$2(2-2t) + (-t-3) - (t+1) - 3 = 0$ da cui si ricava $t = -\frac{1}{2}$, che sostituito nella retta ci dà le coordinate di A:

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

2.3.5 Intersezione retta-retta

Consideriamo il problema di individuare l'intersezione tra due rette nello spazio. È geometricamente evidente che le due rette, pur se non sono parallele, in genere non si intersecano, e cioè sono *sghembe,* perché non esiste un piano che le contenga entrambe. Come si traduce questo in termini analitici? Se le rette sono in forma generale si tratta di mettere a sistema quattro equazioni nelle tre incognite x, y, z. In genere, un sistema con più equazioni che incognite non ha soluzione. Se le rette sono date in forma parametrica, ci si riduce allo studio di un sistema di tre equazioni in due incognite (i due parametri).
Vediamo qualche esempio.

Esempio 2.3.4 1. Proviamo che le rette

$$r : \begin{cases} x = 1 + 2t \\ y = -2 - t \\ z = -3 + 3t \end{cases} ,$$

e

$$r' : \begin{cases} x = 3 + t' \\ y = 2 + 2t' \\ z = -t' \end{cases} ,$$

sono sghembe.
Come nell'esempio precedente, mettiamo a sistema le due rette:

$$\begin{cases} 1 + 2t = 3 + t' \\ -2 - t = 2 + 2t' \\ -3 + 3t = -t' \end{cases} ,$$

da cui

$$\begin{cases} t' = 2t - 2 \\ t = 0 \\ -3 = -t' \end{cases} ,$$

e quindi

$$\begin{cases} t' = -2 \\ t = 0 \\ t' = 3 \end{cases} .$$

Il sistema è impossibile. Le rette dunque non si intersecano. Osservando inoltre che due loro vettori direzione non sono proporzionali, le rette non possono essere parallele. Se ne conclude quindi che sono sghembe.

2. Date le rette di equazioni parametriche

$$r : \begin{cases} x = 1 + 2t \\ y = -2 - t \\ z = -3 + 3t \end{cases} ,$$

$$r' : \begin{cases} x = 3 + t' \\ y = 2 + 2t' \\ z = -5 - t' \end{cases} ,$$

proviamo che sono incidenti e calcoliamo il loro punto di intersezione.
Se le due rette sono incidenti, esiste un punto P che appartiene ad entrambe, un punto cioè le cui coordinate soddisfano entrambe le equazioni delle rette r ed r'.
Mettendo a sistema le due rette otteniamo

$$\begin{cases} 1 + 2t = 3 + t' \\ -2 - t = 2 + 2t' \\ -3 + 3t = -5 - t' \end{cases} ,$$

da cui

$$\begin{cases} t' = 2t - 2 \\ t = 0 \\ -3 = -3 \end{cases},$$

e infine

$$\begin{cases} t' = -2 \\ t = 0 \end{cases}.$$

Sostituendo il valore dei parametri trovati ($t = 0$ nella retta r oppure $t' = -2$ nella retta r'), abbiamo il punto cercato: $P(1, -2, -3)$.

2.4 Approfondimenti

2.4.1 Proprietà dei vettori

In questo paragrafo riportiamo in maniera più sistematica le proprietà delle operazioni sui vettori che abbiamo in parte già descritto precedentemente.

Proprietà della somma

1) Proprietà associativa: dati tre vettori $\overrightarrow{OP}$, $\overrightarrow{OQ}$, $\overrightarrow{OR}$; $(\overrightarrow{OP} + \overrightarrow{OQ}) + \overrightarrow{OR} = \overrightarrow{OP} + (\overrightarrow{OQ} + \overrightarrow{OR})$;
2) Sia $\overrightarrow{OO}$ il vettore nullo; allora: $\overrightarrow{OO} + \overrightarrow{OP} = \overrightarrow{OP}$ (ecco un buon motivo per considerare il vettore nullo);
3) Vettore opposto: Sia $-\overrightarrow{OP}$ il vettore che stessa intensità e direzione di $\overrightarrow{OP}$, ma verso opposto. Allora $\overrightarrow{OP} + (-\overrightarrow{OP}) = \overrightarrow{OO}$;
4) Proprietà commutativa: siano $\overrightarrow{OP}$, $\overrightarrow{OQ}$ due vettori. Allora $\overrightarrow{OP} + \overrightarrow{OQ} = \overrightarrow{OQ} + \overrightarrow{OP}$.

Proprietà del prodotto di un vettore per uno scalare

1. Proprietà associativa: siano $s, t \in \mathbb{R}$ ed $\overrightarrow{OP}$ un vettore, allora $t(s\overrightarrow{OP}) = (ts)\overrightarrow{OP}$;
2. $1 \cdot \overrightarrow{OP} = \overrightarrow{OP}$;
3. $0 \cdot \overrightarrow{OP} = \overrightarrow{OO}$;
4. Legge di annullamento del prodotto: se $t \cdot \overrightarrow{OP} = \overrightarrow{OO}$ allora $t = 0$ oppure $\overrightarrow{OP} = \overrightarrow{OO}$;
5. Proprietà distributive:
   a) $(t + s)\overrightarrow{OP} = t \cdot \overrightarrow{OP} + s \cdot \overrightarrow{OP}$;
   b) $t \cdot (\overrightarrow{OP} + \overrightarrow{OQ}) = t \cdot \overrightarrow{OP} + t \cdot \overrightarrow{OQ}$.

Proprietà del prodotto scalare

1) Proprietà commutativa: $\overrightarrow{OP} \cdot \overrightarrow{OQ} = \overrightarrow{OQ} \cdot \overrightarrow{OP}$;

2) $(t \cdot \overrightarrow{OP}) \cdot (\overrightarrow{OQ}) = t \cdot (\overrightarrow{OP} \cdot \overrightarrow{OQ})$;
3) $\overrightarrow{OP} \cdot (\overrightarrow{OQ} + \overrightarrow{OR}) = \overrightarrow{OP} \cdot \overrightarrow{OQ} + \overrightarrow{OP} \cdot \overrightarrow{OR}$;
4) $(\overrightarrow{OP} + \overrightarrow{OQ}) \cdot \overrightarrow{OR} = \overrightarrow{OP} \cdot \overrightarrow{OR} + \overrightarrow{OQ} \cdot \overrightarrow{OR}$.

2.5 Esercizi

2.5.1 Quesiti a risposta chiusa

Rispondere vero o falso.
Siano $\overrightarrow{v_1}$ e $\overrightarrow{v_2}$ due vettori tali che $|\overrightarrow{v_1}| = 1$,$|\overrightarrow{v_2}| = 2$ e $\overrightarrow{v_1} \cdot \overrightarrow{v_2} = 1$:

1. $|\overrightarrow{v_1} + v_2| = 7$;
2. L'angolo tra $\overrightarrow{v_1}$ e $\overrightarrow{v_2}$ è $\vartheta = \frac{\pi}{3}$;
3. $\overrightarrow{v_2 - v_1}$ e $\overrightarrow{v_1}$ sono perpendicolari.

Sia α il piano di equazione: $-x + y + 2z = 1$. Sia r la retta perpendicolare ad α passante per $P(1, 0, 1)$.

1. $P \in \alpha$;
2. $\overrightarrow{v} = -\overrightarrow{i} + \overrightarrow{j} + 2\overrightarrow{k}$ è un vettore direzione di r;
3. $\overrightarrow{w} = \overrightarrow{i} - \overrightarrow{j} + \overrightarrow{k}$ è un vettore normale ad r;
4. L'equazione generale della retta r è $2x + y + \frac{1}{2}z = \frac{5}{2}$;
5. L'equazione parametrica di r è $\begin{cases} x = 1 - 2t \\ y = 2t \\ z = 1 + 4t \end{cases}$.

Dati i piani α: $x - 2y + z + 3 = 0$ e β: $x + y + z + 4 = 0$:

1. sono paralleli;
2. la retta r: $\begin{cases} x = 2t \\ y = -4t + 3 \\ z = 2t - 1 \end{cases}$ è perpendicolare al piano α;
3. la retta r interseca il piano β;
4. $\overrightarrow{w} = (1, 1, 1)$ è un vettore perpendicolare a β.

2.5.2 Esercizi aperti

Esercizio 2.5.1 Dati i vettori:

$$\overrightarrow{v} = (1, 0, 2) \quad \overrightarrow{u} = (-1, 2, 3) \quad \overrightarrow{w} = (0, 2, 1),$$

calcolare:

1. $\overrightarrow{v} + \overrightarrow{w} - 2\overrightarrow{u}$;

2. $|\overrightarrow{u}|$;
3. $\overrightarrow{u} \cdot \overrightarrow{v}$;
4. l'angolo formato tra $\overrightarrow{v}$ e $\overrightarrow{u}$ e quello formato tra $\overrightarrow{v}$ e $\overrightarrow{w}$.

Esercizio 2.5.2 Dati i vettori:

$$\overrightarrow{v} = 2\overrightarrow{i} - \overrightarrow{j} + \overrightarrow{k} \quad \overrightarrow{w} = \overrightarrow{i} + \overrightarrow{j}$$

calcolare il prodotto scalare $\overrightarrow{v} \cdot \overrightarrow{w}$. Qual è l'angolo formato tra i due vettori?

Esercizio 2.5.3 Scrivere le equazioni delle seguenti rette in forma parametrica ed in forma generale:

1. la retta passante per $P(2,1)$ e $Q(-1,-1)$;
2. la retta passante per $P(2,1)$ ed avente come vettore direzione $\overrightarrow{v} = \overrightarrow{i} - 3\overrightarrow{j}$;
3. la retta passante per $P(-1,3)$ e perpendicolare al vettore $\overrightarrow{v} = 2\overrightarrow{i} + \overrightarrow{j}$.

Esercizio 2.5.4 Sia r la retta di equazione $2x + y - 1 = 0$ e s la retta in forma parametrica $\begin{cases} x = 1 - t \\ y = -1 + 2t \end{cases}$.
Calcolare:

1. un vettore direzione di r;
2. un vettore direzione di s;
3. l'angolo formato tra r ed s;
4. un vettore perpendicolare ad s;
5. per quale valore di b la retta $bx + (b-1)y + 2 = 0$ è perpendicolare ad s?

Esercizio 2.5.5 Scrivere l'equazione delle seguenti rette in forma parametrica:

1. La retta r passante per i punti $(1,0,1)$ e $(-1,-1,0)$;
2. La retta passante per il punto $(-2,1,2)$ ed avente la direzione del vettore $\overrightarrow{v} = \overrightarrow{i} - \overrightarrow{k}$;
3. La retta intersezione dei due piani $x - 2y + z = 0$ e $2x + y - 3z = 0$;
4. La retta parallela ad r e passante per $(1,0,-1)$.

Esercizio 2.5.6 Dati i punti $A(1,-1,2) \quad B(0,-3,3) \quad C(1,0,2)$, verificare che non sono allineati. Scrivere l'equazione della retta passante per A e B, del piano passante per A e perpendicolare a $\overrightarrow{v} = \overrightarrow{BC}$.

3

Matrici

Quando si costruisce un modello matematico di un problema, spesso per trovarne la soluzione bisogna risolvere sistemi con molte equazioni e molte incognite, ed in genere molto complessi da risolvere, se non del tutto intrattabili. Inoltre, la struttura dell'insieme delle loro soluzioni è sconosciuta. Pertanto il sistema ottenuto viene spesso approssimato da un opportuno sistema lineare. Questo ha due vantaggi: prima di tutto i sistemi lineari sono analizzabili da un punto di vista teorico, e l'insieme delle loro soluzioni ha una struttura ben precisa, che ricorre frequentemente in altre situazioni. Secondariamente, dal punto di vista del calcolo delle sue soluzioni, esistono algoritmi potenti che permettono ai calcolatori di risolvere sistemi di dimensioni enormi. Questo, a dispetto del fatto che "a mano", già sistemi con più di tre equazioni e tre incognite sono noiosi e lunghi da risolvere. Il fatto di poter trattare algoritmicamente una quantità spettacolare di dati rende dunque i sistemi lineari utili nei campi più svariati, come ad esempio l'economia o la meteorologia.
Per analizzare la struttura dei sistemi lineari, è utile introdurre un nuovo oggetto matematico, detto matrice. È quanto facciamo in questo capitolo, in cui definiamo le matrici e le principali operazioni che si possono fare fra matrici. Val la pena osservare che l'oggetto matrice, utilizzato soprattutto per la descrizione dei sistemi lineari, risulta poi uno strumento molto utile in altri campi non direttamente collegati all'algebra lineare: vedremo ad esempio che esse descrivono efficacemente un gioco fra due giocatori, finito e strettamente competitivo.

Definizione 3.0.1 *Una* matrice A *è una tabella di numeri reali*[1].

Ad esempio

$$A = \begin{pmatrix} 1 & 2 & -3 \\ 0 & -5 & 4 \end{pmatrix}.$$

[1] È naturalmente possibile considerare matrici fatte da numeri complessi. In queste note non lo faremo praticamente mai

In una matrice sono individuate delle righe e delle colonne. La matrice A di sopra è costituita da due righe e tre colonne. Si dice allora che A è una matrice 2×3.

Definizione 3.0.2 *Si chiama* vettore colonna *una matrice $n \times 1$, ossia una matrice che ha una sola colonna.*

Una matrice A del tipo 2×3 ha in generale la forma $\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix}$. Il primo pedice indica la riga e il secondo la colonna. La matrice A viene anche indicata con $A = \left(a_{ij} \right)$ o anche $A = \left[a_{ij} \right]$, dove i è l'indice di riga e j quello di colonna. Tutto si generalizza in maniera ovvia ad una matrice $n \times m$.

Definizione 3.0.3 *Si chiama* matrice trasposta *di A e si indica con A^t la matrice che si ottiene scambiando le righe e le colonne tra loro. Dato un vettore colonna $\overrightarrow{a}$, il suo trasposto $\overrightarrow{a}^t$ si chiama* vettore riga.

Se A è una matrice $m \times n$ allora A^t è una matrice $n \times m$.

Ad esempio, se $A = \begin{pmatrix} 2 & 4 & 1 \\ 3 & 0 & 5 \end{pmatrix}$, allora $A^t = \begin{pmatrix} 2 & 3 \\ 4 & 0 \\ 1 & 5 \end{pmatrix}$.

Se $\overrightarrow{a} = \begin{pmatrix} a_1 \\ a_2 \\ . \\ . \\ a_n \end{pmatrix}$, allora $(\overrightarrow{a})^t = (a_1, a_2, \ldots, a_n)$, dove gli indici questa volta rappresentano indici di colonna.

Osservazione 3.0.1 $(A^t)^t = A$.

Definizione 3.0.4 *La matrice $n \times m$ i cui elementi sono tutti nulli è detta* matrice nulla.

Definizione 3.0.5 *Si chiama* matrice quadrata di ordine n *una matrice $n \times n$. Gli elementi della forma a_{ii} formano la cosiddetta* diagonale principale *della matrice quadrata.*

Definizione 3.0.6 *Si chiama* matrice unità (di ordine n) *la matrice quadrata I_n che ha gli elementi uguali ad 1 sulla diagonale principale e 0 in tutti gli elementi extra diagonali. Quindi*

$$I_n = [\delta_{ij}] = \begin{cases} 1 & se\ i = j \\ 0 & se\ i \neq j. \end{cases}$$

Esiste una matrice unità per ogni valore di n; mentre la matrice nulla può essere rettangolare, la matrice unità deve essere quadrata.

Definizione 3.0.7 *Una matrice quadrata di ordine n si dice* diagonale *se* $a_{ij} = 0 \quad \forall i \neq j$. *Si indica con* diag $(a_{11}, a_{22}, a_{33}, ..., a_{nn})$.

Ad esempio è diagonale la matrice $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} = \text{diag}\,(1, 2, 3)$.

Definizione 3.0.8 *Una matrice matrice quadrata si dice* simmetrica *se* $a_{ij} = a_{ji} \quad \forall i, j$.

In altre parole, A è simmetrica se $A = A^t$.

Ad esempio è simmetrica la matrice $A = \begin{pmatrix} 1 & 1 & 2 \\ 1 & 4 & 5 \\ 2 & 5 & 3 \end{pmatrix}$.

Definizione 3.0.9 *Ua matrice quadrata si dice* emisimmetrica *se verifica* $A = -A^t$.

Ad esempio è emisimmetrica la matrice $A = \begin{pmatrix} 0 & 1 & 2 \\ -1 & 0 & -5 \\ -2 & 5 & 0 \end{pmatrix}$.

3.1 Operazioni sulle matrici

Ovviamente dobbiamo dire come operiamo sulle matrici. È quindi il momento di introdurre alcune operazioni fra matrici.

1. *Somma:* date due matrici $m \times n$ $A = (a_{ij})$ e $B = (b_{ij})$ definiamo la matrice $A + B = C = (c_{ij})$ come la matrice $m \times n$ tale che $c_{ij} = a_{ij} + b_{ij} \quad \forall i, j$;
   Notiamo che si possono sommare solo matrici di uguale dimensione, e che la somma si fà sommando tra loro gli elementi di posto corrispondente.
2. *Prodotto di una matrice per uno scalare:* sia $A = (a_{ij})$ una matrice $m \times n$ e t uno scalare: definiamo la matrice $tA = D = (d_{ij})$ come la matrice $m \times n$ tale che $d_{ij} = ta_{ij} \quad \forall i, j$;
3. *Prodotto righe per colonne:* sia $A = (a_{ij})$ una matrice $m \times n$ e $B = (b_{ij})$ una matrice $n \times p$. Allora $A \cdot B = C = (c_{ij})$ è una matrice $m \times p$ dove $c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$ con $1 \leq i \leq m \quad 1 \leq k \leq p$. [2]
   È importante osservare che per poter svolgere il prodotto righe per colonne $A \cdot B$ è necessario che il numero delle colonne di A sia uguale al numero di righe di B.

[2] Per indicare il prodotto righe per colonne di A per B useremo indifferentemente le notazioni $A \cdot B$ e AB.

Osservazione 3.1.1 *Prodotto scalare (fra vettori colonna):* il prodotto scalare fra due vettori $\overrightarrow{a}$ e $\overrightarrow{b}$ si può vedere come caso particolare di prodotto righe per colonne:

se $\overrightarrow{a}^t = (a_1, \dots, a_n)$ è un vettore riga, e $\overrightarrow{b} = \begin{pmatrix} b_1 \\ . \\ . \\ . \\ b_n \end{pmatrix}$ un vettore colonna, allora:

$$\overrightarrow{a}^t \cdot \overrightarrow{b} = a_1b_1 + a_2b_2 + ... + a_nb_n = \sum_{k=1}^{n} a_kb_k.$$

Naturalmente i due vettori devono avere la stessa dimensione.

Esempio 3.1.1 Consideriamo ad esempio le matrici $A = \begin{pmatrix} 1 & 0 \\ 2 & 3 \\ -1 & 5 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 2 \end{pmatrix}$. Eseguiamo il prodotto righe per colonne della matrice A per la matrice B. Otterremo una matrice 3×3. $AB = \begin{pmatrix} 1+0 & 0+0 & 0+0 \\ 2+6 & 0+3 & 0+6 \\ -1+10 & 0+5 & 0+10 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 8 & 3 & 6 \\ 9 & 5 & 10 \end{pmatrix}$. Si può anche eseguire il prodotto BA, che fornisce una matrice 2×2. Si ha $BA = \begin{pmatrix} 1 & 0 \\ 2 & 13 \end{pmatrix}$.

Osserviamo esplicitamente che:

- Date due matrici qualunque A, B , non si può in genere fare il loro prodotto;
- Può essere possibile fare un prodotto, ad esempio AB, ma non l'altro (BA);
- Anche se si possono fare entrambi, di solito si ottengono matrici *di dimensioni differenti*;
- Anche nel caso in cui le matrici AB, BA abbiano la stessa dimensione, non è detto che siano la stessa matrice.

Date due matrici quadrate A e B, se $A \cdot B = B \cdot A$ allora si dice che B è *permutabile* con A, e viceversa. $A^2 = A \cdot A$ è sempre permutabile con A, $A \cdot A^2 = A^2 \cdot A = A^3$, in generale $A^n \cdot A^m = A^m \cdot A^n = A^{n+m}$. Se le matrici A e B non sono permutabili allora in generale $A^2 \cdot B^2 \neq (AB)^2$. Infine osserviamo che due matrici diagonali sono sempre permutabili, ed il risultato è una matrice diagonale.

Osservazione 3.1.2 Osserviamo che non vale il teorema sull'annullamento del prodotto per le matrici: se $A \neq O$ con O matrice nulla, e $B \neq O$ allora può essere $A \cdot B = O$. Ad esempio $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$.

Un caso particolare di prodotto fra matrici è il prodotto di una matrice per un vettore colonna. Dati A di dimensione $m \times n$ e $\overrightarrow{x}$ di dimensione $n \times 1$ il prodotto $A \cdot \overrightarrow{x}$ dà un vettore colonna di dimensione m.

Proprietà delle operazioni precedenti

La seguente è una lista lunga e noiosa delle proprietà delle operazioni fra matrici appena introdotte. Più che provare a dimostrare queste proprietà, che sono spesso immediate, sarebbe utile osservare quando, nel seguito, esse vengono utilizzate.

1. $(A+B)+C = A+(B+C)$;
   Vale la proprietà associativa dell'addizione tra matrici.
2. $A+B = B+A$;
   Vale la proprietà commutativa dell'addizione.
3. $A+O = A$ dove con O si indica la matrice nulla della stessa dimensione della matrice A;
   Esistenza dell'elemento neutro rispetto all'operazione di addizione di matrici.
4. Esiste una matrice $B = (b_{ij})$ tale che $A+B = O$; B si si chiama *matrice opposta* di A. B è unica e vale $B = (-1)A$ e viene, naturalmente, indicata con $-A$;
   Esistenza dell'elemento opposto.
5. $t(A+B) = tA+tB$ con $t \in \mathbb{R}$;
   Vale la proprietà distributiva della moltiplicazione per uno scalare rispetto all'addizione di matrici.
6. $(t+s)A = tA+sA$ con $t, s \in \mathbb{R}$;
   Vale la proprietà distributiva della moltiplicazione di matrici rispetto all'addizione di scalari.
7. Sia A una matrice $n \times m$. Allora $A \cdot I_n = A$ e $I_m \cdot A = A$;
   Esistenza dell'elemento neutro rispetto alla moltiplicazione. Attenzione alle dimensioni delle matrici unità.
8. $(A \cdot B)^t = B^t \cdot A^t$;
9. $(A \cdot B) \cdot C = A \cdot (B \cdot C)$;
   Proprietà associativa della moltiplicazione di matrici.
10. $(A+B) \cdot C = (A \cdot C) + (B \cdot C)$;
    Proprietà distributiva della moltiplicazione a destra rispetto all'addizione.
11. $A \cdot (B+C) = A \cdot B + A \cdot C$;
    Proprietà distributiva della moltiplicazione a sinistra rispetto all'addizione.
12. $t(A+B) = tA+tB$ con $t \in \mathbb{R}$.

Come abbiamo detto all'inizio del capitolo, una delle ragioni per introdurre le matrici è che il calcolo matriciale dà la possibilità di analizzare con efficacia la struttura delle soluzioni di un sistema lineare, e di suggerirne il modo per risolverli. Un sistema lineare di m equazioni ed n incognite si può scrivere nella forma:

$$\begin{cases} a_{11}x_1 + ... + a_{1n}x_n = b_1 \\ ... \\ ... \\ a_{m1}x_1 + ... + a_{mn}x_n = b_m \end{cases} . \tag{3.1}$$

Possiamo ora utilizzare il prodotto righe per colonne, per scrivere il sistema (3.1) in forma compatta: indicando con $A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$ (chiamata *matrice dei coefficienti*), con $\overrightarrow{x} = \begin{pmatrix} x_1 \\ \cdot \\ x_n \end{pmatrix}$ il vettore delle incognite, e con $\overrightarrow{b} = \begin{pmatrix} b_1 \\ \cdot \\ b_m \end{pmatrix}$ il vettore dei termini noti. Dunque, in notazione matriciale il sistema diventa

$$A\overrightarrow{x} = \overrightarrow{b}. \tag{3.2}$$

Come analizzare e risolvere, avendo introdotto le matrici, un sistema come questo? Innanzi tutto osserviamo che, se il numero delle equazioni ed il numero delle incognite è lo stesso, molti “indizi” ci inducono a pensare che in genere ci possiamo aspettare un'unica soluzione del sistema. Se il sistema è due per due, o tre per tre, la geometria ci suggerisce che cerchiamo l'intersezioni di due rette nel piano o di tre piani nello spazio, e ci aspettiamo dunque di trovare in entrambi i casi un punto, cioè una soluzione. Se abbiamo una mente più algebrica, possiamo pensare che, ad esempio se il sistema ha due equazioni e come incognite x, y, ricavando ad esempio x in funzione di y nella prima equazione, e sostituendo nella seconda, questa diventa un'equazione nella sola variabile y, che si risolve. Trovato y, si ricava x andando a sostituire nella prima. . . .
È anche chiaro che se il sistema è di tre equazioni e con tre incognite si può iterare il procedimento. Se le incognite si indicano con x, y, z potremmo ricavare x dalla prima, in funzione di y, z, e sostituirlo nella seconda e nella terza, che a questo punto, formano un sistema in due variabili, da cui, come nel caso precedente, si ricavano y e z che, sostituite nella prima equazione, ci permettono di ricavare anche la x. Occorre tenere conto però che la geometria stessa ci indica che dobbiamo aspettarci dei casi “eccezionali”: ad esempio due rette potrebbero essere parallele, ed allora non hanno evidentemente punti in comune.
Tornando al sistema scritto in scrittura matriciale (3.2), iniziamo col fare, per il momento, due ipotesi semplificative, che in seguito elimineremo. Imponiamo per prima cosa che $m = n$, cioè che il numero delle equazioni eguagli quello delle incognite, ovverossia che la matrice dei coefficienti del sistema sia quadrata. Supponiamo anche che esista una matrice $\widetilde{A}$ tale che $\widetilde{A}A = I_n$. Si ha allora, da $A\overrightarrow{x} = \overrightarrow{b}$,

$$\widetilde{A}A\overrightarrow{x} = \widetilde{A}\overrightarrow{b},$$

cioè

$$I_n\overrightarrow{x} = \widetilde{A}\overrightarrow{b}$$

e quindi

$$\overrightarrow{x} = \widetilde{A}\overrightarrow{b}.$$

Questo significa che il sistema è risolto! Naturalmente, il calcolo sopra è un po' troppo formale, ma ci dà un'idea di come procedere. Osserviamo anche che il calcolo precedente motiva l'idea di definire il prodotto righe per colonne e giustifica l'importanza

che si dà alla matrice identità.
Cominciamo col dare la seguente:

Definizione 3.1.1 *La matrice quadrata $A(n \times n)$ si dice* invertibile *se esiste una matrice $\widetilde{A}$ $(n \times n)$ tale che $\widetilde{A}A = A\widetilde{A} = I_n$.*

Più tardi vedremo condizioni necessarie e sufficienti affinché una matrice sia invertibile. Per il momento osserviamo che tale matrice $\widetilde{A}$ è unica: sia infatti B un'altra con la stessa proprietà: allora

$$B = BI_n = BA\widetilde{A} = I_n\widetilde{A} = \widetilde{A}.$$

La matrice $\widetilde{A}$ viene chiamata *inversa di* A ed indicata con A^{-1}. Vedremo dopo come calcolarla. Ora enunciamo le seguenti proprietà:

1. se A è una matrice diagonale, $A = \text{diag}\ (\lambda_1, \lambda_2, ...\lambda_n)$, allora A^{-1} esiste se e solo se tutti gli elementi sulla diagonale sono diversi da zero e $A^{-1} = \text{diag}\ (\frac{1}{\lambda_1}, \frac{1}{\lambda_2}, ..., \frac{1}{\lambda_n})$; si vede con verifica diretta.
2. se A e B sono invertibili allora anche AB è invertibile e $(AB)^{-1} = B^{-1}A^{-1}$.
   Infatti:
   $$(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}IB = B^{-1}B = I;$$
3. se A è invertibile allora anche A^{-1} è invertibile.
   Infatti $(A^{-1})^{-1} = A$.

La definizione di matrice inversa ci permette tra l'altro di estendere il concetto di potenza di una matrice alle potenze negative.

Definizione 3.1.2 *Se A è una matrice invertibile allora possiamo definire $A^{-n} = (A^{-1})^n$.*

Valgono le seguenti proprietà :

1. $A^n \cdot A^m = A^{m+n}$ con $m, n \in Z$;
2. $(A^n)^m = A^{mn}$ con $m, n \in Z$;
3. $A^0 = I$.

Abbiamo dunque messo in evidenza una proprietà (l'invertibilità) che permette, almeno in teoria, di risolvere sistemi lineari di n equazioni in n incognite. Si tratta allora di stabilire, prima di tutto, un criterio che permetta di verificare se una matrice è invertibile e fornire poi una formula per trovare l'inversa. Questo lo vediamo nel prossimo paragrafo. Per concludere questo, osserviamo che l'idea precedente si applica solo ai sistemi con tante equazioni quante incognite, e che l'idea proposta è interessante soprattutto *in teoria.* In pratica, come vedremo, esistono metodi algoritmicamente più efficienti per risolvere i sistemi lineari.

3.2 Determinante di una matrice

Per analizzare e risolvere i sistemi lineari, è noto che si ricorre a manipolazioni che portano a sistemi equivalenti ma via via sempre più facili da risolvere, in modo da poter individuare le soluzioni. Ad esempio nel sistema:

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

è importante saper riconoscere che una delle prime due equazioni è ridondante (e quindi il sistema è riconducibile ad uno 2×2).
Il sistema

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

è poi equivalente al sistema

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

(di immediata soluzione) ottenuto lasciando inalterata la seconda equazione e sostituendo la prima con la somma della prima e della seconda.
Volendo trattare il sistema utilizzando il calcolo matriciale, abbiamo ancora bisogno di un'importante strumento legato alle matrici, che "si comporti bene" rispetto alle operazioni di cui abbiamo visto sopra un esempio. Questo strumento, che riguarda solo le matrici quadrate, è il *determinante.*

Definizione 3.2.1 *Si chiama* sottomatrice di A *una matrice ottenuta da* A *togliendo alcune righe e/o colonne di* A*. Se la sottomatrice di* A *è quadrata, allora si chiama* minore *di* A*.*

Definizione 3.2.2 *Si chiama* minore complementare *dell'elemento* a_{ij}*, e si indica con* M_{ij}*, il determinante della matrice di ordine* $n-1$ *ottenuta da* A *eliminando la* $i-$*esima riga e la* $j-$*esima colonna. Si chiama* complemento algebrico *dell'elemento* a_{ij} *e si denota con* A_{ij}*, il numero* $A_{ij} = (-1)^{i+j} M_{ij}$*.*

Il complemento algebrico di un elemento coincide con il minore complementare se $(i+j)$ è un numero pari, altrimenti ha il segno cambiato.

Siamo ora in grado di dare la definizione di determinante.

Definizione 3.2.3 *1. Se* A *è una matrice quadrata di ordine 1, allora* A *è della forma* $A = [a]$*. Si definisce* determinante *di* A *(in simboli* $\det A$ *o* $|A|$*) ponendo:* $\det A = a$*.*

2. Se A *una matrice quadrata di ordine* 2*, allora* A *è della forma* $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$*. Si definisce il determinante di* A *ponendo:* $\det A = ad - bc$*.*

3. *Consideriamo ora matrici di ordine superiore. Per definire il determinante di una matrice quadrata di ordine $n \geq 3$, si procede in questo modo: si sceglie una riga qualsiasi della matrice e si moltiplicano gli elementi di quella riga per i rispettivi complementi algebrici, sommando quindi i risultati.*
   In formula:
$$|A| = \sum_{k=1}^{n} a_{ik} A_{ik}.$$

Osservazione 3.2.1 Una prima osservazione importante è la seguente: nella definizione precedente il determinante di A è sviluppato scegliendo una riga di riferimento, che è la i-esima. Ma si dimostra che in realtà il determinante è *indipendente* dalla riga lungo la quale lo si sviluppa. In altre parole,

$$|A| = \sum_{k=1}^{n} a_{ik} A_{ik} = \sum_{k=1}^{n} a_{jk} A_{jk},$$

per ogni i, j.
Una seconda osservazione, questa di facile verifica, è che se A è diagonale, $A = \text{diag}$ $(\lambda_1, \lambda_2, ...\lambda_n)$, allora det $A = \lambda_1 \cdot \lambda_2 \cdot ...\lambda_n$.

Esempio 3.2.1 Calcoliamo il determinante delle matrici:

1. $A = \begin{pmatrix} 2 & 3 \\ 1 & 4 \end{pmatrix}$;
2. $B = \begin{pmatrix} 1 & 0 & 2 \\ 3 & 4 & 5 \\ 5 & 6 & 7 \end{pmatrix}$.

   Il determinante di A (matrice quadrata 2×2) si calcola facendo direttamente riferimento alla prima parte della definizione: det $A = 8 - 3 = 5$. Per la matrice B, calcoliamo il determinante sviluppando la formula secondo la prima riga, cominciando quindi a scrivere i minori complementari degli elementi di questa riga: $|M_{11}| = 28 - 30 = -2$, $|M_{12}| = 21 - 25 = -4$, $|M_{13}| = 18 - 20 = -2$. Per calcolare il determinante, sommiamo infine i prodotti dei complementi algebrici (cioè i minori complementari calcolati prima cambiati di segno se la somma degli indici è dispari) per gli elementi corrispondenti: det $A = a_{11}|M_{11}| - a_{12}|M_{12}| + a_{13}|M_{13}| = -2 - 0 - 4 = -6$.

Proprietà dei determinanti

Questo paragrafo è dedicato ad un'altra lunga e noiosa, ma indispensabile, lista di proprietà. Riguarda ovviamente quelle del determinante.

1. Scambiando due righe di A, il determinante di A cambia segno;
   Provare a capire perché, scambiando ad esempio la prima riga con la seconda.

2. Moltiplicando una riga per λ costante, allora il determinante risulta moltiplicato per λ;
   Basta sviluppare il determinante rispetto a quella riga, raccogliendo λ.
3. $\det I = 1$;
4. $\det (A^t) = \det A$;
5. Se una riga della matrice A è nulla, $\det A = 0$;
   Ovvio sviluppando il determinante rispetto a quella riga.
6. Se A ha due righe uguali, $\det A = 0$;
   Conseguenza della 1. Scambiando le due righe il determinante cambia segno, ma la matrice, dopo lo scambio, è la stessa e dunque il determinante deve essere zero.
7. Se A ha due righe proporzionali, $\det A = 0$;
   Conseguenza della 2 e della 6.
8. $\det (\lambda A) = \lambda^n \det A$, se A è una matrice di ordine n e λ una costante;
   Conseguenza della 2, tenendo conto che λA si ottiene moltiplicando per λ tutte le righe di A.
9. (Teorema di Binet) $\det (A \cdot B) = \det A \cdot \det B$.

Esempio 3.2.2 1. La matrice $A = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \\ 4 & 3 & 3 \end{pmatrix}$ ha determinante nullo perché ha due righe uguali.

2. La matrice $A = \begin{pmatrix} 0 & 5 & 3 \\ 8 & 6 & 6 \\ 4 & 3 & 3 \end{pmatrix}$ ha determinante nullo perché ha due righe proporzionali: la seconda riga è il doppio della terza.
3. $B = \begin{pmatrix} 3 & 0 & 6 \\ 9 & 12 & 15 \\ 15 & 18 & 21 \end{pmatrix}$. Questa matrice può essere scritta come $B = 3 \begin{pmatrix} 1 & 0 & 2 \\ 3 & 4 & 5 \\ 5 & 6 & 7 \end{pmatrix}$ e quindi possiamo calcolare il determinante della matrice $B' = \begin{pmatrix} 1 & 0 & 2 \\ 3 & 4 & 5 \\ 5 & 6 & 7 \end{pmatrix}$ che moltiplicato per $3^3 = 27$ (proprietà 8) ci dà il determinante di B. Dato che il determinante di B' è -6 (vedi Esempio 3.2.1), quello di B è -162.

Osservazione 3.2.2 Osserviamo:

1. La proprietà 4 implica che si può calcolare il determinante sviluppandolo lungo una colonna invece che lungo una riga;
2. Il teorema di Binet ha come conseguenza il fatto che una matrice invertibile ha determinante non nullo. Infatti se A è invertibile allora $A \cdot A^{-1} = I$, quindi:
   $$1 = \det I = \det (A \cdot A^{-1}) = \det A \cdot \det (A^{-1}),$$
   da cui si ricava che $\det A \neq 0$ e $\det (A^{-1}) = \frac{1}{\det A}$. Questa condizione di invertibilità non è solo necessaria ma anche sufficiente, come vediamo subito.

Supponendo che A sia la matrice dei coefficienti di un sistema lineare, possiamo osservare che le proprietà 6 e 7 fanno sospettare che l'annullarsi del determinante sia un buon indicatore della presenza nel sistema di equazioni "anomale". Infatti, due righe uguali, o proporzionali, indicano che nel sistema abbiamo due equazioni uguali (o proporzionali), almeno per quanto riguarda il primo membro: se anche il secondo è uguale, o mantiene la stessa costante di proporzionalità, allora un'equazione può essere gettata, se invece il secondo membro non corrisponde, le equazioni sono incompatibili. Dunque possiamo aspettarci che il determinante sia uno strumento adatto a "testare" un sistema lineare, e ad indicare se possiede soluzioni. Le cose stanno proprio così, ed è quanto vedremo in seguito. Sempre la nozione di determinante sarà preziosa anche per trattare il caso in cui il numero di equazioni è differente dal numero delle incognite, anche se in questo caso non si può applicare l'idea di determinante alla matrice dei coefficienti, che non è più quadrata.

3.3 Calcolo dell'inversa

Una matrice quadrata A è invertibile se e solo se det $A \neq 0$. Sia infatti C la matrice costituita dagli elementi $c_{ij} = A_{ij}$ (complementi algebrici della matrice A). Vale allora la seguente formula:

$$A \cdot C^t = C^t \cdot A = (\det A)I.$$

Dunque la matrice inversa di A ha come generico elemento α_{ij}:

$$\alpha_{ij} = \frac{A_{ji}}{\det A}.$$

L'elemento generico ij della matrice AC^t è $\sum_k a_{ik}c_{kj} = \sum_k a_{ik}A_{jk}$. Ora questa espressione rappresenta il determinante della matrice ottenuta da A sostituendo in essa la j-esima riga con la i-esima. Ma allora questa nuova matrice ha determinante nullo, se $j \neq i$, perché ha due righe uguali, mentre se $i = j$ la matrice coincide con la matrice A quindi l'espressione non è altro che det A. Analogamente per C^tA , con scambio di ruolo fra righe e colonne. Un consiglio: provare a verificare quanto detto in una generica matrice 3×3, su un elemento extra diagonale, per esempio l'elemento che sta nella prima riga e seconda colonna.

Esempio 3.3.1 Se $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ con $ad - bc \neq 0$, $A^{-1} = \frac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$

Esempio 3.3.2 Calcoliamo la matrice inversa di $A = \begin{pmatrix} 1 & 1 & 2 \\ 2 & -1 & 2 \\ 4 & 1 & 4 \end{pmatrix}$.

Si ha che det $A = 6 \neq 0$ quindi A è invertibile. Calcoliamo i complementi algebrici degli elementi di A: $A_{11} = -6$, $A_{12} = 0$, $A_{13} = 6$, $A_{21} = -2$, $A_{22} = -4$, $A_{23} = 3$, $A_{31} = 4$, $A_{32} = 2$, $A_{33} = -3$. Costruiamo adesso la matrice che ha come elementi i complementi algebrici che abbiamo calcolato. $C = \begin{pmatrix} -6 & 0 & 6 \\ -2 & -4 & 3 \\ 4 & 2 & -3 \end{pmatrix}$, scriviamo la trasposta $C^t = \begin{pmatrix} -6 & -2 & 4 \\ 0 & -4 & 2 \\ 6 & 3 & -3 \end{pmatrix}$. La matrice inversa di A è $A^{-1} = \frac{1}{\det A} C^t =$ $\frac{1}{6}\begin{pmatrix} -6 & -2 & 4 \\ 0 & -4 & 2 \\ 6 & 3 & -3 \end{pmatrix} = \begin{pmatrix} -1 & -\frac{1}{3} & \frac{2}{3} \\ 0 & -\frac{2}{3} & \frac{1}{3} \\ 1 & \frac{1}{2} & -\frac{1}{2} \end{pmatrix}$.

Provate a verificare il risultato calcolando il prodotto $A \cdot A^{-1}$.

3.4 La regola di Cramer

La determinazione della matrice inversa ha come applicazione la ricerca delle soluzioni di un sistema $n \times n$. Infatti dato $A\overrightarrow{x} = \overrightarrow{b}$, un sistema di equazioni con matrice dei coefficienti quadrata di ordine n, allora, se A è invertibile, la soluzione del sistema è data da: $\overrightarrow{x} = A^{-1} \cdot \overrightarrow{b} = \frac{1}{|A|} C^t \overrightarrow{b}$, cioè $x_j = \frac{1}{|A|} \sum_{k=1}^{n} c_{jk} b_k = \frac{1}{|A|} \sum_{k=1}^{n} b_k A_{kj} = \frac{\det\,(B_j)}{|A|}$ dove con B_j si indica la matrice ottenuta da A mettendo al posto della colonna j-esima il vettore $\overrightarrow{b}$ dei termini noti.

Questa è indicata come *regola di Cramer.*

Come nel caso della costruzione della matrice inversa, anche per capire la regola di Cramer è importante capire che l'espressione $\sum_{k=1}^{n} b_k A_{kj}$ non è altro che il determinante della matrice B_j, sviluppato secondo la jesima colonna.

Esempio 3.4.1 Calcoliamo, se possibile, la soluzione del seguente sistema lineare usando la regola di Cramer:

$$\begin{cases} x + y + z = 11 \\ 2x - y + z = 5 \\ 3x + 2y + z = 24 \end{cases} .$$

Per vedere se è applicabile la regola di Cramer, calcoliamo il determinante della matrice A dei coefficienti del sistema: $A = \begin{pmatrix} 1 & 1 & 1 \\ 2 & -1 & 1 \\ 3 & 2 & 1 \end{pmatrix}$, $\det A = (-1-2)-(2-3)+(4+3) = 5$ (calcolato secondo la prima riga): il determinante è diverso da zero, quindi possiamo applicare la regola. Dunque si ha

$5x = \begin{vmatrix} 11 & 1 & 1 \\ 5 & -1 & 1 \\ 24 & 2 & 1 \end{vmatrix} = 11(-1-2) - (5-24) + (10+24)$, da cui $x = 4$;

La matrice al numeratore della frazione è stata ottenuta sostituendo alla prima colonna della matrice A la colonna dei termini noti del sistema. Analogamente, per calcolare y e z, sostituiamo, rispettivamente nella matrice A, la seconda colonna e la terza colonna con quella dei termini noti. Si ha allora

$5y = \begin{vmatrix} 1 & 11 & 1 \\ 2 & 5 & 1 \\ 3 & 24 & 1 \end{vmatrix}$, da cui $y = \frac{25}{5} = 5$;

e $5z = \begin{vmatrix} 1 & 1 & 11 \\ 2 & -1 & 5 \\ 3 & 2 & 24 \end{vmatrix}$ da cui $z = \frac{10}{5} = 2$;

La soluzione del sistema è dunque:

$$\begin{cases} x = 4 \\ y = 5 \\ z = 2 \end{cases} .$$

Esercizio 3.4.1 Verificare l'esattezza della soluzione trovata nell'esempio precedente.

3.5 Approfondimenti: Prodotto vettore e prodotto misto

In questo approfondimento vediamo una nuova definizione di prodotto tra vettori di $\mathbb{R}^3$. Ricordiamo che abbiamo già definito un prodotto tra vettori, chiamato prodotto scalare perché il risultato fornisce un numero reale. Ora invece definiamo un prodotto tra due vettori che come risultato fornisce un terzo vettore, e che per questo è chiamato prodotto vettoriale. Per la sua definizione, si utilizza un determinante, il che spiega perché viene messo qui.
Siano allora dati due vettori $\overrightarrow{u} = (u_1, u_2, u_3)$, $\overrightarrow{v} = (v_1, v_2, v_3)$ di $\mathbb{R}^3$.

Definizione 3.5.1 *Si chiama* prodotto vettore *tra* $\overrightarrow{u}$ *e* $\overrightarrow{v}$, *e si indica con* $\overrightarrow{u} \wedge \overrightarrow{v}$, *il vettore ottenuto (formalmente) come determinante della seguente matrice:*

$$\begin{pmatrix} \overrightarrow{i} & \overrightarrow{j} & \overrightarrow{k} \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{pmatrix} .$$

Sviluppando il calcolo, si ottiene allora:

$$\overrightarrow{u} \wedge \overrightarrow{v} = (u_2 v_3 - u_3 v_2)\overrightarrow{i} + (u_3 v_1 - u_1 v_3)\overrightarrow{j} + (u_1 v_2 - u_2 v_1)\overrightarrow{k} .$$

Dalla definizione discendono in maniera immediata le seguenti proprietà, significative per $\overrightarrow{u}$, $\overrightarrow{v}$ vettori non nulli:

1. $\overrightarrow{u} \wedge \overrightarrow{u} = \overrightarrow{0}$; più generalmente, se $\overrightarrow{u}$, $\overrightarrow{v}$ sono paralleli, $\overrightarrow{u} \wedge \overrightarrow{v} = \overrightarrow{0}$;
   La matrice ha due righe proporzionali, quindi il suo determinante è nullo;
2. $\overrightarrow{u} \wedge \overrightarrow{v} = -(\overrightarrow{v} \wedge \overrightarrow{u})$;
   Scambio di due righe;

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

Dalla 3. segue subito che $\overrightarrow{u} \wedge \overrightarrow{v}$ è un vettore normale ai vettori $\overrightarrow{u}$ e $\overrightarrow{v}$ (infatti $(\overrightarrow{u} \wedge \overrightarrow{v}) \cdot \overrightarrow{u} = 0$ perché il determinante ha due righe uguali). Quindi la definizione di prodotto vettore permette di identificare un vettore dello spazio perpendicolare a due vettori dati (e quindi al piano generato dai due vettori).

Il prodotto $(\overrightarrow{u} \wedge \overrightarrow{v}) \cdot \overrightarrow{w}$ che interviene nella 3. di sopra viene chiamato *prodotto misto*. Dati i vettori $\overrightarrow{u}, \overrightarrow{v}, \overrightarrow{w}$, di cui $\overrightarrow{u}, \overrightarrow{v}$ non proporzionali, la matrice formata mettendoli come righe ha dunque determinante nullo se $\overrightarrow{w}$ ha direzione normale a quella della direzione di $\overrightarrow{u} \wedge \overrightarrow{v}$, cioè se e solo se $\overrightarrow{w}$ è complanare con $\overrightarrow{u}, \overrightarrow{v}$ il che ci mostra in un altro modo che il determinante è zero se e solo se i tre vettori sono linearmente dipendenti.

Vediamo ora alcune proprietà del prodotto vettore, la cui dimostrazione è lasciata per esercizio (le prime due seguono facilmente dalle varie proprietà dei determinanti, la terza si verifica con noioso calcolo diretto).

Proposizione 3.5.1 *Valgono le proprietà seguenti:*

- $\overrightarrow{u} \wedge (\overrightarrow{v} + \overrightarrow{w}) = \overrightarrow{u} \wedge \overrightarrow{v} + \overrightarrow{u} \wedge \overrightarrow{w}$ *(proprietà distributiva);*
- $a(\overrightarrow{u} \wedge v) = a\overrightarrow{u} \wedge \overrightarrow{v}$*;*
- $||\overrightarrow{u} \wedge \overrightarrow{v}||^2 = ||\overrightarrow{u}||^2||\overrightarrow{v}||^2 - (\overrightarrow{u} \cdot \overrightarrow{v})^2$.

Osserviamo esplicitamente che il prodotto vettore non gode della proprietà associativa; ad esempio

$$\overrightarrow{i} \wedge (\overrightarrow{i} \wedge \overrightarrow{j}) \neq (\overrightarrow{i} \wedge \overrightarrow{i}) \wedge \overrightarrow{j}.$$

Un vettore non nullo, come ben si sa, è caratterizzato da un modulo, una direzione, un verso. Sopra abbiamo visto quale sia la direzione di $\overrightarrow{u} \wedge \overrightarrow{v}$: è la retta perpendicolare al piano che contiene i vettori $\overrightarrow{u}, \overrightarrow{v}$. Nella prossima proposizione sono descritti il modulo ed il verso di $\overrightarrow{u} \wedge \overrightarrow{v}$.

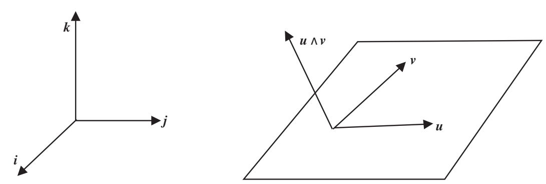
Proposizione 3.5.2 *Siano dati i vettori $\overrightarrow{u}, \overrightarrow{v} \in \mathbb{R}^3$, e sia $0 \leq \theta \leq \pi$ l'angolo da essi formato. Si ha allora:*

- $|\overrightarrow{u} \wedge \overrightarrow{v}| = |\overrightarrow{u}||\overrightarrow{v}| \sin\theta$*;*
- *In un sistema di coordinate destrorse, il verso di $\overrightarrow{u} \wedge \overrightarrow{v}$ si determina con la regola della mano destra: se $\overrightarrow{u}$ si porta con una rotazione su $\overrightarrow{v}$, in modo che le dita della mano destra diano il senso della rotazione piegandosi, allora il pollice della mano destra indica il verso di $\overrightarrow{u} \wedge \overrightarrow{v}$.*

Dimostrazione. Per determinare il verso di $\overrightarrow{u} \wedge \overrightarrow{v}$, consideriamo un sistema di assi coordinati in cui mettiamo l'origine nel punto di applicazione dei due vettori e, ad esempio, il verso positivo dell'asse delle x nel verso di $\overrightarrow{u}$. In altre parole, $\overrightarrow{u} = a\overrightarrow{i}$, con $a > 0$. Poi scegliamo allo stesso modo, nel piano determinato dai vettori $\overrightarrow{u}$, $\overrightarrow{v}$, la direzione dell'asse y in modo che $\overrightarrow{v} = b\overrightarrow{i} + c\overrightarrow{j}$, con $c > 0$ ($b \in \mathbb{R}$). Si ha allora che

$$\overrightarrow{u} \wedge \overrightarrow{v} = \begin{pmatrix} \overrightarrow{i} & \overrightarrow{j} & \overrightarrow{k} \\ a & 0 & 0 \\ b & c & 0 \end{pmatrix} = ac\overrightarrow{k}.$$

Poiché $ac > 0$, allora $\overrightarrow{u} \wedge \overrightarrow{v}$ ha la stessa direzione dell'asse positivo delle z, e questo corrisponde esattamente alla regola della mano destra. Per quanto riguarda la formula della lunghezza di $\overrightarrow{u} \wedge \overrightarrow{v}$, questa si ottiene con un semplicissimo calcolo a partire dall'identità mostrata nel terzo punto della Proposizione 3.5.1. ■



La regola della mano destra

Figura 3.1

Una conseguenza importante della proposizione precedente è:

Proposizione 3.5.3 $|\overrightarrow{u} \wedge \overrightarrow{v}|$ *rappresenta l'area del parallelogramma individuato da* $\overrightarrow{u}$ *e* $\overrightarrow{v}$.

Analogamente si può dimostrare che il prodotto misto $\overrightarrow{u} \wedge \overrightarrow{v} \cdot \overrightarrow{w}$ rappresenta il volume del parallelepipedo generato dai tre vettori.

Il prodotto vettore compare ad esempio in analisi, quando si definisce l'area di una superficie. Questa in genere viene definita come l'immagine, tramite una funzione $\overrightarrow{r}$ definita su un dominio del piano T, ed a valori in $\mathbb{R}^3$: $\Sigma = \overrightarrow{r}(T)$, $T \subset \mathbb{R}^2$. Allora

$$\text{Area}\ (\Sigma) \stackrel{\text{def}}{=} \int_T |\overrightarrow{r}_x \wedge \overrightarrow{r}_y|\, dxdy.$$

Qui $\overrightarrow{r}_x$, $\overrightarrow{r}_y$ rappresentano le derivate parziali di $\overrightarrow{r}$ ed il fattore $|\overrightarrow{r}_x \wedge \overrightarrow{r}_y|$ è un fattore di ingrandimento per l'area, nella trasformazione, tramite $\overrightarrow{r}$, dall'elemento di area (piccolo!) $dxdy$ di T nel corrispondente elemento di superficie (un rettangolo che diventa, nella trasformazione, un parallelogramma [in maniera approssimata]).

Vediamo ora un paio di applicazioni dei concetti precedenti alla geometria analitica.

Una retta in $\mathbb{R}^3$ passante per il punto $P_0 = (x_0, y_0, z_0)$ è data, in forma generale, dall'intersezione di due piani π, ρ:

$$\begin{cases} \pi : a(x-x_0) + b(y-y_0) + c(z-z_0) = 0 \\ \rho : \hat{a}(x-x_0) + \hat{b}(y-y_0) + \hat{c}(z-z_0) = 0, \end{cases}$$

ove $\overrightarrow{n} = (a, b, c)$ e $\overrightarrow{\hat{n}} = (\hat{a}, \hat{b}, \hat{c})$ sono due normali ai piani π e ρ rispettivamente. I vettori normali sono perpendicolari ai due piani. Un vettore perpendicolare ad essi giace allora su entrambi i piani e quindi determina la direzione della retta. Sapendo che $\overrightarrow{n} \wedge \overrightarrow{\hat{n}}$ è perpendicolare ad $\overrightarrow{n}$ e a $\overrightarrow{\hat{n}}$, e che la retta passa per P_0, si ha dunque che la retta si scrive, indicando con P un generico punto della stessa:

$$P - P_0 = t(\overrightarrow{n} \wedge \overrightarrow{\hat{n}}),$$

che è la forma parametrica della retta.

Esempio 3.5.1 Determiniamo la forma parametrica della retta in forma generale:

$$r : \begin{cases} 2x + y + z = 4 \\ x - y = 0 \end{cases} \quad ;$$

Due vettori normali ai piani sono $(2, 1, 1)$ e $(1, -1, 0)$. Facciamo il loro prodotto vettore:

$$\begin{vmatrix} \overrightarrow{i} & \overrightarrow{j} & \overrightarrow{k} \\ 2 & 1 & 1 \\ 1 & -1 & 0 \end{vmatrix} = \overrightarrow{i} + \overrightarrow{j} - 3\overrightarrow{k}.$$

Poiché la retta contiene il punto $(1, 1, 1)$, ne segue che la sua forma parametrica è:

$$\begin{cases} x = 1 + t \\ y = 1 + t \\ z = 1 - 3t \end{cases} \quad .$$

Supponiamo ora di voler trovare il piano che passa per i tre punti non allineati $P_0 = (x_0, y_0, z_0)$, $P_1 = (x_1, y_1, z_1)$, $P_2 = (x_2, y_2, z_2)$. Detto $P = (x, y, z)$ un generico punto del piano, si ha che i tre vettori: $P - P_0$, $P_1 - P_0$, $P_2 - P_0$ devono giacere sullo stesso piano. Quindi il loro prodotto misto deve essere nullo. Si ottiene quindi:

$$\begin{vmatrix} x - x_0 & y - y_0 & z - z_0 \\ x_1 - x_0 & y_1 - y_0 & z_1 - z_0 \\ x_2 - x_0 & y_2 - y_0 & z_2 - z_0 \end{vmatrix} = 0.$$

Esempio 3.5.2 Troviamo l'equazione del piano passante per i punti $(0, 1, 0)$, $(0, 0, 1)$ e $(1, 0, 0)$. Il determinante di sopra diventa in questo caso:

$$\begin{vmatrix} x & y-1 & z \\ 0 & -1 & 1 \\ 1 & -1 & 0 \end{vmatrix},$$

che, posto uguale a zero, dà l'equazione $x + y + z = 1$.

Col prodotto vettore si trova facilmente l'area di un triangolo di vertici A, B, C nello spazio. Basta infatti considerare i vettori $B - A$ e $C - A$ (ad esempio) e calcolare il modulo del loro prodotto vettore. Quel che si ottiene è il doppio dell'area cercata (provare a capire perché, è facile).

Esempio 3.5.3 Calcoliamo l'area del triangolo che ha vertici in $(1, 0, 0)$, $(0, 1, 0)$, $(0, 0, 1)$. Scriviamo il determinante:

$$\begin{vmatrix} \overrightarrow{i} & \overrightarrow{j} & \overrightarrow{k} \\ 1 & 0 & -1 \\ 0 & 1 & -1 \end{vmatrix} = \overrightarrow{i} + \overrightarrow{j}.$$

Quindi l'area del triangolo vale $\frac{\sqrt{2}}{2}$.

I prossimi sono due esempi di come il prodotto vettore ed il prodotto misto servano ad identificare utili grandezze in fisica.

Esempio 3.5.4 Prodotto vettore Data una forza $\overrightarrow{f}$ applicata in un punto P, il suo momento M_O rispetto al punto O è il prodotto vettore $M_O = (P - O) \wedge \overrightarrow{f}$; dalla definizione di prodotto vettore segue allora che il modulo del momento M_O è il prodotto $M_O = fb$ della forza per il suo braccio (cioè la distanza della retta di applicazione della forza dal punto O.)

Esempio 3.5.5 Prodotto misto Data una forza $\overrightarrow{f}$ applicata in P, ed un asse u, chiamiamo $\overrightarrow{u}$ il versore dell'asse, cioè un vettore diretto come l'asse e di modulo 1. Si definisce *momento della forza rispetto all'asse* il prodotto misto

$$M_u = (P - O) \wedge \overrightarrow{f} \cdot \overrightarrow{u},$$

dove O è un qualunque punto dell'asse:il momento rispetto all'asse è quindi la componente sull'asse del momento M_O della forza rispetto ad un punto dell'asse.
Utilizzando la definizione di prodotto scalare e vettore, si può dimostrare (provare per esercizio) che:

1. M_u è indipendente dalla scelta del punto O sull'asse, cioè preso un altro punto O' si ha $M_{O'} \neq M_O$, ma $M_{O'} \cdot \overrightarrow{u} = M_O \cdot \overrightarrow{u}$ comunque si prendano O e O' sull'asse.
2. $M_u = fb$, essendo b il braccio della forza, cioè la distanza della retta di applicazione della forza dall'asse.

3. Se f è parallela all'asse, o incidente sull'asse (cioè la sua retta di applicazione passa per l'asse) allora $M_u = 0$ (come è noto dall'esperienza comune, una porta girevole attorno ad un asse rimane ferma applicando ad essa una forza di qualunque intensità purché parallela o incidente all'asse della porta, mentre si mette in moto se le si applica una forza sghemba rispetto all'asse, e tanto più velocemente quanto più la forza è applicata lontano dall'asse,cioè con un braccio elevato.)

3.6 Esercizi

3.6.1 Quesiti a risposta chiusa

Rispondere vero o falso.

1. Date le matrici $A = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{pmatrix}$ e $B = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{pmatrix}$:
   a) BA è invertibile;
   b) $(BA)^{-1} = B^{-1}A^{-1}$;
   c) $AB = BA$.
2. Se A è una matrice $n \times m$ e A^t è la sua trasposta allora $A^t \cdot A$ è una matrice $n \times n$ simmetrica.
3. Sia A una matrice 4×4 tale che $\det A = -2$. Allora
   a) $\det(3A) = -6$;
   b) $\det(A^t \cdot A) = 4$.
4. Data $A = \begin{pmatrix} -t & 1 & 2 \\ 1 & -t & 1 \\ 2 & 2 & 1-t \end{pmatrix}$:
   a) per $t = -1$ non è invertibile;
   b) la sua inversa per $t = 0$ è $A^{-1} = \frac{1}{5}\begin{pmatrix} -2 & 3 & 1 \\ 1 & -4 & 2 \\ 2 & 2 & -1 \end{pmatrix}$.
5. La matrice $A = \begin{pmatrix} 1 & 2 & 0 \\ 3 & 6 & -1 \\ 2 & 4 & 0 \end{pmatrix}$ è invertibile;
6. Il vettore $(1, t, 1)$ giace sul piano che contiene i vettori $(1, 0, 0)$ e $(2, 0, 2)$ per $t = -1$;
7. Tre vettori non nulli sono complanari se e solo se il loro prodotto misto è nullo;
8. $\overrightarrow{u} \wedge \overrightarrow{v} = \overrightarrow{0}$ e $\overrightarrow{u} \cdot \overrightarrow{v} = 0$ implica o $\overrightarrow{u}$ o $\overrightarrow{v}$ sono il vettore nullo;
9. Dati due punti A, B distinti i punti dello spazio che verificano $\overrightarrow{P - A} \wedge \overrightarrow{P - B} = \overrightarrow{0}$ descrivono la retta per A e B.

3.6.2 Esercizi aperti

Esercizio 3.6.1 Calcolare, se possibile, la somma tra le seguenti matrici:

1. $\begin{pmatrix} 1 & 2 & -3 & 4 \\ 0 & -5 & 1 & -1 \end{pmatrix} + \begin{pmatrix} 3 & -5 & 6 & -1 \\ 2 & 0 & -2 & -3 \end{pmatrix}$;
2. $\begin{pmatrix} 1 & 2 & -3 & 4 \\ 0 & -4 & 1 & -1 \end{pmatrix} + \begin{pmatrix} 3 & 5 \\ 1 & -2 \end{pmatrix}$;
3. $\begin{pmatrix} 1 & 2 & -1 & 0 \\ 4 & 0 & 2 & 1 \\ 2 & -5 & 1 & 2 \end{pmatrix} - \begin{pmatrix} 3 & -4 & 1 & 2 \\ 1 & 5 & 0 & 3 \\ 2 & -2 & 3 & -1 \end{pmatrix}$.

Esercizio 3.6.2 Calcolare, se possibile, il prodotto tra le seguenti matrici:

1. $\begin{pmatrix} 4 & 5 & 6 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix}$;
2. $\begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix} \cdot \begin{pmatrix} 4 & 5 & 6 \end{pmatrix}$;
3. $\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & -1 \end{pmatrix} \cdot \begin{pmatrix} 0 & -1 & -2 \\ 4 & 2 & 3 \\ 1 & 1 & 1 \end{pmatrix}$.
4. Date le matrici $A = \begin{pmatrix} 1 & -1 & 2 \\ 0 & 3 & 4 \end{pmatrix}$, $B = \begin{pmatrix} 4 & 0 & -3 \\ -1 & -2 & 3 \end{pmatrix}$ e $C = \begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix}$, calcolare
   a) $3A - 4B$;
   b) AC;
   c) A^t;
   d) $C^t \cdot A^t$; $C^t \cdot C$; $C \cdot C^t$.

Esercizio 3.6.3 Calcolare, se possibile, il determinante delle seguenti matrici:

1. $\begin{pmatrix} 2 & 3 \\ 1 & 4 \end{pmatrix}$;
2. $\begin{pmatrix} 1 & 0 & 2 \\ 3 & 4 & 5 \\ 5 & 6 & 7 \end{pmatrix}$;

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

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

Esercizio 3.6.4 Calcolare, se possibile, la soluzione dei seguenti sistemi lineari utilizzando la regola di Cramer

1. $\begin{cases} 3x + y - z = -2 \\ 5y + 3z = -1 \\ 7y - 2z = 1 \end{cases}$;

2. $\begin{cases} x - y + 2z = -1 \\ 2x + y - z = -3 \\ x - 4y + 7z = 2 \end{cases}$;

3. $\begin{cases} x + y + 2z = -1 \\ 2x - y + 2z = -4 \\ 4x + y + 4z = -2 \end{cases}$.

Esercizio 3.6.5 Mostrare che $|\overrightarrow{u} \wedge \overrightarrow{v}| = |\overrightarrow{u}||\overrightarrow{v}|$ se e solo se $\overrightarrow{u}$ e $\overrightarrow{v}$ sono complanari.

Esercizio 3.6.6 Dati tre punti distinti A, B nello spazio, dare una condizione, tramite il prodotto vettore, per verificare se sono allineati.

Esercizio 3.6.7 Trovare il piano che contiene l'asse x e il punto $(1, 1, 1)$

Esercizio 3.6.8 Utilizzando il prodotto vettore:

- Determinare la forma parametrica della retta di forma generale $\begin{cases} x + y - z = 1 \\ x + y + z = 0 \end{cases}$;
- determinare il piano che contiene il punto $(1, 1, 1)$ e la retta di equazioni parametriche: $\begin{cases} x = 1 + t \\ y = 1 + t \\ z = t \end{cases}$.

Nel primo caso trovare un punto della retta (es $(1, 0, -1)$), nel secondo caso trovare due punti della retta (es $(1, 1, 0)$ e $(2, 2, 1)$).

4

Spazi Vettoriali

Abbiamo introdotto l'insieme dei vettori (del piano e dello spazio) ed alcune operazioni fra di loro. È ora il momento di mettere in evidenza alcune proprietà strutturali dell'insieme dei vettori. Questo insieme, piuttosto naturalmente, viene chiamato *spazio vettoriale*. Come spesso succede, l'individuazione di certe proprietà specifiche ed astratte permette poi di rendersi conto che anche altri insiemi, non costituiti da vettori geometrici, ma da oggetti più complessi, hanno la stessa struttura di spazio vettoriale. In queste note comunque quel che più ci interessa è utilizzare la struttura vettoriale per discutere la geometria dell'insieme delle soluzioni di un sistema lineare. Prima di cominciare conviene precisare che il termine *vettore* può indicare:

1. I vettori geometrici del piano o dello spazio;
2. Gli elementi di $\mathbb{R}^n$, ossia le n-uple ordinate di numeri reali che scriviamo come vettori colonna: $\overrightarrow{x} = \begin{pmatrix} x_1 \\ \cdot \\ \cdot \\ x_n \end{pmatrix}$, o come vettori riga, interpretati come trasposti di vettori colonna, $\overrightarrow{x}^t = (x_1, \ldots, x_n)$;
3. Oggetti più generali che fanno parte di un insieme che gode di proprietà che ora elenchiamo.

Definizione 4.0.1 *Uno* spazio vettoriale V *sul campo* $\mathbb{R}$ *è un insieme non vuoto nel quale sono definite:*

1. *Un'operazione interna*[1] *di somma che gode delle proprietà associativa, commutativa, che ha l'elemento neutro (il vettore nullo, indicato con* $\overrightarrow{0}$*) e tale che ad ogni elemento associa un altro elemento, detto opposto, con la proprietà che la somma dei due dà l'elemento neutro;*

[1] Operazione interna ad un insieme vuol dire che l'operazione fra due elementi dell'insieme fornisce ancora un elemento dell'insieme

2. *Un'operazione che associa ad una coppia di elementi* $(\overrightarrow{v}, r) \in V \times \mathbb{R}$ *un elemento di* V *(operazione di prodotto per uno scalare) che gode della proprietà distributiva rispetto alla somma di scalari e vettori* $((a+b)\overrightarrow{v} = a\overrightarrow{v} + b\overrightarrow{v}$ *e* $a(\overrightarrow{v_1} + \overrightarrow{v_2}) = a\overrightarrow{v_1} + a\overrightarrow{v_2})$ *e tale che* $a(b\overrightarrow{v}) = (ab)\overrightarrow{v}$ *;* $1 \cdot \overrightarrow{v} = \overrightarrow{v}$.

Come spesso succede, la definizione è un po' astratta e "calata" dall'alto. Per capirla meglio, vediamo ora qualche semplice esempio di spazio vettoriale, in modo da familiarizzarsi con quest'idea fondamentale.

Osservazione 4.0.1 La definizione appena data può essere estesa a spazi vettoriali definiti sul campo $\mathbb{C}$ invece che su $\mathbb{R}$. Gli elementi del campo sono in ogni caso detti scalari, come già menzionato.

1. Il più piccolo spazio vettoriale è lo spazio vettoriale nullo, formato dal solo vettore $\overrightarrow{0}$.
   L'insieme $\{\overrightarrow{0}\}$ soddisfa banalmente le proprietà di spazio vettoriale.
2. L'insieme delle matrici a elementi reali di dimensione $m \times n$, che indichiamo con $M(m \times n)$, è uno spazio vettoriale con le operazioni di somma e prodotto per uno scalare definite nel capitolo precedente[2];
3. L'insieme dei polinomi, con le usuali operazioni di somma e prodotto per un numero reale;
4. L'insieme $\mathbb{C}^n$ delle n-uple di numeri complessi è uno spazio vettoriale se gli scalari si scelgono in $\mathbb{C}$.

Negli esempi precedenti si è sempre specificata quali siano le operazioni di somma e di prodotto per uno scalare considerate sull'insieme in questione. Questo è molto importante, anche se spesso le operazioni stesse sono molto naturali. Come altro esempio, se si considera l'insieme delle funzioni definite su un intervallo ed a valori reali, esso è uno spazio vettoriale se si definisce le funzioni $f + g$ e tf (t scalare) in questo modo:

$$(f+g)(x) = f(x) + g(x),$$
$$(tf)(x) = tf(x).$$

Chi legge scrupolosamente queste note dovrebbe soffermarsi sull'uso che nella formula precedente è fatto delle parentesi, e sul fatto che il + a primo membro *non* è lo stesso di quello a secondo membro: a volte i matematici, un po' perversi, usano lo stesso simbolo per cose differenti...

Definizione 4.0.2 *Un sottoinsieme non vuoto* W *di uno spazio vettoriale* V *si dice* sottospazio vettoriale di V *se:*

[2] Molte delle proprietà elencate delle operazioni fra matrici avevano esattamente lo scopo di mostrare che le operazioni introdotte godono delle proprietà richieste per dare struttura vettoriale allo spazio delle matrici. Osservare che il prodotto righe per colonne *non c'entra nulla* con la struttura vettoriale dello spazio delle matrici

1. $\overrightarrow{x}, \overrightarrow{y} \in W$ *implica* $\overrightarrow{x} + \overrightarrow{y} \in W$*;*
2. $\overrightarrow{x} \in W, t \in \mathbb{R}$ *implica* $t \cdot \overrightarrow{x} \in W$*.*

In altre parole, il sottoinsieme W, ereditando le definizioni di somma e prodotto per uno scalare da V, essendo chiuso rispetto alla somma ed al prodotto per uno scalare, ne eredita anche le proprietà, e quindi è ancora uno spazio vettoriale.

Osservazione 4.0.2 Un sottospazio contiene sempre $\overrightarrow{0}$: se $\overrightarrow{x} \in W$ allora $\overrightarrow{0} = 0 \cdot \overrightarrow{x} \in W$. Il sottospazio nullo e l'intero spazio vettoriale si dicono *sottospazi banali.* In $\mathbb{R}^3$ i sottospazi non banali sono le rette ed i piani passanti per l'origine.

Esempio 4.0.1 Se $\overrightarrow{v} \in V$, allora $\{t\overrightarrow{v} : t \in \mathbb{R}\}$ è un sottospazio di V (geometricamente visualizzabile come una retta passante per l'origine).

Esempio 4.0.2 L'insieme delle funzioni continue su un intervallo I a valori reali è un sottospazio vettoriale dello spazio di tutte le funzioni definite in I a valori reali. Anche l'insieme delle funzioni derivabili in I costituisce uno spazio vettoriale. Infatti somma di funzioni continue o derivabile è ancora funzione continua o derivabile, e questo succede anche per la moltiplicazione per uno scalare.

Osservazione 4.0.3 È facile osservare che l'intersezione di due sottospazi dello spazio V è ancora un sottospazio di V, eventualmente ridotto al sottospazio nullo. L'unione di due sottospazi di uno spazio vettoriale non è un sottospazio a meno che uno non sia incluso nell'altro (provarlo per esercizio).

Definizione 4.0.3 *Si chiama* sottospazio affine *di uno spazio vettoriale* V *un insieme* W *che è il traslato di un sottospazio vettoriale di* V*.*

In altre parole W è sottospazio affine di uno spazio vettoriale se esistono $\overrightarrow{x} \neq \overrightarrow{0}$ e U sottospazio tali che $W = \overrightarrow{x} + U$. In $\mathbb{R}^3$ i sottospazi affini sono i punti, le rette, i piani che non passano per l'origine.

Esempio 4.0.3 Consideriamo i due insiemi: $A = \{\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 : x = 0\}$ e $B = \{\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 : 2x + y = 1\}$ e vediamo se sono sottospazi di $\mathbb{R}^2$.
L'insieme A (che è costituito dai vettori di $\mathbb{R}^2$ con la prima componente nulla) è un sottospazio di $\mathbb{R}^2$ perché:

1. È non vuoto ($\begin{pmatrix} 0 \\ 0 \end{pmatrix} \in A$);
2. È chiuso rispetto alla somma: $\begin{pmatrix} 0 \\ y_1 \end{pmatrix} + \begin{pmatrix} 0 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0 \\ y_1 + y_2 \end{pmatrix} \in A$;

3. È chiuso rispetto al prodotto per uno scalare: se $\begin{pmatrix} 0 \\ y \end{pmatrix} \in A$, allora $t \begin{pmatrix} 0 \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ ty \end{pmatrix} \in A$.

B invece non è un sottospazio di $\mathbb{R}^2$ perché non è chiuso rispetto alla somma; basta ad esempio considerare i vettori: $\overrightarrow{v_1} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ e $\overrightarrow{v_2} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ entrambi appartenenti a B perché le loro componenti soddisfano la relazione $2x + y = 1$. La loro somma è il vettore $\overrightarrow{v_1} + \overrightarrow{v_2} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \notin B$. Un modo più semplice di vedere che B non è un sottospazio vettoriale è quello di accorgersi che lo zero non sta in B: $\begin{pmatrix} 0 \\ 0 \end{pmatrix} \notin B$.

La prossima definizione è molto importante.

Definizione 4.0.4 *Sia V uno spazio vettoriale e siano $\overrightarrow{v_1}, \overrightarrow{v_2}, ...\overrightarrow{v_n} \in V$ e $t_1, t_2, ...t_n$ degli scalari. Allora il vettore*

$$\overrightarrow{v} = \sum_{k=1}^{n} t_k \overrightarrow{v_k} = t_1 \overrightarrow{v_1} + t_2 \overrightarrow{v_2} + ...t_n \overrightarrow{v_n}$$

rappresenta un vettore che si chiama combinazione lineare *dei vettori $\overrightarrow{v_k}$.*

L'insieme di tutte le combinazione lineari dei vettori $\overrightarrow{v_1}, \overrightarrow{v_2}, ..., \overrightarrow{v_n} \in V$ è un sottospazio vettoriale di V che si rappresenta con $L(\overrightarrow{v_1}, \overrightarrow{v_2}, ...\overrightarrow{v_n})$. I vettori $\overrightarrow{v_k}$ si chiamano *generatori* di tale sottospazio.

Per dimostrare che $L(\overrightarrow{v_1}, \overrightarrow{v_2}, ...\overrightarrow{v_n})$ è un sottospazio, siano $\overrightarrow{v} = \sum_{k=1}^{n} t_k \overrightarrow{v_k}$, $\overrightarrow{w} = \sum_{k=1}^{n} s_k \overrightarrow{v_k}$ allora $\overrightarrow{v} + \overrightarrow{w} = \sum_{k=1}^{n} (t_k + s_k) \overrightarrow{v_k} \dots$

Se $L(\overrightarrow{v_1}, \overrightarrow{v_2}, ...\overrightarrow{v_n}) = V$ allora i vettori $\overrightarrow{v_k}$ generano lo spazio vettoriale V.

Esempio 4.0.4 Combinazione lineare di vettori Consideriamo N punti P_i, ciascuno di massa m_i e sia $m = \sum_i m_i$ la massa totale dei punti; il centro di massa (o baricentro) del sistema di punti è il punto G la cui posizione $G - O$ rispetto ad un generico punto O è definita dal vettore

$$\frac{\sum_i m_i (P_i - O)}{m}$$

ed è quindi dato dalla combinazione lineare di vettori posizione $(P_i - O)$ dei singoli punti, ciascuno moltiplicato per il fattore (positivo) $\frac{m_i}{m}$.

Definizione 4.0.5 *I vettori $\overrightarrow{v_1}, \overrightarrow{v_2}, ..., \overrightarrow{v_n}$ di uno spazio vettoriale V si dicono* linearmente dipendenti *se esistono n scalari non tutti nulli $a_1, a_2, ..., a_n$ tali che*

$$a_1\overrightarrow{v_1} + a_2\overrightarrow{v_2} + ... + a_n\overrightarrow{v_n} = \overrightarrow{0}.$$

In caso contrario, i vettori si dicono linearmente indipendenti*.*

Se $\overrightarrow{v_1}$ e $\overrightarrow{v_2}$ sono due vettori di $\mathbb{R}^3$ linearmente indipendenti, allora $L(\overrightarrow{v_1}, \overrightarrow{v_2})$ è geometricamente il piano che contiene i due vettori (che passa quindi per l'origine). L'equazione del piano passante per il punto P_0 e parallelo a $L(\overrightarrow{v_1}, \overrightarrow{v_2})$ ha allora equazione

$$P - P_0 = t\overrightarrow{v_1} + s\overrightarrow{v_2},$$

che rappresenta, come già visto, la forma parametrica del piano.

Esempio 4.0.5 Vediamo quali tra i vettori $\overrightarrow{a} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ e $\overrightarrow{b} = \begin{pmatrix} 2 \\ 2 \\ 3 \end{pmatrix}$ appartiene allo spazio generato dai vettori $\overrightarrow{v_1} = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}$ e $\overrightarrow{v_2} = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$.
Per stabilire se i due vettori dati appartengono allo spazio generato da $\overrightarrow{v_1}, \overrightarrow{v_2}$ $(L(\overrightarrow{v_1}, \overrightarrow{v_2}))$ dobbiamo vedere se è possibile scriverli come combinazione lineare di $\overrightarrow{v_1}$ e $\overrightarrow{v_2}$.

1. $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \overrightarrow{a} = c_1 \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} + c_2 \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$ è equivalente a dire

$$\begin{cases} c_1 + c_2 = 1 \\ 2c_2 = 2 \\ 2c_1 + c_2 = 3 \end{cases},$$

che è un sistema impossibile,quindi $\overrightarrow{a} \notin L$.

2. $\begin{pmatrix} 2 \\ 2 \\ 3 \end{pmatrix} = \overrightarrow{b} = c_1 \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} + c_2 \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$ cioè

$$\begin{cases} c_1 + c_2 = 2 \\ 2c_2 = 2 \\ 2c_1 + c_2 = 3 \end{cases},$$

che ha come soluzione $c_1 = c_2 = 1$ e quindi $\overrightarrow{b} \in L$.

Esempio 4.0.6 I vettori $(1,0)$ e $(1,1)$ sono linearmente indipendenti perché ogni loro combinazione lineare $a_1(1,0)+a_2(1,1)=(a_1+a_2,a_2)=(0,0)$ implica

$$\begin{cases} a_1+a_2=0 \\ a_2=0 \end{cases},$$

cioè $a_1=a_2=0$.
I vettori $(1,0)$, $(0,1)$ e $(3,4)$ sono linearmente dipendenti; infatti $a_1(1,0)+a_2(0,1)+a_3(3,4)=(a_1+3a_3,a_2+4a_3)=(0,0)$ implica

$$\begin{cases} a_1+3a_3=0 \\ a_2+4a_3=0 \end{cases},$$

cioè

$$\begin{cases} a_1=-3a_3 \\ a_2=-4a_3 \end{cases},$$

con a_3 che può assumere qualsiasi valore reale.

Esercizio 4.0.9 Siano dati n vettori $\overrightarrow{v_k}$, $1\le k\le n$.

- Se $n=1$, allora $\{\overrightarrow{v_1}\}$ è linearmente indipendente se e solo se $\overrightarrow{v_1}\neq 0$;
- Se $n=2$, $\overrightarrow{v_1},\overrightarrow{v_2}$ sono dipendenti se e solo se uno dei due è multiplo dell'altro;
- se esiste k tale che $\overrightarrow{v_k}=\overrightarrow{0}$, i vettori $\overrightarrow{v_1},\overrightarrow{v_2},...\overrightarrow{v_n}$ sono linearmente dipendenti.

Vale questo fatto molto importante:

Proposizione 4.0.1 *Se $\overrightarrow{v_1},\overrightarrow{v_2},...\overrightarrow{v_n}$ sono dipendenti almeno uno di essi è combinazione lineare degli altri.*
Sia $a_1\overrightarrow{v_1}+a_2\overrightarrow{v_2}+...+a_n\overrightarrow{v_n}=\overrightarrow{0}$, e supponiamo che uno dei coefficienti, per esempio a_1, sia diverso da zero. Allora $\overrightarrow{v_1}=-(a_2/a_1)\overrightarrow{v_2}-...-(a_n/a_1)\overrightarrow{v_n}$.

Definizione 4.0.6 *Si chiama* base *dello spazio vettoriale V un sistema di generatori linearmente indipendenti.*

Quindi un insieme $\{\overrightarrow{v_1},\overrightarrow{v_2},\dots\overrightarrow{v_n}\}$ è una base dello spazio vettoriale V se:

- $L(\overrightarrow{v_1},\overrightarrow{v_2},\dots\overrightarrow{v_n})=V$, cioè se i vettori generano lo spazio;
- sono linearmente indipendenti.

Dato un sistema di generatori di uno spazio vettoriale, ogni vettore dello spazio può, per definizione, essere scritto come combinazione lineare dei vettori del sistema. Tuttavia i coefficienti della combinazione lineare non sono, di solito, univocamente determinati. Ma, nel caso di una base, vale la seguente importante proposizione.

Proposizione 4.0.2 *Se $\overrightarrow{v_1},\overrightarrow{v_2},...\overrightarrow{v_n}$ sono una base di V, allora ogni vettore $\overrightarrow{v}$ di V si esprime in modo unico come combinazione lineare dei $\overrightarrow{v_k}$.*

Sia $\overrightarrow{v} = \sum_{k=1}^{n} t_k \overrightarrow{v_k} = \sum_{k=1}^{n} s_k \overrightarrow{v_k}$. Allora $\overrightarrow{0} = \sum_{k=1}^{n} (t_k - s_k)\overrightarrow{v_k}$, e quindi $t_k = s_k$ per ogni k, essendo gli elementi della base linearmente indipendenti.

Vale il seguente importante

Teorema 4.0.1 *Sia V uno spazio vettoriale. Allora:*

1. *Ogni insieme di generatori di V contiene una base;*
2. *Ogni insieme linearmente indipendente è contenuto in una base di V;*
3. *Tutte le basi di V hanno lo stesso numero di elementi.*

Definizione 4.0.7 *Dunque il numero di elementi di una base qualsiasi di uno spazio vettoriale V è un invariante dello spazio stesso. Possiamo allora dare un nome a questo numero, che chiameremo* dimensione *di V e si indica con* $\dim V$.

Osservazione 4.0.4 Le osservazioni che seguono sono facili conseguenze del teorema.

1. Se $\overrightarrow{w_1}, \overrightarrow{w_2}, ...\overrightarrow{w_m}$ sono linearmente indipendenti in V allora $m \leq \dim V$ e $m = \dim V$ se e solo se $\overrightarrow{w_1}, \overrightarrow{w_2}, ...\overrightarrow{w_m}$ formano una base di V;
2. Se $\overrightarrow{w_1}, \overrightarrow{w_2}, ...\overrightarrow{w_m}$ generano V, allora $m \geq \dim V$ e $m = \dim V$ se e solo se $\overrightarrow{w_1}, \overrightarrow{w_2}, ...\overrightarrow{w_m}$ sono linearmente indipendenti;
3. Se W è un sottospazio di V allora $\dim W \leq \dim V$ e $\dim W = \dim V$ se e solo se $W = V$.

Osservazione 4.0.5 La dimensione di $\mathbb{R}^n$ è n; infatti se indichiamo con $\overrightarrow{e_1}, \overrightarrow{e_2}, ...\overrightarrow{e_n}$ i vettori $\overrightarrow{e_1} = (1, 0, ..., 0), \overrightarrow{e_2} = (0, 1, 0, ...), ..., \overrightarrow{e_n} = (0, 0, ..., 1)$, questi formano una base, detta *base canonica* di $\mathbb{R}^n$.

Esempio 4.0.7 Calcoliamo una base per $V = L(\overrightarrow{v_1}, \overrightarrow{v_2})$ e $W = L(\overrightarrow{w_1}, \overrightarrow{w_2})$, gli spazi vettoriali generati dai vettori $\overrightarrow{v_1} = \begin{pmatrix} 1 \\ 0 \\ -1 \\ 2 \end{pmatrix} \quad \overrightarrow{v_2} = \begin{pmatrix} 2 \\ 1 \\ 0 \\ -1 \end{pmatrix} \overrightarrow{w_1} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \quad \overrightarrow{w_2} = \begin{pmatrix} 3 \\ 2 \\ 1 \\ 0 \end{pmatrix}$.
Con facili calcoli si vede che i due vettori $\overrightarrow{v_1}, \overrightarrow{v_2}$ sono linearmente indipendenti, ed essendo anche dei generatori di V, sono una base per V. Allo stesso modo, $\overrightarrow{w_1}, \overrightarrow{w_2}$ sono linearmente indipendenti, ed essendo anche dei generatori di W, sono una base per W.

Uno spazio vettoriale ha una struttura interessante da un punto di vista matematico perché i suoi elementi possono essere descritti in maniera semplice: dati $\{\overrightarrow{v_1}, \ldots, \overrightarrow{v_n}\}$ generatori di V, *ogni elemento di V* può essere descritto attraverso $\overrightarrow{v_1}, \ldots, \overrightarrow{v_n}$ usando *combinazioni lineari,* cioè utilizzando le semplici operazioni di somma e moltiplicazione per uno scalare (gli scalari si assumono come elementi "familiari", vedere la

Proposizione 4.0.2). Fra gli insiemi di generatori, le *basi* sono particolarmente importanti per due ragioni: la prima è che rappresentano un insieme di generatori *economico,* cioè minimale, nel senso che nessuno di essi può essere soppresso; la seconda, legata alla prima, è che ogni elemento dello spazio si scrive in maniera *unica,* quando si utilizzano gli elementi di una base.
Riassumiamo quanto detto con un esempio: $\mathbb{R}^n$ è uno spazio ad n dimensioni. Per descriverlo perciò ci bastano gli oggetti seguenti:

1. i numeri reali;
2. definire che cosa è la somma fra due vettori e che cosa significa moltiplicare un vettore per uno scalare;
3. individuarne una base, ad esempio la più naturale:
$$\{(1,0,\dots,0),(0,1,\dots,0),\dots,(0,0,\dots,1)\}.$$

In altre parole, per descrivere completamente uno spazio vettoriale, basta la conoscenza del campo degli scalari ($\mathbb{R}$), le definizioni delle operazioni di somma e di prodotto per uno scalare, ed una base dello spazio.

4.1 Ortogonalità tra spazi

Concludiamo questo capitolo con un risultato molto utile in svariati problemi. Stabilisce che, a partire dagli elementi di una base, si può costruire un'altra base fatta di vettori ortogonali fra loro (ed anche di lunghezza unitaria, se si vuole).

Proposizione 4.1.1 *Data una base* $\{\overrightarrow{b}_1,\dots,\overrightarrow{b}_n\}$ *di* $\mathbb{R}^n$*, è possibile, a partire da essa, costruire un'altra base, fatta di* vettori ortogonali.

Cenno all'idea della costruzione della nuova base, detto metodo di Graham-Schmidt.
La nuova base $\{\overrightarrow{w}_1,\dots,\overrightarrow{w}_n\}$ si costruisce così: si pone
$$\overrightarrow{w}_1=\overrightarrow{b}_1,$$
si pone
$$\overrightarrow{w}_2=\overrightarrow{b}_2-a\overrightarrow{b}_1,$$
con a scelto in modo che $\overrightarrow{w}_2$ sia ortogonale a $\overrightarrow{w}_1$; si pone
$$\overrightarrow{w}_3=\overrightarrow{b}_3-a_1\overrightarrow{b}_1-a_2\overrightarrow{b}_2,$$
con a_1,a_2 scelti in modo che $\overrightarrow{w}_3$ sia ortogonale a $\overrightarrow{w}_1,\overrightarrow{w}_2\dots$
Ognuno di questi problemi descritto dalle formule sopra ha tante incognite (i coefficienti a_i) quante equazioni (le condizioni di perpendicolarità). Non è strano dunque che esistano e siano uniche le soluzioni: facendo un po' di calcoli, si trova che
$$a=\frac{\overrightarrow{b}_2\cdot\overrightarrow{w}_1}{\overrightarrow{w}_1\cdot\overrightarrow{w}_1},\ a_1=\frac{\overrightarrow{b}_3\cdot\overrightarrow{w}_1}{\overrightarrow{w}_1\cdot\overrightarrow{w}_1},\ a_2=\frac{\overrightarrow{b}_3\cdot\overrightarrow{w}_2}{\overrightarrow{w}_2\cdot\overrightarrow{w}_2},\dots$$

Osservazione 4.1.1 Sia $U \subset \mathbb{R}^n$ un insieme qualunque. Allora l'insieme $V \subset \mathbb{R}^n$ dei vettori ortogonali a tutti i vettori di U tali che $V = \{\overrightarrow{v} \in \mathbb{R}^n : \overrightarrow{v} \cdot \overrightarrow{u} = 0 \quad \forall u \in U\}$ è un sottospazio vettoriale di $\mathbb{R}^n$.

Infatti $V \neq 0$ dato che V contiene $\overrightarrow{0}$. Siano $\overrightarrow{v_1}, \overrightarrow{v_2} \in V$ e sia $\overrightarrow{u} \in U$. Allora $\overrightarrow{v_1} \cdot \overrightarrow{u} = \overrightarrow{v_2} \cdot \overrightarrow{u} = \overrightarrow{0}$, e quindi $\overrightarrow{v_1} + \overrightarrow{v_2} \in V$. Siano $t \in \mathbb{R}$, $\overrightarrow{v} \in V$ e $\overrightarrow{u} \in U$. Allora $\overrightarrow{v} \cdot \overrightarrow{u} = 0$ quindi $(t\overrightarrow{u}) \cdot \overrightarrow{v} = 0$ che implica $t\overrightarrow{u} \in V$.
Useremo la notazione $U^\perp$ per indicare lo spazio dei vettori ortogonali ai vettori di U.

Estendendo quanto visto sopra, si può considerare un sottospazio U, e l'insieme V dei vettori ortogonali a tutti gli elementi di U. V è spazio vettoriale, ed è solitamente denotato $V = U^\perp$. Dualmente si ha che $U = V^\perp$ (osservare che $W^\perp$ è uno spazio vettoriale qualunque sia l'insieme W). Si dice allora che U, V sono *ortogonali tra loro.*

Vale il seguente risultato.

Proposizione 4.1.2 *Dato U sottospazio di $\mathbb{R}^n$, si ha che:* $\dim U + \dim U^\perp = n$.

Sia $B_1 = \{\overrightarrow{u}_1, \ldots, \overrightarrow{u}_k\}$ una base, fatta di vettori ortogonali, di U e $B_2 = \{\overrightarrow{v}_1, \ldots, \overrightarrow{v}_j\}$ una base, fatta di vettori ortogonali, di $U^\perp$. Supponiamo $k+j < n$; possiamo allora completare $B_1 \cup B_2$ per ottenere una base ortogonale di $\mathbb{R}^n$: $B = \{\overrightarrow{u}_1, \ldots, \overrightarrow{u}_k, \overrightarrow{v}_1, \ldots, \overrightarrow{v}_j, \overrightarrow{w}_1, \ldots, \overrightarrow{w}_i\}$. Consideriamo il vettore $\overrightarrow{w}_1$. Esso è ortogonale a tutti i vettori $\overrightarrow{u}_i$, quindi $\overrightarrow{w}_1 \in U^\perp$. D'altra parte, è anche ortogonale ai vettori $\overrightarrow{v}_i$, quindi $\overrightarrow{w}_1 \in U$. Dunque $\overrightarrow{w}_1 \in U^\perp \cap U$, e quindi $\overrightarrow{w}_1 = \overrightarrow{0}$. Ma questo è impossibile, perché è un elemento di una base. Questa contraddizione conclude la dimostrazione.

Ne segue, in particolare, che dato un sottospazio U di $\mathbb{R}^n$, si ha che $(U^\perp)^\perp = U$.

In generale, se U, V sono sottospazi, la loro somma $U + V = \{\overrightarrow{x} \in X : \exists \overrightarrow{u} \in U, \overrightarrow{v} \in V : \overrightarrow{x} = \overrightarrow{u} + \overrightarrow{v}\}$ è un sottospazio vettoriale. La sua dimensione è legata alla dimensione di U e di V nel modo che si vede nella prossima proposizione.

Proposizione 4.1.3 $\dim (U + V) = \dim U + \dim V - \dim (U \bigcap V)$.

Siano $\{\overrightarrow{w}_1, \ldots, \overrightarrow{w}_m\}$, $\{\overrightarrow{w}_1, \ldots, \overrightarrow{w}_m, \overrightarrow{u}_1, \ldots, \overrightarrow{u}_k\}$, $\{\overrightarrow{w}_1, \ldots, \overrightarrow{w}_m, \overrightarrow{v}_1, \ldots, \overrightarrow{v}_j\}$ basi di $U \cap V$, U e V, rispettivamente. Per provare quanto affermato, basterà verificare che

$$B = \{\overrightarrow{u}_1, \ldots, \overrightarrow{u}_k, \overrightarrow{v}_1, \ldots, \overrightarrow{v}_j, \overrightarrow{w}_1, \ldots, \overrightarrow{w}_m\}$$

è una base di $U + V$. Intanto è facile verificare che una combinazione lineare di elementi di B è un elemento di $U + V$. Poi è semplice anche verificare che gli elementi di B generano $U + V$. Vediamo allora che sono linearmente indipendenti. Supponiamo allora

$$c_1 \overrightarrow{u}_1 + \cdots + c_k \overrightarrow{u}_k + r_1 \overrightarrow{v}_1 + \cdots + r_j \overrightarrow{v}_j + s_1 \overrightarrow{w}_1 + \cdots + s_m \overrightarrow{w}_m = 0. \tag{4.1}$$

Si ha che

$$c_1 \overrightarrow{u}_1 + \cdots + c_k \overrightarrow{u}_k \in U \cap V.$$

Che stia in U è evidente, che stia in V dipende dal fatto che

$$c_1 \overrightarrow{u}_1 + \cdots + c_k \overrightarrow{u}_k = -r_1 \overrightarrow{v}_1 - \cdots - r_j \overrightarrow{v}_j - s_1 \overrightarrow{w}_1 - \cdots - s_m \overrightarrow{w}_m \in V.$$

Allora esistono dei coefficienti $\alpha_1, \dots \alpha_m$ tali che:

$$c_1 \overrightarrow{u}_1 + \cdots + c_k \overrightarrow{u}_k = \alpha_1 \overrightarrow{w}_1 + \cdots + \alpha_m \overrightarrow{w}_m.$$

Sostituendo in (4.1) si ottiene allora

$$(\alpha_1 + s_1)\overrightarrow{w}_1 + \cdots + (\alpha_m + c_m)\overrightarrow{w}_m + r_1 \overrightarrow{v}_1 + \cdots + r_j \overrightarrow{v}_j = 0,$$

che implica in particolare $r_1 = \cdots = r_j = 0$ (i vettori $\overrightarrow{w}_1, \dots, \overrightarrow{w}_m, \overrightarrow{v}_1, \dots \overrightarrow{v}_j$ sono linearmente indipendenti). In maniera del tutto analoga si mostra che $c_1 = \cdots = c_k = 0$. In conclusione, ritornando alla (4.1) si ha che

$$s_1 \overrightarrow{w}_1 + \cdots + s_m \overrightarrow{w}_m = 0,$$

che implica finalmente $s_1 = \cdots = s_m = 0$.

Nel caso in cui $U \bigcap V = \{0\}$, un elemento qualsiasi $\overrightarrow{x} \in U + V$ si può scrivere in modo *unico* $\overrightarrow{x} = \overrightarrow{u} + \overrightarrow{v}$, con $\overrightarrow{u} \in U$, $\overrightarrow{v} \in V$: si dice allora che la somma di U e V è *diretta* e si scrive $U \oplus V$.

Sia $\overrightarrow{x} = \overrightarrow{a} + \overrightarrow{b} = \overrightarrow{c} + \overrightarrow{d}$, con $\overrightarrow{a}, \overrightarrow{c} \in U$, $\overrightarrow{b}, \overrightarrow{d} \in V$. Allora $\overrightarrow{0} = (\overrightarrow{a} - \overrightarrow{c}) + (\overrightarrow{b} - \overrightarrow{d})$. Poiché si deve dimostrare che $\overrightarrow{a} = \overrightarrow{c}$, $\overrightarrow{b} = \overrightarrow{d}$, basta vedere che se $\overrightarrow{0} = \overrightarrow{u} + \overrightarrow{v}$, con $\overrightarrow{u} \in U$, $\overrightarrow{v} \in V$, allora $\overrightarrow{u} = \overrightarrow{v} = \overrightarrow{0}$. Ma chiaramente $\overrightarrow{v} = -\overrightarrow{u}$ e quindi $\overrightarrow{v} \in U$ perché U è spazio vettoriale. Poiché $U \bigcap V = \{0\}$, si ha allora $\overrightarrow{v} = 0$ e quindi anche $\overrightarrow{u} = 0$.

4.2 Cambiamento di base

Un vettore $\overrightarrow{x}$ di $\mathbb{R}^n$ si scrive come $\overrightarrow{x} = \sum_i x_i \overrightarrow{e}_i$, ove $x_i = \overrightarrow{x} \cdot \overrightarrow{e}_i$, ed i vettori $\overrightarrow{e}_i$ sono quelli della base canonica. Il fatto che gli $\overrightarrow{e}_i$ si chiamino vettori della base *canonica* indica chiaramente che hanno un ruolo privilegiato, che sta appunto nel fatto che la scrittura di un vettore $\overrightarrow{x}$ per mezzo della base canonica risulta particolarmente semplice. Tuttavia ha senso porsi il problema di come scrivere un generico vettore $\overrightarrow{x}$ in termini degli elementi di una base qualsiasi. Indichiamo allo stesso modo, con la lettera B, l'insieme costituito dagli elementi della base, e la matrice ottenuta accostando i vettori colonna della base. Scrivere allora $\overrightarrow{x}$ come combinazione lineare degli elementi della base B si dice, in forma abbreviata, scrivere $\overrightarrow{x}$ *in base* B. Vediamo allora come si ottiene il cambiamento di base.

Si ha

$$\overrightarrow{x} = \sum x'_k \overrightarrow{b}_k = B\overrightarrow{x}',$$

ove le colonne $\overrightarrow{b}_k$ sono gli elementi della base, e gli scalari x'_k rappresentano le incognite del problema, cioè i coefficienti di $\overrightarrow{x}$ in base B. Formalmente, la formula di cambiamento di base è la seguente:

$$\overrightarrow{x}' = B^{-1}\overrightarrow{x}.$$

Useremo la notazione $[\overrightarrow{x}]_B$, al posto di $\overrightarrow{x}'$, per mettere in evidenza la base B utilizzata.

Dunque per trovare $\overrightarrow{x}'$ bisogna invertire B.

Esempio 4.2.1 Dato il vettore $\overrightarrow{x} = \begin{pmatrix} 3 \\ 2 \\ 4 \end{pmatrix}$ scritto nella base canonica, scriviamo il vettore $\overrightarrow{x}'$ nella base B formata dai vettori $\overrightarrow{b}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, $\overrightarrow{b}_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$, $\overrightarrow{b}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$.
Dobbiamo risolvere il sistema nell'incognita $\overrightarrow{x}'$:
$\overrightarrow{x} = B\overrightarrow{x}' \Longrightarrow \begin{pmatrix} 1\,0\,0 \\ 1\,1\,0 \\ 1\,1\,1 \end{pmatrix} \begin{pmatrix} x_1' \\ x_2' \\ x_3' \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \\ 4 \end{pmatrix}$

$$\begin{cases} x_1' = 3 \\ x_1' + x_2' = 2 \\ x_1' + x_2' + x_3' = 4 \end{cases}$$

che ha come soluzione

$$\begin{cases} x_1' = 3 \\ x_2' = -1 \\ x_3' = 2 \end{cases} .$$

Il vettore $\overrightarrow{x}$ nella nuova base ha coordinate $(3, -1, 2)$.

Supponiamo ora che le colonne di B siano ortogonali fra loro, cioè che $\overrightarrow{b}_i \cdot \overrightarrow{b}_j = 0$ se $i \neq j$; si ha allora: $\overrightarrow{x} \cdot \overrightarrow{b}_j = (\sum x_k' \overrightarrow{b}_k) \cdot \overrightarrow{b}_j = x_j' \overrightarrow{b}_j \cdot \overrightarrow{b}_j$, e dunque:

$$x_j' = \frac{\overrightarrow{x} \cdot \overrightarrow{b}_j}{\overrightarrow{b}_j \cdot \overrightarrow{b}_j}.$$

In conclusione:

Teorema 4.2.1 *Data base B, fatta di vettori ortogonali, un vettore $\overrightarrow{x}$ in base B si scrive $\overrightarrow{x} = \sum x_k' \overrightarrow{b}_k$, con $x_k' = \frac{\overrightarrow{x} \cdot \overrightarrow{b}_k}{\overrightarrow{b}_k \cdot \overrightarrow{b}_k}$. Se poi i $\overrightarrow{b}_k$ sono versori, si ha la formula (semplice ed elegante) $x_k' = \overrightarrow{x} \cdot \overrightarrow{b}_k$.*

4.3 Esercizi

4.3.1 Quesiti a risposta chiusa

Rispondere vero o falso.

1. I vettori $\overrightarrow{v}_1 = (1, 2, 1)$, $\overrightarrow{v}_2 = (2, 2, 1)$,e $\overrightarrow{v}_3 = (1, 1, 2)$ formano una base di $\mathbb{R}^3$.

2. I polinomi $1 + x - x^2$, $x + x^3$, $1 + x^2 - 2x^3$ sono linearmente indipendenti.
3. L'insieme dei vettori $\overrightarrow{v}_1 = t_1 \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} + t_2 \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$ forma, al variare dei parametri $t_1, t_2 \in \mathbb{R}$ un sottospazio vettoriale di $\mathbb{R}^3$.
4. Siano $\overrightarrow{v}_1 = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$, $\overrightarrow{v}_2 = \begin{pmatrix} 2 \\ -1 \\ 2 \end{pmatrix}$, $\overrightarrow{v}_3 = \begin{pmatrix} 1 \\ 1 \\ a \end{pmatrix}$:
   a) per ogni valore reale di a i tre vettori dati sono linearmente indipendenti;
   b) $\overrightarrow{v}_3$ appartiene ad $L(\overrightarrow{v}_1, \overrightarrow{v}_2)$ (lo spazio generato da $\overrightarrow{v}_1$ e $\overrightarrow{v}_2$) se e solo se $a = 1$.
5. Dati i vettori $\overrightarrow{v}_1 = \begin{pmatrix} 1 \\ a \\ a \end{pmatrix}$, $\overrightarrow{v}_2 = \begin{pmatrix} a \\ 1 \\ a \end{pmatrix}$, $\overrightarrow{v}_3 = \begin{pmatrix} a \\ a \\ 1 \end{pmatrix}$:
   a) sono linearmente indipendenti per ogni valore di a;
   b) per $a = 1$ formano una base di $\mathbb{R}^3$.
6. Sia W il sottospazio di $\mathbb{R}^4$ generato dai vettori $\overrightarrow{u} = \begin{pmatrix} 1 & -2 & 5 & -3 \end{pmatrix}$, $\overrightarrow{v} = \begin{pmatrix} 2 & 3 & 1 & -4 \end{pmatrix}$, $\overrightarrow{w} = \begin{pmatrix} 3 & 8 & -3 & -5 \end{pmatrix}$:
   a) W ha dimensione 3;
   b) W è generato da $\overrightarrow{u}$ e $\overrightarrow{v}$;
7. se $\overrightarrow{u}, \overrightarrow{v}$ sono ortogonali e di lunghezza 1, allora $\{\overrightarrow{u}, \overrightarrow{v}, \overrightarrow{u} \wedge \overrightarrow{v}\}$ è una base ortonormale di $\mathbb{R}^3$.

4.3.2 Esercizi aperti

Esercizio 4.3.1 Stabilire se i seguenti sono sottospazi vettoriali di di $\mathbb{R}^3$:

1. $E = \left\{ \begin{pmatrix} x \\ y \\ 0 \end{pmatrix} \in \mathbb{R}^3 : x^2 + y^2 = 1 \right\}$;
2. $E = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 : |x| = |y| \right\}$.

Esercizio 4.3.2 Stabilire se i seguenti vettori sono linearmente indipendenti:

1. $\overrightarrow{v}_1 = \begin{pmatrix} 12 \\ 0 \\ 13 \\ 0 \end{pmatrix}$; $\overrightarrow{v}_2 = \begin{pmatrix} 4 \\ 0 \\ 4 \\ 1 \end{pmatrix}$; $\overrightarrow{v}_3 = \begin{pmatrix} 4 \\ 0 \\ -3 \\ -4 \end{pmatrix}$;
2. $v_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$; $\overrightarrow{v}_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$; $\overrightarrow{v}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$.

Esercizio 4.3.3 Quale tra i vettori $\overrightarrow{a} = \begin{pmatrix} 3 \\ 2 \\ 4 \end{pmatrix}$, $\overrightarrow{b} = \begin{pmatrix} 3 \\ 2 \\ 5 \end{pmatrix}$ appartiene allo spazio vettoriale generato da $\overrightarrow{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}$ e $\overrightarrow{v}_2 = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$?

Esercizio 4.3.4 Siano $V = L(\overrightarrow{v}_1, \overrightarrow{v}_2)$, $W = L(\overrightarrow{w}_1, \overrightarrow{w}_2)$ gli spazi vettoriali generati dai vettori $\overrightarrow{v}_1 = \begin{pmatrix} 1 \\ 0 \\ -1 \\ 2 \end{pmatrix}$ $\overrightarrow{v}_2 = \begin{pmatrix} 2 \\ 1 \\ 0 \\ -1 \end{pmatrix}$ $\overrightarrow{w}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$ $\overrightarrow{w}_2 = \begin{pmatrix} 3 \\ 2 \\ 1 \\ 0 \end{pmatrix}$. Trovare una base di V e di W.

Esercizio 4.3.5 Siano $\overrightarrow{u}$, $\overrightarrow{v}$ vettori linearmente indipendenti. Allora $\{\overrightarrow{u}, \overrightarrow{v}, \overrightarrow{u} \wedge \overrightarrow{v}\}$ formano una base di $\mathbb{R}^3$.

5

Sistemi Lineari

In questo capitolo ci occupiamo di estendere quanto visto precedentemente per i sistemi lineari che hanno tante equazioni quante incognite. Trattiamo infatti il caso in cui non necessariamente il numero delle equazioni sia uguale a quello delle incognite.

Supponiamo dunque di avere un sistema lineare di m equazioni ed n incognite. In generale, se $m = n$ ci si aspetta una ed una sola soluzione. Nel caso $m = 2$ ad esempio, dal punto di vista geometrico, risolvere il sistema significa cercare l'intersezione fra due rette, che di solito è un punto. Il che corrisponde a dire che il sistema ha appunto una ed una sola soluzione. Per $m = 3$, si tratta invece di determinare l'intersezione fra tre piani, che di solito fornisce un punto. Se invece il numero di equazioni è minore del numero delle incognite, di solito si hanno infinite soluzioni (per $m = 2$, $n = 3$ si ha intersezioni fra due piani, quindi una retta), anche se può succedere che non ce ne sia nessuna (piani paralleli). Se infine le equazioni sono più delle incognite, il rischio è quello di avere troppi vincoli, che quindi potrebbero non essere soddisfatti contemporaneamente (tre rette in un piano di solito non hanno punti in comune). Proprio gli esempi geometrici ricordati mostrano però che queste regole possono ammettere eccezioni: due rette potrebbero essere parallele, quindi il sistema non ha soluzione, o coincidenti, che significa infinite soluzioni. Esistono allora casi *degeneri*, di cui occorre tenere conto. Vediamo ora come il calcolo matriciale e l'idea di spazio vettoriale ci permettano di dare una risposta brillante allo studio della determinazione dell'esistenza o meno di soluzioni di un sistema, ed alla sua struttura.

Definizione 5.0.1 *Lo* spazio colonna *di una matrice A è lo spazio vettoriale generato dalle colonne di A, e si indica con $SC(A)$. Se $A^1, A^2, ...A^n$ sono le colonne di A, allora $SC(A) = L(A^1, A^2, ...A^n)$.*

Definizione 5.0.2 *Il* rango *di una matrice A (in simboli* $\mathrm{rk}\ A$*) è la dimensione dello spazio colonna di A:* $\mathrm{rk}\ A = \dim SC(A)$.

Osservazione 5.0.1 Il rango r di A rappresenta il numero massimo di colonne linearmente indipendenti che si possono estrarre da A. r colonne linearmente indipendenti formano quindi una base di $SC(A)$.

Definizione 5.0.3 *Lo* spazio riga *di una matrice A è lo spazio generato dalle righe di A. Se $A_1, A_2, ...A_m$ sono le righe di A, $SR(A) = L(A_1, A_2, ...A_m)$.*

Il prossimo teorema è molto importante.

Teorema 5.0.1 *Lo spazio riga e lo spazio colonna della matrice A hanno la stessa dimensione: cioè* $\text{rk } A = \dim SR(A)$.

Osservazione 5.0.2 dal teorema precedente seguono queste due semplici conseguenze:

1. Data una matrice A, $SC(A) = SR(A^t)$;
2. Si può calcolare il rango di una matrice usando le righe anziché le colonne. In particolare, il rango di A è anche il massimo numero di righe linearmente indipendenti e quindi, se A è una matrice $n \times m$, allora $\text{rk } A \leq \min\{m, n\}$.

Come si calcola il rango di una matrice? Una prima risposta è data dal seguente

Teorema 5.0.2 *Sia A una matrice quadrata $n \times n$. Allora* $\text{rk } A = n$ *se e solo se* $\det A \neq 0$.

Per dimostrare il teorema precedente, facciamo un'osservazione cruciale.

Se denotiamo con $A^1, A^2, ...A^n$ le colonne della matrice A e con $\overrightarrow{x} = \begin{pmatrix} x_1 \\ . \\ . \\ x_n \end{pmatrix}$ allora il sistema:

$$A\overrightarrow{x} = \overrightarrow{b} \tag{5.1}$$

si può scrivere

$$A\overrightarrow{x} = \sum_{k=1}^{n} x_k A^k. \tag{5.2}$$

D'altra parte, il sistema $A\overrightarrow{x} = \overrightarrow{b}$ ha soluzione unica, qualunque $\overrightarrow{b}$, se e solo se il determinante di A è non nullo. Equivalentemente, il sistema ha unica soluzione se e solo se esiste unico $(x_1, \ldots, x_n)$ tale che $\overrightarrow{b} = \sum_{k=1}^{n} x_k A^k$. Ma questo, visto che $\overrightarrow{b}$ è arbitrario, significa che SC(A)= $\mathbb{R}^n$, quindi che $\text{rk } A = n$.

Vediamo come procedere ora per determinare il rango di una matrice che non è quadrata, o che ha determinante nullo.

Definizione 5.0.4 *Data una matrice A, un* determinante minore *(o semplicemente* minore*) di ordine r è il determinante di una sottomatrice quadrata di A di ordine r.*

Teorema 5.0.3 *Sia A una matrice. Allora il rango di A è il massimo degli ordini dei minori non nulli di A.*

Osservazione 5.0.3 Valgono questi due fatti:

1. I minori di A di ordine maggiore del rango sono tutti uguali a zero, mentre esiste almeno un minore di ordine uguale al rango di A, diverso da zero;
2. Se B è una sottomatrice quadrata di A di ordine $r = \text{rk } A$ avente determinante non nullo, allora le righe di A che si sono scelte per formare B sono una base per $SR(A)$, mentre le colonne di B formano una base per $SC(A)$ (per dimostrare questo si utilizza il Teorema 5.0.1).

Esempio 5.0.1 Calcoliamo il rango della matrice: $\begin{pmatrix} 1 & -2 & -1 & 3 \\ 2 & -4 & -2 & 6 \\ 3 & -6 & -3 & 9 \end{pmatrix}$.

La matrice è 3×4 quindi il rango può essere al massimo 3. Consideriamo il determinante dei minori di ordine 3:

1. $\begin{pmatrix} 1 & -2 & -1 \\ 2 & -4 & -2 \\ 3 & -6 & -3 \end{pmatrix} = 0$ perché la seconda colonna è il doppio della terza;
2. $\begin{pmatrix} 1 & -2 & 3 \\ 2 & -4 & 6 \\ 3 & -6 & 9 \end{pmatrix} = 0$ perché la terza colonna è il triplo della prima;
3. $\begin{pmatrix} -2 & -1 & 3 \\ -4 & -2 & 6 \\ -6 & -3 & 9 \end{pmatrix} = 0$ perché la prima colonna è il doppio della seconda;
4. $\begin{pmatrix} 1 & -1 & 3 \\ 2 & -2 & 6 \\ 3 & -3 & 9 \end{pmatrix} = 0$ perché la terza colonna è il triplo della prima.

Tutti i minori di ordine 3 hanno determinante nullo, quindi consideriamo quelli di ordine 2. Si verifica che anche questi hanno tutti determinante nullo (in ognuno c'è sempre una colonna multipla di un'altra) e quindi il rango della matrice è 1. Si può dire la stessa cosa in modo un po' diverso. Chiamate $A^1, \ldots, A^4$ le colonne di A, si ha che $A^2 = -2A^1$, $A^3 = -A^1$, $A^4 = 3A^1$. Questo significa che A^2, A^3, A^4 sono proporzionali ad A^1, e quindi lo spazio generato dalle colonne di A è lo spazio generato, ad esempio da A^1, che è ovviamente uno dimensionale. Dunque il rango di A è uno.

Torniamo ora al problema iniziale. Consideriamo il sistema $A\overrightarrow{x} = \overrightarrow{b}$ e vediamo di stabilire se questo ammette soluzioni.

Definizione 5.0.5 *La matrice $[A|\overrightarrow{b}]$ (cioè la matrice A a cui aggiungiamo la colonna dei termini noti) si chiama la* matrice ampliata *del sistema $A\overrightarrow{x} = \overrightarrow{b}$.*

Ripetiamo quanto detto sopra, perché è la chiave per dare una condizione necessaria e sufficiente per la solvibilità di un sistema lineare.

Se $A^1, A^2, ...A^n$ sono le colonne della matrice A e se $\overrightarrow{x} = \begin{pmatrix} x_1 \\ \cdot \\ \cdot \\ x_n \end{pmatrix}$ allora il sistema:

$$A\overrightarrow{x} = \overrightarrow{b} \tag{5.3}$$

si può scrivere

$$A\overrightarrow{x} = \sum_{k=1}^{n} x_k A^k. \tag{5.4}$$

Quindi, dire che il sistema (5.1) ammette soluzione, è equivalente a dire che esistono $x_1, \ldots, x_n$ che verificano (5.2): ma questo significa esattamente che $\overrightarrow{b}$ *è combinazione lineare di* $A^1, A^2, ...A^n$.

Dunque vale il seguente teorema.

Teorema 5.0.4 *(di Rouché-Capelli) Sia* $A\overrightarrow{x} = \overrightarrow{b}$ *con* A *matrice* $m \times n$, $\overrightarrow{x} \in \mathbb{R}^n$, $\overrightarrow{b} \in \mathbb{R}^m, r = \text{rk } A, r' = \text{rk } ([A|\overrightarrow{b}])$.

1. *Se* $r' > r$, *il sistema è impossibile;*
2. *Se* $r' = r$ *il sistema ammette soluzioni.*

$r' > r$ significa che la colonna aggiunta a quelle della matrice A, e cioè $\overrightarrow{b}$, è *indipendente* dalle altre, perché avendola aggiunta si è aumentato il rango della matrice, quindi lo la dimensione dello spazio generato dalle colonne. Dunque il sistema è impossibile.

Osservazione 5.0.4 Supponiamo che il sistema sia risolubile. r, il rango della matrice A è (anche) la dimensione dello spazio generato dalle righe di A. Dunque r rappresenta *il numero delle equazioni del sistema effettivamente importanti per il sistema stesso*; le altre si possono buttare (nel sistema

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

un'equazione fra le prime due può esssere buttata via). Quali equazioni tenere? Basta scegliere quelle che intervengono in un minore di ordine r con determinante diverso da zero. Dunque, nel caso che il sistema abbia soluzioni, si può procedere così: supponiamo che il minore di ordine r con determinante diverso da zero sia quello di Nord-Ovest[1] (tutt'al più riordiniamo equazioni ed incognite perché questo succeda). Allora si buttano via le equazioni dalla $r + 1$ in poi, e si portano a secondo membro

[1] cioè $\begin{pmatrix} a_{11} & \ldots & a_{1r} \\ \ldots & \ldots & \ldots \\ a_{r1} & \ldots & a_{rr} \end{pmatrix}$

le parti delle equazioni relative alle incognite $x_{r+1}, \dots, x_n$, che ora vengono trattate come *parametri.* Siamo allora di fronte ad un sistema di r equazioni ed r incognite, con determinante della matrice dei coefficienti non nullo: la regola di Cramer ci permette di determinare l'unica soluzione. La quale dipende dai parametri $x_{r+1}, \dots, x_n$, che sono liberi. Ecco perché, tra l'altro, si dice che in questo caso ci sono ∞^{n-r} soluzioni.

Vediamo qualche esempio.

Esempio 5.0.2 Discutiamo i sistemi lineari:

1.
$$\begin{cases} x + y - z = 1 \\ 3x + 2y - 3x = \frac{7}{4} \\ -3x + y + 3z = 2 \end{cases} ,$$

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

Per il primo sistema di tre equazioni in tre incognite il determinante della matrice dei coefficienti $A = \begin{pmatrix} 1 & 1 & -1 \\ 3 & 2 & -3 \\ -3 & 1 & 3 \end{pmatrix}$ è nullo. Il minore (di ordine 2) $\begin{pmatrix} 1 & 1 \\ -3 & 1 \end{pmatrix}$ ha determinante diverso da zero, quindi il rango di A è 2. Il rango della matrice ampliata $C = \begin{pmatrix} 1 & 1 & -1 & 1 \\ 3 & 2 & -3 & \frac{7}{4} \\ -3 & 1 & 3 & 2 \end{pmatrix}$ è 2: tutti i minori di ordine 3 ottenuti utilizzando l'ultima colonna (quella dei termini noti) hanno determinante nullo. Per trovare le soluzioni, riscriviamo il sistema come:

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

(Abbiamo considerato come minore non nullo il minore $\begin{pmatrix} 1 & 1 \\ -3 & 1 \end{pmatrix}$, ottenuto considerando la prima e terza equazione e le variabili x, y. Per questo buttiamo via la seconda equazione e e per risolvere il sistema consideriamo la variabile z come parametro). A z fissato, il sistema ha un'unica soluzione, che si può trovare con Cramer, ad esempio: si ottiene allora $(z - \frac{1}{4}, \frac{5}{4})$. Il sistema allora ha le ∞^1 soluzioni: $(z - \frac{1}{4}, \frac{5}{4}, z)$.

Per risolvere il secondo sistema, osserviamo che, in questo caso, abbiamo più incognite che equazioni(due equazioni e tre incognite). La matrice A ha rango 2 perché il minore $\begin{pmatrix} 2 & 3 \\ 1 & 5 \end{pmatrix}$ ha determinante diverso da zero. Le due equazioni sono indipendenti (i due piani sono incidenti). Ovviamente il rango della matrice ampliata non può aumentare, perché si aggiunge una colonna. Dunque esistono soluzioni: $(7y - 2, y, 1 - 3y)$.

5.1 Algoritmo di Gauss

I teoremi di Cramer e di Rouché-Capelli sono estremamente utili per capire se e quando un sistema ha soluzioni, quale è la struttura dello spazio delle soluzioni, ma non sono efficaci dal punto di vista operativo, quando si tratta di calcolare effettivamente le soluzioni. In applicazioni industriali può succedere di dover calcolare le soluzioni di un sistema con anche un milione di incognite, e dunque un sistema efficace è molto più utile di uno meno efficace, in quanto i tempi d'uso macchina, con un enorme numero di equazioni e variabili, si dilatano moltissimo. In questo paragrafo illustreremo il metodo comunemente usato per risolvere i sistemi lineari, e che permette un notevole risparmio rispetto all'utilizzo della regola di Cramer. Per fare questo, cominciamo col vedere una definizione.

Definizione 5.1.1 *Due sistemi si dicono* equivalenti *se hanno le stesse soluzioni.*

Ad esempio sono sistemi equivalenti

$$\begin{cases} x_1 - 2x_2 + 3x_3 = 1 \\ x_1 - 4x_2 + x_3 = 0 \end{cases}$$

$$\begin{cases} 2x_1 - 6x_2 + 4x_3 = 1 \\ 2x_2 + 2x_3 = 1 \end{cases} .$$

Le matrici complete che rappresentano questi sistemi sono:

$$(A|b) = \begin{pmatrix} 1 & -2 & 3 & 1 \\ 1 & -4 & 2 & 0 \end{pmatrix}$$

e

$$(A'|b') = \begin{pmatrix} 2 & -6 & 4 & 1 \\ 0 & 2 & 1 & 1 \end{pmatrix} .$$

È possibile, partendo da un sistema lineare complesso, ottenerne con semplici passaggi uno ad esso equivalente ma più semplice da risolvere?
Sì , è possibile utilizzando tre operazioni "elementari", che ora descriviamo.

Prima operazione
Scambio di posto fra due equazioni del sistema.

Esempio 5.1.1 I sistemi:

$$\begin{cases} 2x_1 - x_2 + x_3 = 7 \\ -3x_1 - x_2 + 2x_3 = 3 \\ x_2 + x_3 = 0 \end{cases} \quad \text{e} \quad \begin{cases} -3x_1 - x_2 + 2x_3 = 3 \\ 2x_1 - x_2 + x_3 = 7 \\ x_2 + x_3 = 0 \end{cases}$$ sono (ovviamente) equivalenti.

Seconda operazione
Moltiplicazione di un'equazione per un coefficiente non nullo.

Esempio 5.1.2 I sistemi:

$$\begin{cases} 2x_1 - x_2 + x_3 = 7 \\ -3x_1 - x_2 + 2x_3 = 3 \\ x_2 + x_3 = 0 \end{cases} \quad \text{e} \quad \begin{cases} 2x_1 - x_2 + x_3 = 7 \\ -3x_1 - x_2 + 2x_3 = 3 \\ 3x_2 + 3x_3 = 0 \end{cases} \quad \text{sono equivalenti.}$$

Terza operazione
Somma di un'equazione con un multiplo di un'altra.

Esempio 5.1.3 I sistemi:

$$\begin{cases} 2x_1 - x_2 + x_3 = 7 \\ -3x_1 - x_2 + 2x_3 = 3 \\ x_2 + x_3 = 0 \end{cases} \quad \text{e} \quad \begin{cases} 2x_1 + 3x_2 + 5x_3 = 7 \\ -3x_1 - x_2 + 2x_3 = 3 \\ x_2 + x_3 = 0 \end{cases} \quad \text{sono equivalenti.}$$

(Il secondo è stato ottenuto dal primo sostituendo alla prima equazione la somma della stessa con il quadruplo della terza)

Quando il sistema di partenza è in forma matriciale, queste tre operazioni corrispondono a:

1. Scambiare le righe della matrice completa del sistema;
2. Moltiplicare una riga della matrice completa del sistema per un numero non nullo;
3. Sostituire ad una riga della matrice completa del sistema la somma della riga stessa con un multiplo di un'altra riga.

Osserviamo esplicitamente che queste operazioni *non* cambiano il rango della matrice, come deve evidentemente essere, dal momento che i sistemi sono equivalenti.

Le tre operazioni che abbiamo introdotto permettono di risolvere in modo piuttosto semplice i sistemi lineari attraverso un algoritmo (*Algoritmo di Gauss*) che trasforma le matrici in una forma particolare detta *a gradini.*

Definizione 5.1.2 *Sia A una matrice. Il* pivot i-esimo *P_i è il primo elemento non nullo della riga i-esima. Se una riga è fatta di tutti zeri, diciamo che la riga non ha pivot.*

Esempio 5.1.4 Nella matrice $A = \begin{pmatrix} 1 & 2 & 0 \\ 0 & -1 & 1 \\ 4 & 0 & 3 \\ 0 & 0 & 0 \end{pmatrix}$ $P_1 = 1$, $P_2 = -1$, $P_3 = 4$, P_4 non esiste.

Definizione 5.1.3 *Una matrice si dice* a gradini *se ogni pivot si trova in una delle colonne più a destra rispetto al pivot della riga precedente. Nel caso abbia una riga nulla, allora è a gradini se verifica la condizione precedente sulle righe non nulle e se la i-esima è la prima fatta di tutti zeri, allora le righe i sono fatte tutte di zeri.*

Esempio 5.1.5 Sono a gradini le matrici:

$$A = \begin{pmatrix} 1 & -1 & 3 \\ 0 & 2 & 1 \\ 0 & 0 & -1 \end{pmatrix} \quad B = \begin{pmatrix} 2 & 0 & 2 & 0 \\ 0 & 0 & 5 & 3 \\ 0 & 0 & 0 & 2 \end{pmatrix} \quad C = \begin{pmatrix} 1 & -1 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad D = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Esempio 5.1.6 Non sono a gradini le seguenti matrici:

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

Vediamo ora, con un esempio, l'algoritmo di riduzione di una matrice qualunque ad una matrice a gradini.

1. Si cerca una riga che abbia un elemento non nullo sulla prima colonna (se non la trovo passo alla seconda colonna); ad esempio supponiamo che la matrice sia la seguente:
$$A = \begin{pmatrix} 0 & 1 & -1 & 2 \\ 3 & -1 & -2 & 0 \\ 1 & 1 & 2 & 4 \end{pmatrix};$$
in tal caso la seconda riga va bene (come la terza, del resto);
2. Si porta la riga così determinata in prima posizione per avere il primo pivot in posizione a_{11}; nell'esempio
$$A = \begin{pmatrix} 3 & -1 & -2 & 0 \\ 0 & 1 & -1 & 2 \\ 1 & 1 & 2 & 4 \end{pmatrix}.$$
3. Ora si deve far comparire lo zero nei posti a_{21} e a_{31}. Si sostituisce la terza riga con la differenza fra la prima ed il triplo della terza:
$$A = \begin{pmatrix} 3 & -1 & -2 & 0 \\ 0 & 1 & -1 & 2 \\ 0 & -4 & -8 & -12 \end{pmatrix};$$
4. Si sostituisce ora la terza riga con la somma del quadruplo della seconda riga con la terza:
$$A = \begin{pmatrix} 3 & -1 & -2 & 0 \\ 0 & 1 & -1 & 2 \\ 0 & 0 & -12 & -4 \end{pmatrix}.$$

Vediamo un altro esempio: la matrice di partenza sia:

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

Vediamo la procedura, anzi una procedura, senza commenti.

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

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

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

In questo caso abbiamo sfruttato il fatto che la seconda riga aveva già zero ai posti a_{21} e a_{22}, agiamo quindi direttamente sulla terza per portare uno zero al posto a_{31} e poi facciamo uno scambio fra la seconda e la terza.

Torniamo ora al problema della soluzione di un sistema lineare, sfruttando il metodo precedente. Supponiamo dunque che la matrice in questione sia la matrice completa $(A|b)$ che rappresenta un sistema lineare $A\overrightarrow{x} = \overrightarrow{b}$. Trasformando $(A|b)$ in modo da ottenere una matrice a gradini, il nuovo sistema (equivalente a quello iniziale!) è facilmente risolubile: vediamo un esempio.

Esempio 5.1.7 Risolviamo il sistema:

$$\begin{cases} x_1 + x_2 + x_3 = 1 \\ 2x_1 + 3x_2 - 4x_3 = -3 \\ 3x_1 - 4x_2 + 5x_3 = -2 \end{cases}.$$

Scriviamo la matrice completa

$$(A|b) = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 2 & 3 & -4 & -3 \\ 3 & -4 & 5 & -2 \end{pmatrix}$$

e riduciamola a gradini.

1. Sottraiamo alla seconda riga il doppio della prima

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & -6 & -5 \\ 3 & -4 & 5 & -2 \end{pmatrix}$$

2. Facciamo la differenza tra la terza riga ed il triplo della prima

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & -6 & -5 \\ 0 & -7 & 2 & -5 \end{pmatrix}$$

3. Sommiamo la terza riga alla seconda moltiplicata per sette

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & -6 & -5 \\ 0 & 0 & -40 & -40 \end{pmatrix}.$$

Il sistema assume la forma

$$\begin{cases} x_1 + x_2 + x_3 = 1 \\ x_2 - 6x_3 = -5 \\ -40x_3 = -40 \end{cases}.$$

È semplice adesso ricavare x_3 dall'ultima equazione, x_2 dalla penultima ed x_1 dalla prima. La soluzione del sistema è

$$\begin{cases} x_1 = -1 \\ x_2 = 1 \\ x_3 = 1 \end{cases}.$$

Esempio 5.1.8 Risolviamo il sistema

$$\begin{cases} -x_1 + x_2 - x_3 = 2 \\ 3x_1 + 3x_2 + x_3 = 2 \\ x_1 + 5x_2 - x_3 = 6 \end{cases}$$

$$(A|b) = \begin{pmatrix} -1 & 1 & -1 & 2 \\ 3 & 3 & 1 & 2 \\ 1 & 5 & -1 & 6 \end{pmatrix}$$

1. Sommiamo alla seconda riga il triplo della prima

$$\begin{pmatrix} -1 & 1 & -1 & 2 \\ 0 & 6 & -2 & 8 \\ 1 & 5 & -1 & 6 \end{pmatrix}$$

2. Sommiamo la prima e la terza riga

$$\begin{pmatrix} -1 & 1 & -1 & 2 \\ 0 & 6 & -2 & 8 \\ 0 & 6 & -2 & 8 \end{pmatrix}$$

3. Facciamo la differenza tra la terza e la seconda riga
$$\begin{pmatrix} -1 & 1 & -1 & 2 \\ 0 & 6 & -2 & 8 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$
Notiamo che sono tutti nulli gli elementi dell'ultima riga della matrice: questo significa che un'equazione è "sparita". Rimangono allora solo due equazioni in tre incognite e siccome la sottomatrice di Nord-Ovest di ordine 2 ha determinante diverso da zero, questo significa che il sistema ha soluzioni.

Il sistema assume la forma
$$\begin{cases} -x_1 + x_2 - x_3 = 2 \\ 6x_2 - 2x_3 = 8 \end{cases}$$
che ha ∞^1 soluzioni date da:
$$\begin{cases} x_1 = \frac{-2-2x_3}{3} \\ x_2 = \frac{4+x_3}{3} \\ x_3 = x_3 \end{cases}$$

Esempio 5.1.9 Risolviamo il sistema
$$\begin{cases} -x_1 + 2x_2 - 3x_3 = 0 \\ 2x_1 - 4x_2 + 6x_3 = 5 \end{cases}$$
$$(A|b) = \begin{pmatrix} -1 & 2 & -3 & 0 \\ 2 & -4 & 6 & 5 \end{pmatrix}.$$

1. Sommiamo la seconda riga con il doppio della prima
$$\begin{pmatrix} -1 & 2 & -3 & 0 \\ 0 & 0 & 0 & 5 \end{pmatrix}.$$

Il sistema è impossibile perché, essendo il pivot dell'ultima riga in ultima colonna, la seconda equazione del sistema è impossibile:
$$\begin{cases} -x_1 + 2x_2 - 3x_3 = 0 \\ 5 = 0 \end{cases}$$

Vediamo ora un ultimo esempio, interessante per il fatto che mostra come risparmiare calcoli nel caso di dover risolvere due sistema lineari che hanno la stessa matrice dei coefficienti, ma termini noti differenti.

Consideriamo i seguenti sistemi:

$$\begin{cases} x+2y+z+u=1 \\ 2x+4y+u=0 \\ x+3z+u=1 \\ 2x+6y-2z+u=0 \end{cases},$$

$$\begin{cases} x+2y+z+u=1 \\ 2x+4y+u=2 \\ x+3z+u=0 \\ 2x+6y-2z+u=1 \end{cases}.$$

Scriviamo ora la seguente matrice, ottenuta accostando i due termini noti alla matrice dei coefficienti: $\begin{pmatrix} 1 & 2 & 1 & 1 & 1 & 1 \\ 2 & 4 & 0 & 1 & 0 & 2 \\ 1 & 0 & 3 & 1 & 1 & 0 \\ 2 & 6 & -2 & 1 & 0 & 1 \end{pmatrix}$. Cominciamo col mettere, al posto della seconda riga, la seconda meno due volte la prima. Si ottiene allora: $C = \begin{pmatrix} 1 & 2 & 1 & 1 & 1 & 1 \\ 0 & 0 & -2 & -1 & -2 & 0 \\ 1 & 0 & 3 & 1 & 1 & 0 \\ 2 & 6 & -2 & 1 & 0 & 1 \end{pmatrix}$.
Ora, al posto della terza, la terza meno la prima: $\begin{pmatrix} 1 & 2 & 1 & 1 & 1 & 1 \\ 0 & 0 & -2 & -1 & -2 & 0 \\ 0 & -2 & 2 & 0 & 0 & -1 \\ 2 & 6 & -2 & 1 & 0 & 1 \end{pmatrix}$. Scambiamo seconda e terza equazione: $\begin{pmatrix} 1 & 2 & 1 & 1 & 1 & 1 \\ 0 & -2 & 2 & 0 & 0 & -1 \\ 0 & 0 & -2 & -1 & -2 & 0 \\ 2 & 6 & -2 & 1 & 0 & 1 \end{pmatrix}$. Al posto della quarta, la quarta meno due volte la prima: $\begin{pmatrix} 1 & 2 & 1 & 1 & 1 & 1 \\ 0 & -2 & 2 & 0 & 0 & -1 \\ 0 & 0 & -2 & -1 & -2 & 0 \\ 0 & 2 & -4 & -1 & -2 & -1 \end{pmatrix}$. Al posto della quarta, la seconda più la quarta: $\begin{pmatrix} 1 & 2 & 1 & 1 & 1 & 1 \\ 0 & -2 & 2 & 0 & 0 & -1 \\ 0 & 0 & -2 & -1 & -2 & 0 \\ 0 & 0 & -2 & -1 & -2 & -2 \end{pmatrix}$.

Osserviamo la terza e la quarta riga. I primo quattro coefficienti sono uguali. Questo significa che, nel sistema, abbiamo due equazioni con identico primo membro. Anche il quinto coefficiente è uguale. Essendo essi i termini noti della terza e quarta equazione del primo sistema, questo ci dice che nel primo sistema si può buttare via un'equazione, ad esempio la quarta. Analogamente si ragiona sul secondo sistema, osservando che in questo caso però i termini noti della terza e quarta equazione sono diversi, e quindi il sistema è impossibile. Cerchiamo le soluzioni del primo, agendo sull'ultima matrice, cui cancelliamo l'ultima colonna (relativa al secondo sistema, quindi non interessante per noi). Riscriviamo il sistema:

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

Abbiamo un sistema di tre equazioni e quattro incognite, per cui una incognita può essere scelta in modo arbitrario:

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

Si è fatta un po' di fatica, ma adesso abbiamo un sistema facilissimo da risolvere, in quanto si ricava z dall'ultima equazione, ottenuto z si ricava y nella seconda, infine, avendo z ed y, si ricava x dalla prima.

Esercizio 5.1.1 Risolvere i sistemi lineari, provando a confrontare i metodi di Cramer Rouché-Capelli con quello di eliminazione:

1. $$\begin{cases} 2x_1 + x_2 - x_3 + 4x_4 = 5 \\ x_1 + x_2 + x_3 = 1 \\ x_1 + 2x_3 - x_4 = 1 \end{cases}$$

2. $$\begin{cases} 2x_1 + 3x_2 - x_3 + x_4 = 1 \\ 3x_1 - x_2 + x_3 - x_4 = 1 \\ x_1 - 4x_2 + 2x_3 - 2x_4 = 1 \end{cases}$$

3. $$\begin{cases} x_1 + x_2 - x_3 = 1 \\ x_1 - x_2 - x_3 = 1 \\ x_1 + x_2 + x_3 = 1 \end{cases}$$

4. $$\begin{cases} x_1 + x_2 - x_3 + x_4 = 1 \\ x_1 - x_2 + x_3 - x_4 = 1 \\ x_1 + x_2 + x_3 + x_4 = 1 \end{cases}$$

5.2 L'insieme delle soluzioni di un sistema

Gli esempi precedenti hanno messo in chiaro che quando un sistema lineare ammette più di una soluzione, allora queste sono infinite e dipendono da certi parametri liberi. Si può allora immaginare che, quando ci sono infinite soluzioni, l'insieme delle soluzioni abbia una struttura ben precisa. Intuizione ben fondata, come vediamo in questo paragrafo.

Definizione 5.2.1 *Si chiama* sistema omogeneo *un sistema della forma* $A\overrightarrow{x} = \overrightarrow{0}$*; si chiama* sistema completo, *il sistema* $A\overrightarrow{x} = \overrightarrow{b}$ *con* $\overrightarrow{b} \neq 0$.

Osservazione 5.2.1 Alcune importanti osservazioni:

1. Un sistema omogeneo ammette sempre almeno una soluzione: il vettore $\overrightarrow{0}$ lo risolve qualunque sia la matrice A;
2. Se $\overrightarrow{x}_1, \overrightarrow{x}_2$ sono soluzioni del sistema omogeneo $A\overrightarrow{x} = \overrightarrow{0}$, allora $A\overrightarrow{x_1} = \overrightarrow{0}$ e $A\overrightarrow{x_2} = \overrightarrow{0}$. Sommando membro a membro si ottiene $A(\overrightarrow{x_1} + \overrightarrow{x_2}) = \overrightarrow{0}$, e cioè anche $(\overrightarrow{x}_1 + \overrightarrow{x}_2)$ è soluzione del sistema omogeneo; lo stesso vale per $r\overrightarrow{x}_1$, qualunque $r \in \mathbb{R}$;
3. Se $\overrightarrow{x}_1, \overrightarrow{x}_2$ sono soluzioni del sistema lineare $A\overrightarrow{x} = \overrightarrow{b}$, allora $A\overrightarrow{x_1} = \overrightarrow{b}$, $A\overrightarrow{x_2} = \overrightarrow{b}$ e sottraendo membro a membro si ottiene $A(\overrightarrow{x_1} - \overrightarrow{x_2}) = \overrightarrow{0}$, e quindi $(\overrightarrow{x_1} - \overrightarrow{x_2})$ è soluzione del sistema omogeneo $A\overrightarrow{x} = \overrightarrow{0}$;
4. Se $\overrightarrow{z}$ risolve il sistema omogeneo, e $\overrightarrow{x_0}$ quello completo, allora $\overrightarrow{x_0} + \overrightarrow{z}$ è ancora soluzione del sistema completo.

Usiamo la notazione Sol $(A, \overrightarrow{b})$ per indicare l'insieme delle soluzioni del sistema lineare $A\overrightarrow{x} = \overrightarrow{b}$. Abbiamo allora:

Teorema 5.2.1 Sol $(A, \overrightarrow{0})$ *è uno spazio vettoriale.*

Sol $(A, \overrightarrow{0})$ è un sottoinsieme non vuoto (punto 1 osservazione precedente) di $\mathbb{R}^n$, chiuso rispetto alle operazioni di somma e di moltiplicazione per uno scalare, come visto nel punto 2. dell'osservazione di sopra. Dunque è un sottospazio di $\mathbb{R}^n$ e quindi uno spazio vettoriale.

Teorema 5.2.2 Sol $(A, \overrightarrow{b})$ *è il traslato di uno spazio vettoriale:* Sol $(A, \overrightarrow{b}) = \overrightarrow{\overline{x}} +$ Sol $(A, \overrightarrow{0})$*, dove* $\overrightarrow{\overline{x}}$ *indica una qualsiasi soluzione del sistema completo.*

Discende immediatamente dal teorema precedente e dai punti 3. e 4. dell'osservazione di sopra.

Esempio 5.2.1 Consideriamo il sistema, già visto nel paragrafo precedente:

$$\begin{cases} x_1 + x_2 - x_3 = 1 \\ 3x_1 + 2x_2 - 3x_3 = \frac{7}{4} \\ -3x_1 + x_2 + 3x_3 = 2 \end{cases},$$

che ha come soluzioni $(z - \frac{1}{4}, \frac{5}{4}, z)$. Esse possono essere riscritte in forma vettoriale come: $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} (t - \frac{1}{4}) \\ \frac{5}{4} \\ t \end{pmatrix} = \begin{pmatrix} t \\ 0 \\ t \end{pmatrix} + \begin{pmatrix} -\frac{1}{4} \\ \frac{5}{4} \\ 0 \end{pmatrix} = t \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + \begin{pmatrix} -\frac{1}{4} \\ \frac{5}{4} \\ 0 \end{pmatrix}$.

Le soluzioni del sistema omogeneo associato sono $\begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} = t \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$. La dimensione di Sol $(A, \overrightarrow{0})$ è 1 ed una sua base è data dal vettore $\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$. Una soluzione particolare del sistema non omogeneo è data dal vettore: $\begin{pmatrix} -\frac{1}{4} \\ \frac{5}{4} \\ 0 \end{pmatrix}$.

Concludiamo il capitolo con la seguente osservazione. Supponiamo di avere a che fare con un sistema lineare in cui $\overrightarrow{u_1}^t, \overrightarrow{u_2}^t, \ldots \overrightarrow{u_m}^t$ sono le righe di una matrice A $(m \times n)$. Allora $A\overrightarrow{x}$ (ove $\overrightarrow{x} \in \mathbb{R}^n$) è il vettore di componenti $\overrightarrow{u_i}^t \overrightarrow{x_i} \in \mathbb{R}^m$.

Dunque, se consideriamo il sistema omogeneo:

$$A\overrightarrow{x} = \overrightarrow{0},$$

le sue soluzioni sono esattamente quei vettori che sono ortogonali alle righe di A, e quindi anche allo spazio generato dalle righe stesse.

Ne segue che Sol $(A, \overrightarrow{0}) = U^\perp$, dove U indica lo spazio generato dalle righe di A. Osserviamo ancora che, se il rango della matrice è r, allora Sol $(A, \overrightarrow{0})$ ha dimensione $n - r$ (n incognite, r equazioni effettive.)

Infine, se il sistema è completo, consideriamo $Y = A\mathbb{R}^n$, l'insieme di tutti i vettori che si ottengono moltiplicando righe per colonne A per un generico vettore di $\mathbb{R}^n$: se $\overrightarrow{b} \in A\mathbb{R}^n$ allora il sistema è risolubile ed esiste unico vettore $\overrightarrow{x_0} \in U$ tale che $A\overrightarrow{x_0} = \overrightarrow{b}$. Tutte le altre soluzioni si trovano sommando a $\overrightarrow{x_0}$ i vettori di Sol $(A, \overrightarrow{0})$.

Vediamo quanto detto con un esempio semplice. Sia A il vettore $(-1, 1)$ (una matrice particolarmente semplice!) Si ha allora che $U = \{(x, -x) : x \in \mathbb{R}\}$ (lo spazio generato dalle righe di A). D'altra parte Sol $(A, 0)$, cioè l'insieme delle soluzioni dell'equazione $-x + y = 0$, è dato da Sol $(A, 0) = \{(x, x) : x \in \mathbb{R}\} = U^\perp$. L'equazione $A\overrightarrow{x} = b$ ha come insieme di soluzione $\{(b, -b) + (x, x) : x \in \mathbb{R}\}$. Come si vede, ogni soluzione di $A\overrightarrow{x} = b$ scrive come somma di un ben preciso vettore di U $((b, -b))$ e di un qualunque vettore di $U^\perp$.

Consideriamo ora il caso un po' meno semplice di un sistema di tre equazioni in tre incognite con matrice A dei coefficienti di rango $r = 2$. La matrice A rappresenta una trasformazione da $\mathbb{R}^3 \longrightarrow \mathbb{R}^3$. Nello spazio di partenza, dove stanno i vettori $\overrightarrow{x}$, è possibile individuare il sottospazio U generato dalle righe di A, che ha dimensione

2 (2 è il rango della matrice) e che geometricamente è un piano per l'origine, ed il suo sottospazio ortogonale, che è di dimensione 1 e quindi rappresenta una retta. Nello spazio di arrivo $\mathbb{R}^3$, possiamo individuare l'immagine della trasformazione. Questo è un sottospazio di dimensione 2, cioè un piano (in quanto il nucleo della trasformazione ha dimensione 1, è la retta perpendicolare a U). Considerando ora il sistema $A\overrightarrow{x} = \overrightarrow{b}$, si presentano due casi. Se $\overrightarrow{b}$ non appartiene al sottospazio di dimensione 2 che è l'immagine della trasformazione, allora il sistema non ha soluzione, se ci appartiene, ogni soluzione del sistema non omogeneo si scrive allora come somma di un ben preciso vettore che giace nel piano individuato da U, e di un generico vettore della retta ad esso perpendicolare.

Esercizio 5.2.1 Data una matrice A $m \times n$, dimostrare che $Y = A\mathbb{R}^n$ è un sottospazio di $\mathbb{R}^m$.

5.3 Esercizi

5.3.1 Quesiti a risposta chiusa

Rispondere vero o falso.

1. Il rango della matrice $\begin{pmatrix} 1 & 1 & 0 & 1 \\ 2 & 1 & 1 & -1 \\ 1 & 0 & 1 & -2 \end{pmatrix}$ è 2;
2. Si consideri il sistema $\begin{cases} x + y + w = 1 \\ 2x + y + z - w = 1 \\ x + z - 2w = 0 \end{cases}$
   a) il sistema non ha soluzione;
   b) il sistema ha un'unica soluzione;
   c) l'insieme delle soluzioni del sistema è un sottospazio di $\mathbb{R}^4$.
3. Il sistema $\begin{cases} kx + y + (k-1)z = 1 \\ x + ky = k + 1 \\ (k+1)x + (k+1)y + (k-1)z = 2 + k \end{cases}$:
   a) ammette soluzione per ogni valore di k;
   b) ammette una soluzione solo se $k = 0$;
   c) esiste un valore reale di k per cui il sistema ammette un'unica soluzione.
4. Si consideri il sistema lineare omogeneo $\begin{cases} x + y - z + w = 0 \\ x - y + z + w = 0 \\ 2x + 2w = 0 \\ 2y - 2z = 0 \end{cases}$:

a) $\begin{cases} x = s + 2t \\ y = s + r \\ z = s + r \\ w = -s - 2t \end{cases}$ è soluzione del sistema per ogni valore dei parametri r, s, t;

b) il sistema ammette ∞^3 soluzioni.

5. Dato il sistema $\begin{cases} ax + y - az = a \\ x + y + z = 1 \\ x + ay - z = 1 \end{cases}$:

a) è lineare;
b) per $a \neq 0$ ammette una ed una sola soluzione;
c) esistono dei valori di a per cui il sistema impossibile;
d) per $a = 0$ è un sistema omogeneo.

5.3.2 Esercizi aperti

Esercizio 5.3.1 Determinare il rango della matrice: $A = \begin{pmatrix} 1 & 2 & 3 \\ 5 & 3 & 0 \\ 7 & 7 & 6 \end{pmatrix}$.

Esercizio 5.3.2 Determinare, per ogni valore del parametro, il rango delle matrici: $A = \begin{pmatrix} t & 0 & 0 \\ t & t-1 & 0 \\ t & 2t-2 & 2t-2 \end{pmatrix}$, $B = \begin{pmatrix} 0 & a & 0 \\ a(a-1) & a & 0 \\ 0 & a & a(a-1) \end{pmatrix}$.

Esercizio 5.3.3 Discutere e se possibile risolvere i seguenti sistemi lineari. Nel caso il sistema ammetta soluzioni, scrivere le soluzioni nella forma Sol $(A, \overrightarrow{b}) = \overrightarrow{x} +$ Sol $(A, \overrightarrow{0})$ dove $\overrightarrow{x}$ indica una qualsiasi soluzione del sistema completo.

1. $\begin{cases} x + ay = a \\ ax + y = a \end{cases}$.

2. $\begin{cases} ax + (a+1)y + 3z = a \\ 2x + 4ay + (a+5)z = 2 \\ 2x + ay + 5z = 1 \end{cases}$.

3. $\begin{cases} x + y - z = 1 \\ 3x + 2y + az = \frac{7}{4} \\ ax + y + 3z = 2 \end{cases}$.

6

Trasformazioni lineari

Abbiamo visto che uno spazio vettoriale è un insieme su cui sono definite due operazioni: quella di somma, operazione fra due elementi dello spazio, e quella di moltiplicazione di un elemento dello spazio per uno scalare, di solito un numero reale ma eventualmente anche un numero complesso. Quando si passa a studiare le funzioni che agiscono fra spazi vettoriali, diventano allora particolarmente importanti quelle operazioni che tengono conto delle operazioni definite sopra. Ma che cosa significa "tenere conto di"? La prossima definizione rende preciso questo concetto.

Definizione 6.0.1 *Una funzione $f : U \to V$, dove U, V sono spazi vettoriali, si dice* trasformazione lineare *se soddisfa le due condizioni seguenti:*

1. $f(x + y) = f(x) + f(y)$, *per ogni* $x, y \in U$*;*
2. $f(tx) = tf(x)$ *per ogni* $t \in \mathbb{R}$, *per ogni* $x \in U$[1]*.*

Quel che dice la definizione è chiaro: una trasformazione lineare *rispetta* la struttura vettoriale degli spazi fra cui agisce, nel senso che trasforma le somme nello spazio di partenza U in somme nello spazio di arrivo V.

Nonostante le funzioni siano in genere indicate con lettere minuscole $(f, g \dots)$ e che le trasformazioni lineari siano funzioni, d'ora in poi usiamo lettere maiuscole per indicare le trasformazioni lineari, proprio per indicare che stiamo usando funzioni particolari.

Esempio 6.0.1 Dato U, spazio vettoriale, la trasformazione $I : U \to U$ tale che $I(x) = x$ per ogni $x \in U$ è evidentemente una trasformazione lineare. Le funzioni $L_a : \mathbb{R} \to \mathbb{R}$ definite come $L_a(x) = ax$ sono lineari. $f(x) = \sin x$ non è una trasformazione lineare. Se P è lo spazio dei polinomi, la trasformazione $D : P \to P$ tale che $D(p) =$

[1] In questo capitolo, per alleggerire le notazioni, non mettiamo la solita freccia sulle lettere per denotare i vettori. Dove si parla di vettori è chiaro dal contesto. Eccezioni: lo 0 di uno spazio vettoriale è ancora indicato con $\overrightarrow{0}$ per non confonderlo con il numero reale 0, le colonne di una matrice lo stesso, per mettere in evidenza che sono vettori colonna

p', che associa al polinomio P la sua derivata (che è ancora un polinomio) è una trasformazione lineare.

Sia $T : V \to W$ una trasformazione lineare di uno spazio vettoriale V in uno spazio vettoriale W.

Definizione 6.0.2 *Si chiama* immagine *di T l'insieme* Im $(T) = \{w \in W : \exists v \in V : T(v) = w\}$.

Definizione 6.0.3 *Si chiama* nucleo *di T l'insieme* Ker $T = \{v \in V : T(v) = \overrightarrow{0}\}$, *cioè l'insieme dei vettori v che hanno come immagine nella trasformazione T il vettore nullo.*

Teorema 6.0.1 Im *(T) è un sottospazio vettoriale di W.*

Per dimostrare che Im (T) è un sottospazio di W basta verificare che:

1. se $T(x) \in W$, $T(y) \in W$ allora $T(x) + T(y) \in W$: infatti, per linearità, $T(x) + T(y) = T(x + y)$, allora $T(x) + T(y) \in W$ come immagine di $x + y$;
2. se $T(x) \in W$ allora $cT(x) = T(cx) \in W$. Osserviamo che, ponendo $c = 0$ si ha che $T(\overrightarrow{0}) = \overrightarrow{0}$. Dunque in particolare le trasformazioni lineari mandano sempre lo zero di V nello zero di W.

Teorema 6.0.2 *Il nucleo di T è un sottospazio di T.*

Se $x, y \in$ Ker (T), allora $\overrightarrow{0} + \overrightarrow{0} = T(x) + T(y) = T(x + y)$, e $\overrightarrow{0} = cT(x) = T(cx)$. Questo implica che $x + y \in$ Ker (T), $cx \in$ Ker (T).

La dimensione di Ker (T) si dice *nullità* di T. La dimensione di Im (T) si dice *rango* di T.

Il seguente importante teorema, ci dice che le dimensioni di questi due spazi vettoriali sono legate a quella di V.

Teorema 6.0.3 Teorema della "nullità più rango"
La somma della nullità più il rango di una trasformazione lineare è uguale alla dimensione del dominio:

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

Supponiamo che $k = \dim \operatorname{Ker}(T)$ e $n = \dim V$. Consideriamo una base $v_1, v_2, \ldots v_k$ di Ker (T) e completiamola ad una base $v_1, v_2, \ldots v_k, v_{k+1}, \ldots v_n$. Vogliamo vedere che $T(v_{k+1}), \ldots T(v_n)$ sono una base per Im (T). Da ciò ne segue che $\dim \operatorname{Im}(T) = n - k$ cioè $\dim \operatorname{Im}(T) + k = n$ ovvero $\dim \operatorname{Ker}(T) + \dim \operatorname{Im}(T) = \dim V$. Per vedere che $\{T(v_{k+1}), \ldots T(v_n)\}$ costituisce una base per Im (T), proviamo dapprima che $T(v_{k+1}), \ldots T(v_n)$ sono linearmente indipendenti. Supponiamo allora

$$c_1 T(v_{k+1}) + \ldots c_n T(v_{k+n}) = 0,$$

e verifichiamo che $c_i = 0$ per ogni i. Per linearità si ha:

$$T(c_1 v_{k+1} + \dots c_n v_{k+n}) = 0,$$

e questo implica (perché?)

$$c_1 v_{k+1} + \dots c_n v_{k+n} = 0.$$

Essendo i vettori v_{k+i}, $i = 1, \dots, n$ linearmente indipendenti, questo implica $c_i = 0$ per ogni i. Ora dobbiamo vedere che $\{T(v_{k+1}), \dots T(v_n)\}$ genera l'immagine di T. Sia $w \in \mathrm{Im}\,(T)$. Allora esiste $v \in V$ tale che $T(v) = w$. Allora esistono c_i tali che $v = \sum_{i=1}^{k+n} c_i v_i$, e quindi

$$T(v) = \sum_{i=1}^{k+n} c_i T(v_i) = \sum_{i=k+1}^{k+n} c_i T(v_i).$$

Esempio 6.0.2 1. Sia $T : \mathbb{R}^2 \to \mathbb{R}^2$ la trasformazione $T(x, y) = (x - y, x + y)$. $T(x, y)$ è una trasformazione lineare, infatti:
   a) $T(x, y) + T(z, t) = (x - y, x + y) + (z - t, z + t) = (x + z - (y + t), x + z + (y + t)) = T(x + z, y + t)$;
   b) $cT(x, y) = (cx - cy, cx + cy) = T(c(x, y))$.

   Calcoliamo adesso la dimensione di Ker (T) e Im (T). Gli elementi del nucleo sono i vettori (x, y) che hanno come immagine nella trasformazione T, il vettore nullo: $T(x, y) = (x - y, x + y) = (0, 0)$ che implica $\begin{cases} x - y = 0 \\ x + y = 0 \end{cases}$, soddisfatta solo dal vettore $(0, 0)$. Quindi $\dim \mathrm{Ker}\,(T) = 0$. Se applichiamo la formula delle dimensioni, dato che $\dim V = 2$, ricaviamo $\dim \mathrm{Im}\,(T) = 2$ (verificare direttamente che l'immagine di T è $\mathbb{R}^2$);

2. Sia $T : \mathbb{R}^3 \to \mathbb{R}^2$ la trasformazione $T(x, y, z) = (x, 2y + z)$. Verifichiamo che si tratta di una trasformazione lineare:
   a) $T(x, y, z) + T(x_1, y_1, z_1) = (x, 2y + z) + (x_1, 2y_1 + z_1) = (x + x_1, 2(y + y_1) + z + z_1) = T(x + x_1, y + y_1, z + z_1)$;
   b) $cT(x, y, z) = c(x, y + z) = (cx, cy + cz) = T(cx, cy, cz)$.

   Calcoliamo la dimensione del nucleo: $T(x, y, z,) = (0, 0) \Rightarrow (x, 2y + z) = (0, 0)$ cioè $x = 0,\ y = -\frac{1}{2}z$ quindi i vettori del nucleo sono della forma $(x, y, z) = t(0, 1, -2)$ e dunque la dimensione è 1. La dimensione dell'immagine è $3 - 1 = 2$ (verificare direttamente che l'immagine di T è $\mathbb{R}^2$).

Di quante informazioni abbiamo bisogno per identificare una trasformazione lineare? Vale la seguente proposizione, che è fondamentale.

Proposizione 6.0.1 *Siano $T, S : U \to V$ due trasformazioni lineari, sia B una base di U. Se*

$$T(b) = S(b) \qquad \forall b \in B,$$

allora $T(x) = S(x)$ per ogni $x \in X$, cioè $T = S$.

Sia $x \in U$. Allora, se $B = \{b_1, \dots, b_n\}$, esistono $c_1, \dots, c_n$ tali che $x = \sum_{i=1}^n c_i b_i$. si ha allora:

$$T(x) = T(\sum_{i=1}^n c_i b_i) = \sum_{i=1}^n c_i T(b_i) = \sum_{i=1}^n c_i S(b_i) = S(\sum_{i=1}^n c_i b_i) = S(x).$$

In altre parole, una trasformazione lineare è nota qualora si sappia quanto vale sugli elementi di una base. In maniera equivalente, data una base, si può definire una funzione in modo *arbitrario* sugli elementi della base; esiste poi un'unica trasformazione lineare che assuma quei valori dati sugli elementi della base.

La proposizione precedente ci dà informazioni preziose sulla struttura dell'insieme delle trasformazioni lineari da uno spazio vettoriale in un altro. Ad esempio, come sono fatte le trasformazioni lineari da $\mathbb{R}$ in sé? Poiché una base di $\mathbb{R}$ contiene un solo elemento (ad esempio 1), esse sono identificate da un solo numero reale. L'insieme delle trasformazioni lineari da $\mathbb{R}$ in sé si può allora identificare con $\mathbb{R}$ stesso. Infatti il grafico di una tale trasformazione lineare non è altro che una retta passante per l'origine, che può essere identificata tramite il suo coefficiente angolare (cioè, guarda caso, proprio il valore che assume nel punto 1). Analogamente, l'insieme delle trasformazioni lineari da $\mathbb{R}^n$ in $\mathbb{R}$ si può allora identificare con $\mathbb{R}^n$ stesso. Che cosa significa che il vettore $(x_1, \dots, x_n)$ identifica una trasformazione T di questo tipo? Significa che, avendo ad esempio fissato la base canonica $\{e_1, \dots, e_n\}$ di $\mathbb{R}^n$, allora $T(e_i) = x_i$. Appare allora chiaro come le strutture di spazio vettoriale e di applicazione lineare siano così importanti. Per conoscere una funzione reale di variabile reale, ad esempio $\sin x$, dobbiamo sapere quanto vale in almeno un intervallo, per conoscere una funzione lineare da $\mathbb{R}^n$ in $\mathbb{R}$ basta conoscerla in n punti (opportuni).

Una matrice A $m \times n$ identifica in maniera naturale una trasformazione lineare $T : \mathbb{R}^n \to \mathbb{R}^m$: basta porre $T(x) = A\overrightarrow{x}$.

Scriviamo la formula delle dimensioni in questo caso.

Supponiamo di avere una trasformazione lineare $T_A : \mathbb{R}^n \to \mathbb{R}^m$, Ker (T_A) è l'insieme degli $x \in \mathbb{R}^n$ tali che $T_A(x) = \overrightarrow{0}$ cioè degli $\overrightarrow{x}$ tali che $A\overrightarrow{x} = \overrightarrow{0}$. Im (T_A) è l'insieme degli $y \in \mathbb{R}^m$ tali che esiste x tale che $A\overrightarrow{x} = \overrightarrow{y}$. Ma questo è l'insieme delle combinazioni lineari delle colonne di A e dunque, per definizione, l'insieme è lo spazio colonna di A. Quindi Im $(T_A) = SC(A)$ e la formula delle dimensioni afferma che $\dim \mathrm{Ker}\ (T_A) + \dim \mathrm{Im}\ (T_A) = \dim R^n = n$.

Supponiamo ora di avere una trasformazione lineare $T_A : \mathbb{R}^n \to \mathbb{R}^n$, associata alla matrice A, come spiegato sopra. Supponiamo anche che si decida di cambiare base, passando dalla base canonica ad un'altra base B. Come si è già visto, il generico vettore $\overrightarrow{x}$ si trasforma, nella nuova base, in $[\overrightarrow{x}]_B = B^{-1}\overrightarrow{x}$. Analogamente, il trasformato $\overrightarrow{y} = A\overrightarrow{x}$ si trasforma in $[\overrightarrow{y}]_B = B^{-1}\overrightarrow{y}$. Quindi,

$$[\overrightarrow{y}]_B = B^{-1}\overrightarrow{y} = B^{-1}A\overrightarrow{x} = B^{-1}AB[\overrightarrow{x}]_B.$$

In altre parole, la trasformazione di partenza L, descritta dalla matrice A (in base canonica), può essere ugualmente descritta dalla matrice $B^{-1}AB$ (in base B). Ne possiamo concludere che:

Data una matrice A ed una matrice non singolare B, le matrici A e $B^{-1}AB$ descrivono la stessa trasformazione lineare.

Perché fare questa fatica? Semplicemente perché la forma della matrice permette di descrivere in modo più o meno semplice la trasformazione lineare. Supponiamo per esempio di poter descrivere una trasformazione lineare L per mezzo di una matrice D *diagonale*, con elementi d_i sulla diagonale principale. Allora la trasformazione $y = Lx$ si scrive come:

$$\begin{cases} y_1 = d_1 x_1, \\ y_2 = d_2 x_2 \\ \dots \\ y_n = d_n x_n. \end{cases}$$

Dovendo effettuare dei calcoli, appare evidente il vantaggio di una tale rappresentazione: per ogni i la variabile y_i dipende solo dalla variabile x_i: si dice che abbiamo separato le variabili.

Dalle considerazioni precedenti, appare chiaro che le prossime due definizioni possono essere utili.

Definizione 6.0.4 *Date due matrici A e B (quadrate e della stessa dimensione) si dice che sono* simili *se esiste una matrice (non singolare) S tale che $A = S^{-1}BS$.*

Definizione 6.0.5 *Una matrice A si dice* diagonalizzabile *se è simile ad una matrice diagonale.*

La relazione di similitudine è una *relazione di equivalenza* (che si indica col simbolo $\simeq$): questo significa che è una relazione che gode delle proprietà seguenti:

- *riflessiva*: $A \simeq A$;
- *simmetrica*: $A \simeq B$ implica $B \simeq A$;
- *transitiva*: $A \simeq B$ e $B \simeq C$ implica $A \simeq C$.

Provate a fare le verifiche per esercizio.

Ricordiamo ancora una volta che la proprietà di similitudine fra matrici si interpreta, in termini delle trasformazioni lineari associate, dicendo che esse individuano la stessa trasformazione.

Ci interessiamo ora al problema della diagonalizzazione di una matrice. Premettiamo che non tutte la matrici sono diagonalizzabili anche se, come vedremo, il sottoinsieme, particolarmente importante, delle matrici simmetriche, lo è. Ma ora vediamo che relazioni devono esistere fra una matrice diagonalizzabile e la sua diagonalizzata. D'ora in poi usiamo i simboli Λ per indicare una matrice diagonale, U per indicare una matrice non singolare che permette di passare da A a Λ[2].

[2] U viene a volte chiamata *matrice modale* di A

La relazione fondamentale

$$\Lambda = U^{-1}AU$$

si scrive anche, ovviamente,

$$AU = U\Lambda.$$

Ora, se indichiamo con $\overrightarrow{u_1}, \dots, \overrightarrow{u_n}$ le colonne di U, allora AU è la matrice che ha per colonne

$$A\overrightarrow{u_1}, \dots, A\overrightarrow{u_n}.$$

La matrice $U\Lambda$ invece ha come colonne:

$$\lambda_1\overrightarrow{u_1}, \dots, \lambda_n\overrightarrow{u_n}.$$

Uguagliando le due espressioni si ottiene allora che $AU = U\Lambda$ si esprime mediante le uguaglianze fra vettori:

$$A\overrightarrow{u_1} = \lambda_1\overrightarrow{u_1}, \dots, A\overrightarrow{u_n} = \lambda_n\overrightarrow{u_n}.$$

$\lambda_1, \dots, \lambda_n$ sono detti gli *autovalori* della matrice A, $\overrightarrow{u_1}, \dots, \overrightarrow{u_n}$ sono autovettori associati, i quali sono certamente non nulli perchè colonne di una matrice non singolare. Osserviamo subito che se $\overrightarrow{u}$ è un autovettore associato all'autovalore λ, allora anche $t\overrightarrow{u}$ è un autovettore associato allo stesso autovalore, per ogni $0 \neq t \in \mathbb{R}$.

Riassumiamo i risultati precedenti nei seguenti teoremi.

Teorema 6.0.4 *Una matrice A di dimensione n è diagonalizzabile se e solo se esiste una base fatta di autovettori di A.*

Infatti la matrice U tale che $A = U^{-1}\Lambda U$, ovviamente non singolare, deve essere fatta da autovettori. Per essere non singolare, questi devono essere linearmente indipendenti. Poiché devono essere n, costituiscono una base.

Teorema 6.0.5 *Gli autovalori della matrice A si trovano risolvendo l'equazione*

$$\det\,(A - \lambda)I = 0.$$

Infatti λ è un autovalore per A se e solo se esiste $\overrightarrow{u} \neq \mathbf{0}$ tale che $(A - \lambda I)\overrightarrow{u} = \overrightarrow{0}$. Questo significa che il sistema omogeneo, in n equazioni ed n incognite, che ha per matrice dei coefficienti $(A - \lambda I)$, ammette soluzioni diverse dalla soluzione nulla, e quindi la matrice dei coefficiente deve essere singolare.

Il polinomio $\det\,(A - \lambda)I$ è detto *polinomio caratteristico* della matrice A.

Nei prossimi capitoli studieremo più in dettaglio il problema della diagonalizzazione di una matrice.

6.1 Esercizi

6.1.1 Quesiti a risposta chiusa

Rispondere vero o falso.

1. La trasformazione $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2 \quad T(x, y) = (x-y, 3x+y)$ è iniettiva e suriettiva;
2. La trasformazione $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^3 \quad T(x, y) = (2x, -3y, x-y)$ è suriettiva;
3. La trasformazione $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^3 \quad T(x, y, z) = (2x-y, 2x+y+z, 2x-3y-z)$ è iniettiva.

6.1.2 Esercizi aperti

Esercizio 6.1.1 Stabilire se le seguenti trasformazioni sono lineari ed in caso affermativo, calcolare la dimensione del nucleo e dell'immagine:

1. $T : \mathbb{R}^2 \to \mathbb{R}^2 \qquad T((x, y)) = (2x-y, x+y)$;
2. $T : \mathbb{R}^3 \to \mathbb{R}^2 \qquad T((x, y, z)) = (x+1, y)$;
3. $T : \mathbb{R}^2 \to \mathbb{R}^3 \qquad T((x, y)) = (x^2+y^2, x, y)$;
4. $T : \mathbb{R}^3 \to \mathbb{R}^4 \qquad T((x, y, z)) = (x-y, x+y-z, z-x, x)$;
5. $T : \mathbb{R}^2 \to \mathbb{R}^2 \qquad T(x, y) = (x^2, -y^2)$;
6. $T : \mathbb{R}^2 \to \mathbb{R}^2 \qquad T(x, y) = (x, 1)$;
7. $T : \mathbb{R}^2 \to \mathbb{R}^2 \qquad T(x, y) = (y, y)$.

Esercizio 6.1.2 Dopo aver scritto il polinomio caratteristico delle seguenti matrici calcolarne gli autovalori.

1. $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$;
2. $B = \begin{pmatrix} 2 & 1 & 1 \\ 2 & 3 & 4 \\ -1 & -1 & -2 \end{pmatrix}$;
3. $C = \begin{pmatrix} 1 & 3 & 0 \\ 3 & -2 & -1 \\ 0 & -1 & 1 \end{pmatrix}$.

Esercizio 6.1.3 Mostrare, usando il teorema di Binet, che matrici simili hanno lo stesso polinomio caratteristico.

7

Diagonalizzabilità

In questo capitolo siamo interessati al problema di *diagonalizzare* una data matrice A $n \times n$. Si tratta cioè di determinare, se esistono, una matrice U ed una matrice Λ diagonale, tali che valga la formula:

$$A = U\Lambda U^{-1}.$$

Ricordiamo che una matrice $m \times n$ A naturalmente definisce una trasformazione lineare $L : \mathbb{R}^n \to \mathbb{R}^m$: $Lx = Ax$, e che se $A = U\Lambda U^{-1}$, allora le due matrici U e Λ rappresentano la stessa trasformazione lineare, scritta in basi diverse. Quindi diagonalizzare una matrice significa, in termini della trasformazione lineare associata, cercare un cambiamento di base tale che la trasformazione stessa venga rappresentata in un modo particolarmente semplice. Il risultato diventa poi ancora più efficace nel caso in cui la matrice U di trasformazione è fatta di vettori ortogonali: nel prossimo capitolo vedremo quando ciò accade.

La formula $A = U\Lambda U^{-1}$ si può evidentemente anche scrivere in maniera equivalente

$$AU = U\Lambda;$$

se indichiamo con $\overrightarrow{u_j}$ una generica colonna di U, l'espressione precedente può essere riscritta come:

$$A\overrightarrow{u_j} = \lambda_j \overrightarrow{u_j},$$

il che dice due cose allo stesso tempo:

- La matrice Λ deve avere sulla diagonale principale gli autovalori di A;
- La matrice U che è non singolare perché invertibile, deve essere fatta da autovettori. Per essere diagonalizzabile dunque A deve avere n autovettori *linearmente indipendenti.*

Osserviamo che la seconda condizione è anche sufficiente: se la matrice A ammette n autovettori linearmente indipendenti, allora la matrice U le cui colonne sono gli autovettori permette la diagonalizzazione di A.

Dunque per risolvere il problema di diagonalizzare una matrice bisogna andare a studiare i suoi autovalori ed autovettori.

L'equazione caratteristica associata alla matrice A:

$$p(\lambda) = \det (A - \lambda I) = 0$$

è sempre di grado n, infatti il coefficiente di λ^n è $(-1)^n$, qualunque sia la matrice A. Il numero di radici dell' equazione dipende dall'insieme in cui l'equazione viene letta. Anche se la matrice è a coefficienti reali, non è detto che in $\mathbb{R}$ l'equazione abbia soluzioni (o potrebbe averne "poche"), ma in $\mathbb{C}$ l'equazione ha n soluzioni (alcune delle quali possono coincidere) e il polinomio $p(\lambda)$ si può scrivere:

$$p(\lambda) = (-1)^n(\lambda - \lambda_1)(\lambda - \lambda_2)...(\lambda - \lambda_n) :$$

è il teorema fondamentale dell'algebra di cui abbiamo parlato nel primo capitolo. In questa formula non necessariamente tutti i λ_i sono distinti.

Dunque una prima condizione *necessaria* per la diagonalizzabilità di una matrice è che i suoi autovalori siano reali. Questo però, come vedremo, non è sufficiente. Intanto però vediamo un primo risultato interessante.

Proposizione 7.0.1 *Autovettori corrispondenti ad autovalori distinti sono linearmente indipendenti.*

Supponiamo che $\overrightarrow{u_1}$,$\overrightarrow{u_2}$, corrispondenti ai due autovalori $\lambda_1 \neq \lambda_2$, siano linearmente dipendenti; allora esistono t_1, t_2 non contemporaneamente nulli e tali che

$$t_1\overrightarrow{u_1} + t_2\overrightarrow{u_2} = 0. \tag{7.1}$$

Sia ad esempio $t_2 \neq 0$. Allora, moltiplicando la relazione precedente per A si ottiene

$$0 = A(t_1\overrightarrow{u_1} + t_2\overrightarrow{u_2}) = t_1A\overrightarrow{u_1} + t_2A\overrightarrow{u_2} = t_1\lambda_1\overrightarrow{u_1} + t_2\lambda_2\overrightarrow{u_2}. \tag{7.2}$$

Moltiplicando in (7.1) per λ_1, si ottiene

$$\lambda_1t_1\overrightarrow{u_1} + \lambda_1t_2\overrightarrow{u_2} = 0. \tag{7.3}$$

Ora facendo la differenza fra (7.2) e (7.3) si ottiene

$$t_2(\lambda_2 - \lambda_1)\overrightarrow{u_2} = 0,$$

che implica $\overrightarrow{u_2} = 0$ perché $t_2 \neq 0$, e $\lambda_2 \neq \lambda_1$. Ma questo è impossibile perché $\overrightarrow{u_2}$ è un autovettore. Il ragionamento si applica senza cambiamenti ad un numero $p > 2$ di autovalori distinti.

Corollario 7.0.1 *Se una matrice A di ordine n ammette n autovalori reali e distinti, allora è diagonalizzabile.*

In tal caso infatti la matrice ha n autovettori linearmente indipendenti.

Vediamo ora il caso in cui alcuni autovalori siano *multipli.*

Definizione 7.0.1 *Si chiama* molteplicità algebrica dell'autovalore λ *la molteplicità di* λ *come radice del polinomio caratteristico.*

Ad esempio, se per una matrice A, $\lambda_1 = \lambda_2 = \lambda_3$, e gli altri autovalori sono diversi, allora λ_1 è autovalore di molteplicità 3.

Esempio 7.0.1 Calcoliamo gli autovalori della matrice $\begin{pmatrix} 1 & 0 \\ 1 & 3 \end{pmatrix}$.
Dobbiamo costruire la matrice $A-\lambda I$ e vedere per quali valori di λ il suo determinante è uguale a zero. $\begin{pmatrix} 1-\lambda & 0 \\ 1 & 3-\lambda \end{pmatrix}$, $\det(A-\lambda I) = 0 \Longrightarrow (1-\lambda)(3-\lambda) = 0$ (equazione caratteristica), che ha come soluzioni $\lambda = 1$, $\lambda = 3$, due autovalori distinti.

Esempio 7.0.2 Calcoliamo gli autovalori della matrice $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$.
$A - \lambda I = \begin{pmatrix} -\lambda & 1 \\ 0 & -\lambda \end{pmatrix}$ e $\det(A-\lambda I) = 0 \Longrightarrow \lambda^2 = 0$ che ha come soluzione $\lambda = 0$ di molteplicità algebrica 2.

Si chiama *autospazio* relativo all'autovalore λ lo spazio vettoriale generato dagli autovettori relativi a λ. Si dimostra che la sua dimensione è compresa tra 1 e la molteplicità algebrica dell'autovalore.

Definizione 7.0.2 *La dimensione dell'autospazio relativa all'autovalore* λ *si chiama* molteplicità geometrica dell'autovalore.

Ecco il teorema che enuncia quando una matrice è diagonalizzabile.

Teorema 7.0.1 *Una matrice avente autovalori reali* $\lambda_1, \lambda_2, ...\lambda_k$ *di molteplicità algebrica* $m_1, m_2, ...m_k$ *tali che* $\sum m_i = n$ *e molteplicità geometrica* $p_1, p_2, ...p_k$ *è diagonalizzabile se e solo se* $\forall k \quad p_k = m_k$, *cioè se e solo se la molteplicità geometrica e la molteplicità algebrica di ogni singolo autovalore è la stessa.*

Osserviamo esplicitamente che la condizione $\sum m_i = n$ implica che tutti gli autovalori sono reali.

Come procedere per vedere se una matrice è diagonalizzabile

Data una matrice quadrata A:

1. Si calcolano gli autovalori trovando gli zeri del polinomio caratteristico, cioè risolvendo l'equazione $\det(A-\lambda I) = 0$;
2. Se tutti gli autovalori sono reali e distinti, allora la matrice è diagonalizzabile;

3. Se gli autovalori sono reali ma non sono distinti, si va a vedere se la molteplicità algebrica di un autovalore multiplo coincide con la dimensione dell'autospazio. Questo significa vedere da quanti vettori linearmente indipendenti è generato l'autospazio: se il numero di questi vettori coincide con la molteplicità algebrica dell'autovalore, e se questo succede per ogni autovalore multiplo, la matrice è diagonalizzabile; in caso contrario non lo è;
4. Una volta verificato che la matrice è diagonalizzabile una matrice U che ha per colonne una base degli autospazi determinati dai vari autovalori è quella che permette di diagonalizzare A, cioè U è una matrice *modale* di A .

Come si trovano gli autovettori

Seguendo la definizione, gli autovettori relativi all'autovalore λ si determinano risolvendo il sistema $A\overrightarrow{x} = \lambda\overrightarrow{x}$, o equivalentemente $(A-\lambda I)\overrightarrow{x} = \overrightarrow{0}$. Proprio per definizione di autovalore, questo sistema ha soluzioni non banali (cioè diverse dal vettore nullo). Ad esempio, nel caso che A sia una matrice 3×3, il sistema diventa:

$$\begin{pmatrix} a_{11}-\lambda & a_{12} & a_{13} \\ a_{21} & a_{22}-\lambda & a_{23} \\ a_{31} & a_{32} & a_{33}-\lambda \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

cioè

$$\begin{cases} (a_{11}-\lambda)x_1 + a_{12}x_2 + a_{13}x_3 = 0 \\ a_{21}x_1 + (a_{22}-\lambda)x_2 + a_{23}x_3 = 0 \\ a_{31}x_1 + a_{32}x_2 + (a_{33}-\lambda)x_3 = 0 \end{cases}.$$

Questo sistema ha infinite soluzioni: se la matrice $A-\lambda I$ ha rango $r<3$, allora sono ∞^{3-r}. Una base dello spazio delle soluzioni rappresenta una scelta per le colonne della matrice U relative all'autovalore λ.

Esempio 7.0.3 Studiamo la diagonalizzabilità delle seguenti due matrici.
$A = \begin{pmatrix} 0 & -\frac{1}{2} & -1 \\ -12 & 4 & 6 \\ 10 & -4 & -7 \end{pmatrix}$; $B = \begin{pmatrix} 2 & 1 & 1 \\ 2 & 3 & 4 \\ -1 & -1 & -2 \end{pmatrix}$.
Calcoliamo gli autovalori della matrice A: $(A-\lambda I) = \begin{pmatrix} -\lambda & -\frac{1}{2} & -1 \\ -12 & 4-\lambda & 6 \\ 10 & -4 & -7-\lambda \end{pmatrix}$ det $(A-\lambda I) = -\lambda((4-\lambda)(-7-\lambda)+24)+\frac{1}{2}(-12(-7-\lambda)-60)-48+40+10\lambda = -3\lambda^2-\lambda^3+4 = (1-\lambda)(\lambda+2)^2$ det $(A-\lambda I)=0 \Longrightarrow \lambda = 1$, autovalore semplice, e $\lambda=-2$, autovalore doppio. Calcoliamo adesso gli autovettori relativi ai due autovalori trovati.
Per $\lambda = -2$:
$A+2I = \begin{pmatrix} 2 & -\frac{1}{2} & -1 \\ -12 & 6 & 6 \\ 10 & -4 & -5 \end{pmatrix} \Longrightarrow \begin{pmatrix} 2 & -\frac{1}{2} & -1 \\ -12 & 6 & 6 \\ 10 & -4 & -5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ cioè

$$\begin{cases} 2x_1 - \frac{1}{2}x_2 - x_3 = 0 \\ -12x_1 + 6x_2 + 6x_3 = 0 \\ 10x_1 - 4x_2 - 5x_3 = 0 \end{cases},$$

che ha come soluzioni $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}$; un autovettore relativo all'autovalore $\lambda = -2$ è dunque $\overrightarrow{u} = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}$.

Dato che l'autospazio ha dimensione 1, diversa dalla molteplicità dell'autovalore $\lambda = -2$, la matrice non è diagonalizzabile.
Procediamo nello stesso modo con la matrice B.
Autovalori:
$(B - \lambda I) = \begin{pmatrix} 2-\lambda & 1 & 1 \\ 2 & 3-\lambda & 4 \\ -1 & -1 & -2-\lambda \end{pmatrix}$; det $(B - \lambda I) = (2-\lambda)((3-\lambda)(-2-\lambda)+4) - (-4-2\lambda+4)+(-2+3-\lambda) = (\lambda+1)(3-\lambda)(\lambda-1)$ e quindi det $(B-\lambda I) = 0 \Longrightarrow \lambda = 1$, $\lambda = 3$, $\lambda = -1$. La matrice ha tre autovalori distinti, quindi è diagonalizzabile.
Calcoliamo gli autovettori relativi ai tre autovalori trovati:
$\lambda = 1$,
$A - I = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 2 & 4 \\ -1 & -1 & -3 \end{pmatrix} \Longrightarrow \begin{pmatrix} 1 & 1 & 1 \\ 2 & 2 & 4 \\ -1 & -1 & -3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ cioè

$$\begin{cases} x_1 + x_2 + x_3 = 0 \\ 2x_1 + 2x_2 + 4x_3 = 0 \\ -x_1 - x_2 - 3x_3 = 0 \end{cases},$$

che ha come soluzioni

$$\begin{cases} x_1 = -x_2 \\ x_2 = x_2 \\ x_3 = 0 \end{cases},$$

che possiamo scrivere come: $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_2 \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$. Un autovettore relativo all'autovalore $\lambda = 1$ è $\overrightarrow{u_1} = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$.

Allo stesso modo, calcoliamo gli altri autovettori:
per $\lambda = -1$,
$A + I = \begin{pmatrix} 3 & 1 & 1 \\ 2 & 4 & 4 \\ -1 & -1 & -1 \end{pmatrix} \Longrightarrow \begin{pmatrix} 3 & 1 & 1 \\ 2 & 4 & 4 \\ -1 & -1 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$, cioè

$$\begin{cases} 3x_1 + x_2 + x_3 = 0 \\ 2x_1 + 4x_2 + 4x_3 = 0 \\ -x_1 - x_2 - x_3 = 0 \end{cases},$$

che ha come soluzioni

$$\begin{cases} x_1 = 0 \\ x_2 = x_2 \\ x_3 = -x_2 \end{cases},$$

che possiamo scrivere nella forma: $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_2 \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$. Un autovettore relativo all'autovalore $\lambda = -1$ è $\overrightarrow{u_2} = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$;

per $\lambda = 3$,

$A - 3I = \begin{pmatrix} -1 & 1 & 1 \\ 2 & 0 & 4 \\ -1 & -1 & -5 \end{pmatrix} \Longrightarrow \begin{pmatrix} -1 & 1 & 1 \\ 2 & 0 & 4 \\ -1 & -1 & -5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$, cioè

$$\begin{cases} 3x_1 + x_2 + x_3 = 0 \\ 2x_1 + 4x_3 = 0 \\ -x_1 - x_2 - 5x_3 = 0 \end{cases},$$

che ha come soluzioni

$$\begin{cases} x_1 = 2x_3 \\ x_2 = 3x_3 \\ x_3 = -x_3 \end{cases},$$

che si possono riscrivere come: $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_2 \begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix}$. Un autovettore relativo all'autovalore $\lambda = 3$ è $\overrightarrow{u_3} = \begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix}$. Una matrice U tale che $U^{-1}AU = \Lambda$ è dunque $U = \begin{pmatrix} -1 & 0 & 2 \\ 1 & 1 & 3 \\ 0 & -1 & -1 \end{pmatrix}$.

Concludiamo questo capitolo con un'osservazione *metodologicamente* importante. Quanto abbiamo visto ci permette di risolvere il problema della diagonalizzazione *da un punto di vista teorico*. Inoltre, siamo capaci di fare i conti per matrici 2×2, ci spazientiamo con quelle 3×3 e qualche sadico costruisce degli esempi di matrici 4×4

che fa impazzire chi deve fare i conti relativi. Ma a parte il caso in cui i conti sono noiosi eppure fattibili, non bisogna dimenticare che il polinomio caratteristico di una matrice $n \times n$ è di grado n. È vero che il teorema fondamentale dell'algebra ci dice che ha n soluzioni, ma un altro bel teorema ci dice che, a parte casi particolari, è ben difficile fare i calcoli per n superiore a 3. Dunque la teoria sviluppata, che è bella ed elegante, richiede poi all'atto pratico un lavoro, spesso altrettanto interessante, per scovare sistemi che permettano, ad esempio al computer, di diagonalizzare matrici anche di dimensioni maggiori di 3.

7.1 Approfondimenti: il teorema di Cayley-Hamilton

Il prossimo teorema stabilisce un importante risultato che caratterizza una matrice ed il suo polinomio caratteristico. Vediamo dapprima un esempio.

Esempio 7.1.1 Sia data la matrice $A = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$. Il suo polinomio caratteristico è $P(\lambda) = (\lambda - 1)^2 = \lambda^2 - 2\lambda + 1$, come si verifica facilmente col calcolo (o anche senza calcoli, osservando che è triangolare). Mostriamo ora che $A^2 - 2A + I = 0$. Si ha che $A^2 = \begin{pmatrix} 1 & 4 \\ 0 & 1 \end{pmatrix}$, $2A = \begin{pmatrix} 2 & 4 \\ 0 & 2 \end{pmatrix}$, da cui si ottiene facilmente il risultato. Abbiamo dunque visto che, se nel polinomio caratteristico di A sostituiamo a λ la matrice stessa A, si ottiene come risultato la matrice nulla.

Quanto visto sopra non è per nulla casuale, infatti il teorema di Cayley-Hamilton asserisce che questo è sempre vero.

Teorema 7.1.1 *Ogni matrice è uno zero del suo polinomio caratteristico.*

Dimostrazione. Per dimostrare il teorema, scriviamo intanto il determinante di $\lambda I - A$ nella forma:

$$|\lambda I - A| = \lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0. \tag{7.4}$$

Consideriamo poi la matrice $C(\lambda)$ che ha come generico elemento c_{ij} il complemento algebrico corrispondente della matrice $\lambda I - A$: $c_{ij} = (\lambda I - A)_{ji}$, e che ha la proprietà che

$$(\lambda I - A)C(\lambda) = |\lambda I - A| I. \tag{7.5}$$

Osserviamo ora gli elementi di $C(\lambda)$ sono polinomi di grado al massimo $n-1$ (perché ottenuti calcolando determinanti di matrici in cui abbiamo cancellato una riga ed una colonna di $\lambda I - A$). Scriviamo ora $C(\lambda)$ come polinomio di matrici, nella forma

$$C(\lambda) = C_{n-1}\lambda^{n-1} + C_{n-2}\lambda^{n-2} + \cdots + C_1\lambda + C_0. \tag{7.6}$$

Da (7.4), (7.5) e (7.6) abbiamo allora

$$(\lambda I - A)(C_{n-1}\lambda^{n-1} + c_{n-2}\lambda^{n-2} + \cdots + C_1\lambda + C_0) = (\lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0)I. \tag{7.7}$$

Nella equazione (7.7) possiamo ora fare le moltiplicazioni a sinistra, e poi eguagliare i coefficienti delle potenze uguali di λ. Otteniamo così:

$$\begin{aligned} C_{n-1} &= I \\ C_{n-2} - AC_{n-1} &= a_{n-1}I \\ C_{n-3} - AC_{n-2} &= a_{n-2}I \\ \dots & \quad \dots \\ C_0 - AC_1 &= a_1 I \\ -AC_0 &= a_0 I. \end{aligned}$$

Ora moltiplichiamo ambo i membri della prima riga per A^n, della seconda per A^{n-1} ... della penultima per A e dell'ultima per I. Otteniamo allora:

$$\begin{aligned} A^n C_{n-1} &= A^n \\ A^{n-1}C_{n-2} - A^n C_{n-1} &= a_{n-1}A^{n-1} \\ A^{n-2}C_{n-3} - A^{n-1}C_{n-2} &= a_{n-2}A^{n-2} \\ \dots & \quad \dots \\ AC_0 - A^2C_1 &= a_1 A \\ -AC_0 &= a_0 I. \end{aligned}$$

Non ci resta che sommare membro a membro. Osserviamo che a sinistra viene la matrice nulla, mentre a destra viene esattamente il polinomio caratteristico di A calcolato proprio nella matrice A. ■

Esempio 7.1.2 Sia $A = \begin{pmatrix} 1 & 1 & 0 \\ -1 & 0 & 2 \\ 2 & 1 & 0 \end{pmatrix}$ calcolare le potenze di $A : A^2, A^3, A^4$.

Scriviamo il polinomio caratteristico $p_A(t)$ della matrice A :

$p_A(t) = \det \begin{pmatrix} 1-t & 1 & 0 \\ -1 & -t & 2 \\ 2 & 1 & -t \end{pmatrix} = -t^3 + t^2 + t + 2 = 0$, allora $P(A) = -A^3 + A^2 + A + 2I = 0$ che ci permette di ricavare $A^3 = A^2 + A + 2I$.

Essendo $A^2 = A \cdot A = \begin{pmatrix} 1 & 1 & 0 \\ -1 & 0 & 2 \\ 2 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ -1 & 0 & 2 \\ 2 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 2 \\ 3 & 1 & 0 \\ 1 & 2 & 2 \end{pmatrix}$ abbiamo la potenza terza di $A : A^3 = \begin{pmatrix} 3 & 2 & 2 \\ 2 & 3 & 2 \\ 3 & 3 & 4 \end{pmatrix}$.

Ora, $A^4 = A^3 \cdot A = \begin{pmatrix} 3 & 2 & 2 \\ 2 & 3 & 2 \\ 3 & 3 & 4 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ -1 & 0 & 2 \\ 2 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 5 & 5 & 4 \\ 3 & 4 & 6 \\ 8 & 7 & 6 \end{pmatrix}$.

Per calcolare potenze molto elevate, se la matrice A è diagonalizzabile conviene procedere così: portiamo la matrice A in forma diagonale: $M^{-1}AM = D = \begin{pmatrix} d_1 & \dots & 0 \\ 0 & \dots & d_n \end{pmatrix}$. Ne segue che $M^{-1}A^kM = \begin{pmatrix} d_1^k & \dots & 0 \\ 0 & \dots & d_n^k \end{pmatrix}$ e quindi $A^k = M \begin{pmatrix} d_1^k & \dots & 0 \\ 0 & \dots & d_n^k \end{pmatrix} M^{-1}$.

Esempio 7.1.3 Calcolare A^{50} con $A = \begin{pmatrix} 1 & -2 \\ 1 & -2 \end{pmatrix}$.
Gli autovalori di A sono due distinti $\lambda_1 = 0, \lambda_2 = -1$ e due autovettori relativi a questi autovalori sono $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$ e $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$; ne segue che possiamo scegliere M come $M = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$, da cui $M^{-1} = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}$, $D = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$.
Se ne ricava $A^{50} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} -1 & 2 \\ -1 & 2 \end{pmatrix}$.

7.2 Esercizi

7.2.1 Quesiti a risposta chiusa

Rispondere vero o falso.

1. Una matrice $n \times n$ è diagonalizzabile se e solo se ha n autovalori distinti;
2. La matrice $\begin{pmatrix} 1 & -1 & 2 \\ 1 & 2 & 1 \\ 0 & 1 & -1 \end{pmatrix}$ è diagonalizzabile;
3. La matrice $\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{pmatrix}$ è diagonalizzabile.
4. Sia $A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$. Esiste una matrice invertibile U tale che U^tAU è diagonale.
5. Sia $A = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix}$:
   a) A non è diagonalizzabile;

b) $\begin{pmatrix} 0 \\ 4 \\ 0 \end{pmatrix}$ e $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ sono due autovettori linearmente indipendenti di A;

c) $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$ è una base di $\mathbb{R}^3$ formata da autovettori di A.

6. $A = \begin{pmatrix} 1 & a \\ -a & 3 \end{pmatrix}$ è diagonalizzabile per $-1 < a < 1$.

7.2.2 Esercizi aperti

Esercizio 7.2.1 Stabilire se sono diagonalizzabili le seguenti matrici, ed in caso affermativo scrivere una matrice U tale che $U^{-1}AU$ sia diagonale.

1. $\begin{pmatrix} 0 & -\frac{1}{2} & -1 \\ -12 & 4 & 6 \\ 10 & -4 & -7 \end{pmatrix}$;
2. $\begin{pmatrix} 2 & 1 & 1 \\ 2 & 3 & 4 \\ -1 & -1 & -2 \end{pmatrix}$.

Esercizio 7.2.2 Determinare i valori del parametro k per i quali la matrice $A = \begin{pmatrix} k & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & k \end{pmatrix}$ ammette come autovettore $\overrightarrow{x} = \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}$. Fissato questo valore di k, dire se la matrice corrispondente è diagonalizzabile e, in caso affermativo, trovare la matrice diagonale a lei simile.

8

Matrici simmetriche, ortogonali, emisimmetriche

8.1 Matrici simmetriche

Le matrici simmetriche sono una sottoclasse molto importante delle matrici quadrate. In Analisi per esempio, data una funzione $f = f(x, y)$ di due variabili, nello studio dei suoi punti critici (massimi, minimi, selle) interviene una matrice fatta con le sue derivate parziali seconde: $A = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix}$.
In realtà però si ha che, se f è sufficientemente regolare, $f_{xy} = f_{yx}$ e questo dà luogo ad una matrice simmetrica, detta *hessiana* di f.

Tra le proprietà delle matrici simmetriche una delle più rilevanti è che sono sempre diagonalizzabili.

Ricordiamo che le matrici simmetriche sono quelle che verificano la relazione $A = A^t$, e cioè $a_{ij} = a_{ji}$ per ogni i, j.
Scambiando le righe con le colonne la matrice non cambia.

Le seguenti sono proprietà fondamentali delle matrici simmetriche, relative ai loro autovalori ed autovettori. Nel capitolo degli approfondimenti daremo tutte le dimostrazioni delle affermazioni seguenti, dimostrazioni utili da capire ma che si possono evitare ad una prima lettura. Ecco le proprietà:

1. Gli autovalori sono reali;
2. La molteplicità algebrica e geometrica degli autovalori è la stessa;
3. Autovalori distinti danno autovettori ortogonali.

Dimostrazione della proprietà 3. Siano α, β autovalori con $\overrightarrow{u}, \overrightarrow{v}$ autovettori associati. Allora

$$A\overrightarrow{u} = \alpha\overrightarrow{u}, \quad A\overrightarrow{v} = \beta\overrightarrow{v}.$$

Si ha allora:

$$\alpha\overrightarrow{u} \cdot \overrightarrow{v} = \alpha\,\overrightarrow{u}^t\overrightarrow{v} \tag{8.1}$$

$$\alpha\overrightarrow{u} \cdot \overrightarrow{v} = (A\overrightarrow{u})^t\overrightarrow{v} = \overrightarrow{u}^tA^t\overrightarrow{v} = \overrightarrow{u}^tA\overrightarrow{v} = \overrightarrow{u}^t\beta\overrightarrow{v} = \beta\overrightarrow{u}^t\overrightarrow{v}. \tag{8.2}$$

Uguagliando (8.1) e (8.2) si ha allora

$$(\alpha - \beta)\overrightarrow{u}^t\overrightarrow{v} = 0,$$

il che implica o $\alpha = \beta$, oppure che $\overrightarrow{u}$, $\overrightarrow{v}$ sono ortogonali.

Dalla proprietà 2. segue immediatamente:

Teorema 8.1.1 *Sia A una matrice simmetrica. Allora A è diagonalizzabile.*

Dato un autovalore multiplo, l'autospazio corrispondente può essere descritto da una base qualunque $\{b_1, \ldots, b_n\}$. Tuttavia è molto comodo avere una base fatta di vettori ortogonali (ne spiegheremo il motivo nel prossimo paragrafo). Il fatto che per matrici ortogonali autovalori distinti generano autovettori ortogonali ha l'importante conseguenza che la matrice modale U può essere fatta di vettori ortogonali fra loro. Vediamolo innanzitutto con un esempio.

Esempio 8.1.1 Sia data la matrice A:

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

Si ha che $P(\lambda) = (\lambda-1)(\lambda-4)^2$. A $\lambda = 1$ corrisponde l'autovettore $\overrightarrow{u}_1 = (1, -1, 1)$ (o un suo multiplo). Per trovare gli autovettori associati a $\lambda = 4$ si arriva all'equazione $-x + y + z = 0$, da cui, ad esempio, si ottiene $\overrightarrow{u}_2 = (1, 0, 1)$, $\overrightarrow{u}_3 = (1, 1, 0)$.

Verificare che $\overrightarrow{u}_1$ è ortogonale agli altri, ma $\overrightarrow{u}_2, \overrightarrow{u}_3$ non sono ortogonali.

Vogliamo ora, a partire da $\{\overrightarrow{u}_1, \overrightarrow{u}_2, \overrightarrow{u}_3\}$, costruire una base $\{\overrightarrow{w}_1, \overrightarrow{w}_2, \overrightarrow{w}_3\}$ di autovettori ortogonali. L'osservazione cruciale, tanto elementare quanto importante, è che siccome $\overrightarrow{u}_2$ e $\overrightarrow{u}_3$ sono autovettori associati allo stesso autovalore, allora *ogni loro combinazione lineare* è ancora un autovettore relativo allo stesso autovalore. Allora basta porre $\overrightarrow{u}_1 = \overrightarrow{w}_1$, $\overrightarrow{u}_2 = \overrightarrow{w}_2$ e costruire $\overrightarrow{w}_3$, combinazione lineare di $\overrightarrow{u}_2$ e $\overrightarrow{u}_3$, in modo che sia ortogonale $\overrightarrow{w}_1$ e $\overrightarrow{w}_2$. L'ortogonalità rispetto a $\overrightarrow{w}_1$ è automatica. Poniamo allora $\overrightarrow{w}_3 = \overrightarrow{u}_3 - a\overrightarrow{w}_2$, con a da determinarsi attraverso la condizione di perpendicolarità. Ponendo uguale a zero il prodotto scalare di $\overrightarrow{w}_3$ con $\overrightarrow{w}_2$, si ottiene allora $a = 1/2$, e quindi $\overrightarrow{w}_3 = (\frac{1}{2}, -\frac{1}{2}, 1)$, o anche $(1, -1, 2)$.

Posto

$$U = \begin{pmatrix} 1 & 1 & 1 \\ -1 & 1 & -1 \\ -1 & 0 & 2 \end{pmatrix},$$

si ha che

$$U^{-1} = \begin{pmatrix} \frac{1}{3} & -\frac{1}{3} & -\frac{1}{3} \\ \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{6} & -\frac{1}{6} & \frac{1}{3} \end{pmatrix},$$

ed infine

$$U^{-1}AU = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{pmatrix},$$

il che, ricordando che $\lambda = 1$ è autovalore semplice e $\lambda = 4$ è autovalore doppio, ci induce a pensare che i conti che abbiamo fatto potrebbero essere giusti.

8.2 Matrici ortogonali

La formula del cambiamento di base $[\overrightarrow{x}]_B = B^{-1}\overrightarrow{x}$ vista nell'ultimo paragrafo del capitolo sugli spazi vettoriali diventa particolarmente semplice ed elegante nel caso che la matrice B rispetto alla quale si vuole scrivere un generico vettore sia fatta di versori ortogonali fra loro, come visto nel Teorema 4.2.1. Dunque ha certamente interesse studiare le matrici fatte da versori ortogonali. È quanto facciamo adesso.

Definizione 8.2.1 *Si chiama* ortogonale *una matrice B tale che le sue colonne sono versori tra loro ortogonali.*

In altre parole, le colonne di B verificano le relazioni $\overrightarrow{b_i} \cdot \overrightarrow{b_j} = 0$ se $i \neq j$, $\overrightarrow{b_i} \cdot \overrightarrow{b_i} = 1$ per ogni i.

La seguente proposizione è importante.

Proposizione 8.2.1 *Una matrice ortogonale B è invertibile ed inoltre $B^{-1} = B^t$.*

Infatti si ha che $(B^tB)_{ij} = \overrightarrow{b_i} \cdot \overrightarrow{b_j}$

Ecco una proprietà fondamentale delle matrici ortogonali.

Teorema 8.2.1 *Le matrici ortogonali conservano i prodotti scalari, le lunghezze, gli angoli. Cioè, per ogni $\overrightarrow{x}_1, \overrightarrow{x}_2$*

$$B\overrightarrow{x}_1 \cdot B\overrightarrow{x}_2 = \overrightarrow{x}_1 \cdot \overrightarrow{x}_2, \quad \|B\overrightarrow{x}_1\| = \|\overrightarrow{x}_1\|, \quad \text{angolo}\,(B\overrightarrow{x}_1, B\overrightarrow{x}_2) = \text{angolo}\,(\overrightarrow{x}_1, \overrightarrow{x}_2).$$

Infatti $B\overrightarrow{x}_1 \cdot B\overrightarrow{x}_2 = (B\overrightarrow{x}_1)^t B\overrightarrow{x}_2 = \overrightarrow{x}_1^t B^t B\overrightarrow{x}_2 = \overrightarrow{x}_1^t \overrightarrow{x}_2$. Da questo segue anche $\|B\overrightarrow{x}_1\|^2 = B\overrightarrow{x}_1 \cdot B\overrightarrow{x}_1 = \overrightarrow{x}_1 \cdot \overrightarrow{x}_1 = \|\overrightarrow{x}_1\|^2$. La relazione sugli angoli è lasciata per esercizio.

Per questo motivo le trasformazioni $\overrightarrow{y} = B\overrightarrow{x}$, con B ortogonale si chiamano *isometrie*.

Proposizione 8.2.2 *Sia data un'isometria B. Allora* $\det B = \pm 1$. *E se λ è un autovalore di B, allora $|\lambda| = 1$.*

Infatti $1 = \det I = \det(B^{-1}B) = \det B^tB = (\det B^t)\det B = (\det B)^2$. E se $B\overrightarrow{v} = \lambda\overrightarrow{v}$, allora $\|\overrightarrow{v}\| = \|B\overrightarrow{v}\| = |\lambda|\|\overrightarrow{v}\|$, da cui segue $|\lambda| = 1$.

8.3 Matrici ortogonali di ordine 2

Studiamo ora in dettaglio le matrici ortogonali 2×2. Conviene intanto scrivere tali matrici nella forma:

$$\begin{pmatrix} \cos\alpha & \cos\beta \\ \sin\alpha & \sin\beta \end{pmatrix},$$

in modo che i vettori colonna abbiano automaticamente norma uno, per ogni α, β. Dobbiamo ora determinare α e β in modo tale che le colonne siano tra loro ortogonali. Questo porta alla relazione

$$0 = \cos\alpha\cos\beta + \sin\alpha\sin\beta = \cos(\alpha - \beta).$$

Ciò accade se e solo se $\alpha - \beta = \pm\frac{\pi}{2}$. Le matrici ortogonali allora sono del tipo:

$$\begin{pmatrix} \cos\alpha & \cos(\alpha + \frac{\pi}{2}) \\ \sin\alpha & \sin(\alpha + \frac{\pi}{2}) \end{pmatrix},$$

e

$$\begin{pmatrix} \cos\alpha & \cos(\alpha - \frac{\pi}{2}) \\ \sin\alpha & \sin(\alpha - \frac{\pi}{2}) \end{pmatrix},$$

che usando semplici formule di trigonometria diventano:

$$\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$$

e

$$\begin{pmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix}.$$

Osserviamo che, come deve essere, il determinante di tali matrici è o 1 o -1.

Chiamiamo S_α le matrici della forma:

$$\begin{pmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix},$$

ed R_α le matrici della forma:

$$\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}.$$

Notiamo subito che:

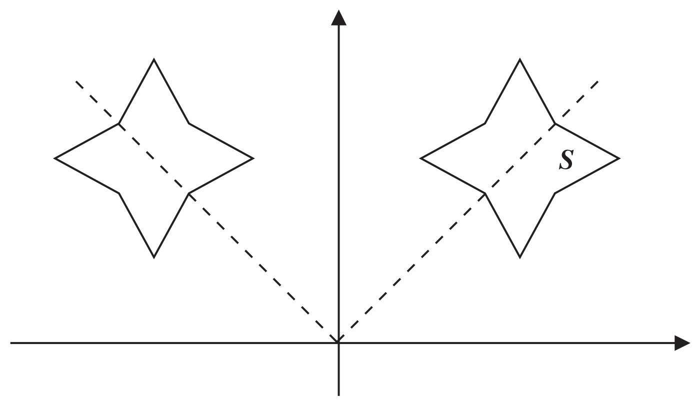
1. S_α è simmetrica e $\det S_\alpha = -1$;
2. $\det R_\alpha = 1$.

8.3.1 Le matrici R_α

Le matrici R_α sono matrici di *rotazione*. Infatti, dalla relazione $v = R_\alpha v'$, con $v = (x, y)$, $v' = (x', y')$,

si ha

$$\begin{cases} x = x' \cos\alpha - y' \sin\alpha \\ y = x' \sin\alpha + y' \cos\alpha \end{cases} .$$



Rotazione di $\frac{\pi}{2}$
Figura 8.1

Casi particolari: $\alpha = (\pi/2)$, $\alpha = \pi$, che danno, rispettivamente:

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

e

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

Esercizio 8.3.1 Fare la verifica della proprietà, dall'ovvio significato geometrico,

$$R_\alpha R_\beta = R_{\alpha+\beta} .$$

L'esercizio precedente mette in particolare in evidenza un'altra proprietà interessante delle matrici di rotazione (di dimensione 2): esse *commutano* tra loro. Infatti $R_\alpha R_\beta = R_{\alpha+\beta} = R_\beta R_\alpha$. L'insieme delle matrici di rotazione ha una struttura particolare, che è fondamentale in matematica, perché interviene in moltissime situazioni, spesso anche in modo inaspettato. Si tratta della struttura di *gruppo*. Chi fosse curioso di sapere la definizione, e vedere qualche esempio di gruppo, può leggerne qualcosa nell'ultimo capitolo degli approfondimenti.

Esercizio 8.3.2 Verificare che, per $\alpha \neq 0, \pi$, R_α non ammette autovalori reali. Questo corrisponde al fatto che una rotazione, diversa da $0, \pi$, non lascia invariata nessuna retta.

8.3.2 Le matrici S_α

Si verifica facilmente che gli autovalori sono $\lambda = 1$ e $\lambda = -1$, per ogni α. Si verifica invece, con conti noiosi, che, per $\alpha \neq 0, \pi$, gli autovettori sono (a meno di multipli) $(1 + \cos\alpha, \sin\alpha)$, corrispondente ad $\lambda = 1$, e $(-1 + \cos\alpha, \sin\alpha)$, corrispondente a $\lambda = -1$. Per S_0 si ha autovettori $(0, 1)$ (per $\lambda = -1$) e $(1, 0)$ (per $\lambda = 1$); per S_π gli autovettori sono $(1, 0)$ (per $\lambda = 1$) e $(0, 1)$ (per $\lambda = 1$).

Esempio 8.3.1

$$S_{\pi/2} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad S_{\pi/2}(x, y) = (y, x).$$

La bisettrice $y = x$ del primo e terzo quadrante è l'autospazio generato da $\lambda = 1$, mentre la bisettrice $y = -x$ del secondo e quarto quadrante è l'autospazio generato da $\lambda = -1$. L'autospazio funziona da *asse di simmetria della trasformazione.*

8.4 Matrici emisimmetriche

Vediamo ora alcune semplici proprietà delle matrici emisimmetriche, cioè quelle matrici A che verificano $A^t = -A$. Negli approfondimenti spieghiamo come questi risultati, che qui sembrano solo molto formali, abbiano un'interpretazione interessante dal punto di vista della teoria dei giochi a somma zero.

Proposizione 8.4.1 *Sia A una matrice emisimmetrica. Allora, per ogni vettore $\overrightarrow{x}$, si ha $\overrightarrow{x}^t A \overrightarrow{x} = 0$.*

Si ha infatti:

$$\overrightarrow{x}^t A \overrightarrow{x} = (\overrightarrow{x}^t A \overrightarrow{x})^t = x^t A^t \overrightarrow{x} = -\overrightarrow{x}^t A \overrightarrow{x}.$$

Proposizione 8.4.2 *Se A è una matrice emisimmetrica $n \times n$ ed n è dispari, allora $\det A = 0$.*

Per le proprietà del determinante, si ha che $\det A = \det A^t$. D'altra parte, sempre per le proprietà del determinante, si ha che $\det -A = (-1)^n \det A$. Ricordando che $A^t = -A$, si conclude che

$$\det A = (-1)^n \det A.$$

Corollario 8.4.1 *Se A è una matrice emisimmetrica $n \times n$ ed n è dispari, allora* 0 *è un autovalore per A.*

8.5 Approfondimenti 1: Matrici simmetriche e diagonalizzazione

Ricordiamo i fatti seguenti, che sono molto utili per il seguito.

1. Date due matrici quadrate A e B di dimensione $n \times n$ la matrice prodotto $C = AB$ ha come colonna di posto j, $C_j = AB_j$, ove ovviamente B_j rappresenta la colonna di posto j di B;
2. Date due matrici A e Λ, con Λ diagonale, la matrice $A\Lambda$ la colonna di posto j è data da $\lambda_j A_j$.

Il prossimo risultato, molto semplice, è utilissimo per dimostrare le proprietà di diagonalizzabilità delle matrici simmetriche.

Proposizione 8.5.1 *Sia A una matrice $n \times n$ e siano $\lambda_1, \ldots, \lambda_s$ autovalori di A cui corrispondono autovettori $\overrightarrow{x_1}, \ldots, \overrightarrow{x_s}$ linearmente indipendenti. Allora esiste una matrice $n \times n$ P tale che*

$$A = PCP^{-1},$$

con C della forma

$$C = \left(\begin{array}{ccc|c} \lambda_1 & \ldots & 0 & \\ \vdots & \ddots & \vdots & T \\ 0 & \ldots & \lambda_s & \\ \hline 0 & \ldots & 0 & \\ 0 & \ldots & 0 & B \end{array}\right)$$

Dimostrazione. Basta evidentemente trovare una matrice P che ci permetta di verificare la relazione voluta. Prendiamo s autovettori linearmente indipendenti, relativi agli autovalori dati. Aggiungiamo ad essi $n - s$ vettori in modo da formare una base dello spazio. Costruiamo P mettendo come colonne ordinatamente gli autovettori, ed aggiungendo (anche alla rinfusa) gli altri della base. È facile verificare che $P^{-1}AP$ ha la forma della matrice C. ■

La matrice C, come si vede, contiene un minore $s \times s$ di Nord-Ovest che è in forma diagonale. Sotto di esso c'è la matrice nulla di dimensione $(n - s) \times s$. Completano la matrice $(n \times n)$ due opportune matrici T e B di dimensioni $s \times (n - s)$ e $(n - s) \times (n - s)$ (ne mettiamo due per comodità di scrittura, si potrebbe considerare un'unica matrice di dimensione $n \times (n - s)$).

Esempio 8.5.1 Consideriamo di nuovo la matrice seguente, già vista nell'esempio 7.0.3. $A = \begin{pmatrix} 0 & -\frac{1}{2} & -1 \\ -12 & 4 & 6 \\ 10 & -4 & -7 \end{pmatrix}$; avevamo trovato che i suoi autovalori sono $\lambda_1 = 1$, e $\lambda_2 = \lambda_3 = -2$, con molteplicità doppia. Un autovettore per -2 è dato da $(1, 0, 2)$, per -1 da $(0, 2, -1)$. Per formare la matrice P consideriamo ad esempio il vettore $(0, -1, 0)$ (scelto col criterio che formi una base con gli altri 2, che contenga tanti zeri per semplificare i calcoli, e che P abbia determinante 1 sempre per semplificare i calcoli). Si ha allora che $P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -2 & -1 \\ 2 & 1 & 0 \end{pmatrix}$. La matrice C è della forma $C = \begin{pmatrix} -2 & 0 & a \\ 0 & 1 & b \\ 0 & 0 & c \end{pmatrix}$, con a, b, c da determinare. Uguagliando i prodotti: $AP = PC$, si ottiene infine: $C = \begin{pmatrix} -2 & 0 & -\frac{1}{2} \\ 0 & 1 & 3 \\ 0 & 0 & -2 \end{pmatrix}$.

Vogliamo ora dimostrare quanto affermato sopra sulle matrici simmetriche, e cioè che hanno autovalori reali e sono diagonalizzabili. Per fare questo, consideriamo allora vettori a componenti complesse, non solo reali, e matrici a valori complessi. Definiremo sempre autovettore di una matrice complessa A, associato all'autovalore complesso λ, un vettore complesso $\overrightarrow{z}$ non nullo che verifica la relazione $A\overrightarrow{z} = \lambda\overrightarrow{z}$.

La prima cosa che vogliamo verificare è che gli autovalori di una matrice simmetrica sono reali. Per esercizio, vediamo intanto il caso di una matrice 2×2. In tal caso si ha $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$. Il suo polinomio caratteristico è

$$\lambda^2 - (a + c)\lambda + (ac - b^2).$$

il discriminante dell'equazione è dunque

$$(a - c)^2 + 4b^2,$$

il che dice che gli autovalori sono reali, e distinti tranne il caso in cui $a = c$, $b = 0$, che implica che A sia un multiplo della matrice identica. Ne segue allora, tra l'altro, che A è diagonalizzabile.

Passiamo ora a dimostrare che gli autovalori di una matrice simmetrica sono reali. Cominciamo col definire che cosa sia la matrice trasposta di una matrice complessa A. Se $A = (z_{ij})$ con $z_{ij} = a_{ij} + ib_{ij}$, definiamo $A^t = w_{ij}$, con $w_{ij} = a_{ji} - ib_{ji}$. In particolare, si ha, per un vettore complesso $\overrightarrow{a} + i\overrightarrow{b}$, che il suo trasposto è: $(a_1 - ib_1, \ldots, a_n - ib_n)$. Notare inoltre che, per $\lambda \in \mathbb{C}$, si ha che $(\lambda\overrightarrow{z})^t = \bar{\lambda}\overrightarrow{z}^t$. Notare anche che se A è reale simmetrica, anche se la vediamo nei complessi vale sempre la formula $A = A^t$.

Proposizione 8.5.2 *Sia A una matrice simmetrica (reale). Allora i suoi autovalori sono reali.*

Dimostrazione. Sia λ un autovalore di A, e mostriamo che è reale. Sia $\overrightarrow{z}$ un autovettore (complesso) della matrice A, che pensiamo a valori complessi. Indichiamo con a il numero reale positivo $a = \overrightarrow{z}^t\overline{\overrightarrow{z}}$. Si ha allora

$$\lambda a = (\lambda\overrightarrow{z}^t)\overline{\overrightarrow{z}} = (A\overrightarrow{z})^t\overline{\overrightarrow{z}} = \overrightarrow{z}^t A\overline{\overrightarrow{z}}.$$

Da questo si ricava:

$$\lambda a = \overrightarrow{z}^t A \overline{\overrightarrow{z}} = (A\overrightarrow{z})^t \overline{\overrightarrow{z}} = (\lambda \overrightarrow{z})^t \overline{\overrightarrow{z}} = \bar{\lambda} \overrightarrow{z}^t \overline{\overrightarrow{z}} = \bar{\lambda} a,$$

da cui si conclude. ■

Vogliamo adesso vedere che una matrice simmetrica è diagonalizzabile. Intanto, i suoi autovalori sono reali. Se sono distinti, allora il risultato è ovvio in quanto autovettori associati sono automaticamente linearmente indipendenti. Vediamo come dimostrare il risultato nel caso di autovalori non necessariamente distinti.

Teorema 8.5.1 *Sia A una matrice simmetrica. Allora A è diagonalizzabile.*

Dimostrazione. Lo dimostriamo per induzione sul numero di righe/colonne della matrice A. Se A è 1×1 o 2×2, è diagonalizzabile. Supponiamo ora che le matrici $n-1 \times n-1$ siano diagonalizzabili, e sia A una matrice $n \times n$. Sia λ un suo autovalore. Utilizziamo la Proposizione 8.5.1 per scrivere

$$A = PCP^{-1},$$

con P matrice *ortogonale* (la prima colonna di P è un autovalore associato a λ, che prendiamo di lunghezza 1, completiamo poi P con colonne tra loro ortogonali e di lunghezza 1, ad esempio col procedimento di diagonalizzazione di Graham-Schmidt) e C della forma

$$C = \left(\begin{array}{c|c} \lambda_1 & \mathrm{T} \\ \hline 0 & \mathrm{B} \end{array}\right).$$

Ora è facile vedere che C è simmetrica, poiché lo è A e P è ortogonale. Dunque, necessariamente $T = 0$ e B simmetrica. Ma B è di dimensione $n-1 \times n-1$, per cui diagonalizzabile per ipotesi induttiva, e questo completa la dimostrazione. ■

8.6 Approfondimenti 2: Matrici emisimmetriche e giochi

Le matrici emisimmetriche compaiono naturalmente nell'ambito dello studio dei giochi fra due giocatori, che hanno a disposizione un numero finito di mosse, e che sono strettamente competitivi. Un gioco a due persone si dice strettamente competitivo (o a somma zero) se quel che guadagna l'uno è esattamente quel che perde l'altro. Un tipico gioco di questa natura è quello in cui uno dei due vince, oppure si ottiene il pareggio, ad esempio la dama, gli scacchi, la morra cinese etc. Giochi di questo tipo si possono, almeno in linea teorica, descrivere tramite matrici. Specificando infatti quel che un giocatore potrebbe fare in ogni situazione in cui potrebbe essere chiamato a muovere, e facendo questo per i due giocatori, è chiaro che si riesce a descrivere completamente l'evoluzione del gioco, e quindi il suo esito. Tuttavia fare questo, anche a partire da un gioco semplice, come ad esempio il tris, non è una cosa banale, e per giochi più complicati, tipo gli scacchi, non è possibile nemmeno al più potente dei computers. Ma supponendo che i giochi (come succede nella maggior parte dei casi) abbiano un numero finito di evoluzioni possibili (e quindi di strategie), possiamo pensare di prendere le strategie come dato primitivo del problema, e modellizzare il gioco appunto con una matrice, che si costruisce così: supponiamo che un giocatore (chiamato convenzionalmente il primo) abbia m scelte (strategie) possibili ed il secondo n. Ad ogni coppia di scelte dei giocatori corrisponde un esito del gioco: supponiamo di indicare con a_{ij} quanto il secondo paga al primo se il primo gioca la scelta i ed il secondo la scelta j. Allora è chiaro che la matrice $m \times n$ $A = (a_{ij})$ descrive efficacemente il gioco. Vediamo un esempio, forse il più semplice: due giocatori giocano a pari e dispari, mostrando contemporaneamente le dita di una mano. Il primo vince se la somma delle dita mostrate è pari, altrimenti vince il secondo. Tutto questo è efficacemente descritto dalla seguente matrice 2×2, ove si assegna convenzionalmente 1 a chi vince:

$$\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}.$$

Abbiamo assegnato, ad esempio, la scelta di un numero pari da parte del primo (secondo) alla prima riga (colonna), ed abbiamo riportato in tabella quanto guadagna il primo. Ad esempio se il primo sceglie la prima riga (tira pari) ed il secondo la seconda colonna (tira dispari), la somma viene dispari, vince il secondo, e quindi il coefficiente a_{12} della matrice diventa -1, avendo convenzionalmente assegnato la quantità 1 a chi vince. La teoria dei giochi a somma zero è molto interessante, ma non possiamo descriverla qui. Però possiamo almeno osservare che le matrici emisimmetriche, cioè le matrici A che verificano la relazione $A = -A^t$, hanno un ruolo importante in questa teoria, in quanto descrivono i giochi "onesti", cioè quelli che danno le stesse chances ai due giocatori. Infatti, essendo ad esempio la prima riga uguale alla prima colonna, ma con i segni cambiati, questo significa che quel che può vincere il primo con la prima riga è esattamente quel che può vincere il secondo con la prima colonna. Perciò ora cerchiamo di vedere alcune semplici proprietà delle matrici emisimmetriche, anche alla luce dell'interpretazione che ne possiamo dare in base alla teoria dei giochi. Senza entrare in troppi dettagli, è chiaro che alcuni di questi giochi non hanno la proprietà che il loro esito è prevedibile *a priori*, anche supponendo che i giocatori siano intelligenti. Nel gioco del pari e del dispari, è abbastanza evidente che l'esito è imprevedibile. Tuttavia, soprattutto se lo giochiamo parecchie volte contro lo stesso avversario, è chiaro che ci sono modi più intelligenti e modi meno intelligenti di giocarlo. Ad esempio, se fossi il primo, che vince quando la somma è pari, scegliere la strategia di gioco: la prima volta tiro a caso, poi tiro sempre quel che il secondo ha giocato al round precedente, non è intelligente, perché dopo un po' il mio avversario se ne accorge e mi frega per sempre.
Dunque, è lecita la domanda se anche nel caso in cui apparentemente il gioco non ha equilibrio, esista un modo ottimale per giocarlo. L'idea diventa allora di giocare ogni riga (colonna) con una certa probabiltà, e di calcolare il guadagno come valore atteso. Questo significa che se vinco 1 con una probabilità del 50% e perdo 1 con una probabilità del 50%, io ne deduco che guadagno zero $(0 = (1/2)1 + (1/2)(-1))$ con certezza[1]. Supponiamo allora che il secondo giocatore giochi la colonna j con

[1] Probabilmente le prime volte che si legge una cosa del genere qualche dubbio ci rimane che questo sistema di calcolare i guadagni "attesi" sia davvero corretto: in fondo non è molto

probabilità $y_j \geq 0$, $j = 1, \ldots, n$, con $y_1 + \cdots + y_n = 1$. Indicando con $\overrightarrow{y}$ il vettore colonna corrispondente, che significato possiamo dare al vettore $A\overrightarrow{y}$? È chiaro che la sua i-esima componente è quanto il giocatore 1 si aspetta di ottenere se giocasse la iesima riga. E quindi, se il vettore $\overrightarrow{x}$ indica le probabilità che il primo decide di usare il risultato atteso finale è rappresentato dal numero reale $\overrightarrow{x}^t A\overrightarrow{y}$. Alla luce di questa interpretazione, il risultato seguente, che abbiamo già visto, è davvero naturale.

Proposizione 8.6.1 *Sia A una matrice emisimmetrica. Allora, per ogni vettore $\overrightarrow{x}$, si ha $\overrightarrow{x}^t A\overrightarrow{x} = 0$.*

Questo significa semplicemente che in un gioco equo, se i giocatori fanno le stesse mosse, il risultato (atteso!) è il pareggio.

Una strategia *ottimale* per il primo giocatore è una strategia $\overrightarrow{\overline{x}}$ tale che

$$\overrightarrow{\overline{x}}^t A\overrightarrow{y} \geq 0,$$

per ogni strategia $\overrightarrow{y}$ del secondo: infatti in questo modo il primo vince (o pareggia, e sempre come valore atteso) contro ogni strategia del secondo. Analogamente per il secondo (con i segni cambiati, naturalmente).

Proposizione 8.6.2 *Se $\overrightarrow{\overline{x}}$ è una strategia ottimale per il primo giocatore, lo è anche per il secondo.*

Si ha che

$$0 \leq (\overrightarrow{\overline{x}}^t A\overrightarrow{y})^t = \overrightarrow{y}^t A^t \overrightarrow{\overline{x}} = -\overrightarrow{y}^t A\overrightarrow{\overline{x}},$$

da cui si ottiene facilmente il risultato.

Proposizione 8.6.3 *Se $\overrightarrow{\overline{x}}$ e $\overrightarrow{\overline{y}}$ sono strategie ottimali per il primo e per il secondo, rispettivamente, si ha allora $\overrightarrow{\overline{x}}^t A\overrightarrow{\overline{y}} = 0$.*

Le proposizioni precedenti, sebbene facili da dimostrare, hanno una dimostrazione assai noiosa da leggere. Però, alla luce della loro interpretazioni come giochi a somma zero, i loro risultati sono molto naturali.

Dunque, come ci si poteva aspettare, un gioco equo a somma zero fra due giocatori intelligenti dà come risultato, in media, il pareggio. Un problema interessante in tali giochi è quello di trovare strategie ottimali di equilibrio, questione non banale se le strategie in generale sono molte.

Che cosa deve fare il giocatore uno? Deve cercare $\overrightarrow{\overline{x}}$ tale che $\overrightarrow{\overline{x}}^t A\overrightarrow{y} \geq 0$ qualunque $\overrightarrow{y}$. Indicando con a_j la j-esima colonna della matrice A, allora è necessario e sufficiente che

intuitivo pensare che se due giocano una sola volta il pari e dispari tirando una moneta, allora il loro guadagno sarà zero, visto che questo *non accade mai.* Tuttavia, considerando che i modelli sono interessanti soprattutto se ripetuti, appare molto intuitivo che giocando il gioco del pari e dispari parecchie volte affidandosi ad una moneta per decidere, alla fine saremo davvero più o meno in una situazione di pareggio

$$\langle a_j, \overrightarrow{x} \rangle \geq 0$$

per ogni j (infatti $\overrightarrow{x}^t Ay = \sum_j \langle a_j, \overrightarrow{x} \rangle$). Supponiamo inoltre che $\langle a_j, \overrightarrow{x} \rangle \geq 0$ per ogni $j \neq h$, e che $\langle a_k, \overrightarrow{x} \rangle > 0$. Allora, dato un vettore $\overrightarrow{y}$, se $y_k > 0$, si ha che $\overrightarrow{x}^t A \overrightarrow{y} > 0$. Questo dice la cosa, molto ragionevole, che se io sono in grado di trovare una strategia mista che mi fa guadagnare almeno zero contro tutte le colonne, e più di zero contro la colonna k, allora per il giocatore due non è conveniente giocare la colonna k, che quindi lui giocherà con probabilità nulla.

In conclusione, se si sapesse che *tutte* le strategie vanno giocate con probabilità positiva, una strategia ottimale si trova semplicemente risolvendo il sistema $A\overrightarrow{x} = 0$ (con le condizioni ausiliarie $x_i > 0 \sum_i x_i = 1$). D'altra parte, perché tale sistema abbia soluzione non nulla, occorre che il determinante di A sia nullo. Ecco che il risultato già visto, che riportiamo qui sotto, ci dà utili informazioni in questo senso.

Proposizione 8.6.4 *Se A è una matrice emisimmetrica $n \times n$ ed n è dispari, allora det $A = 0$.*

8.7 Esercizi

8.7.1 Quesiti a risposta chiusa

Rispondere vero o falso.

1. Siano A una matrice simmetrica e Λ una matrice diagonale della stessa dimensione. Allora $A\Lambda = (\Lambda A)^t$.
2. Siano A B matrici simmetriche della stessa dimensione.
   a) $A + B$ è una matrice simmetrica;
   b) AB è una matrice simmetrica.
3. Gli autovettori relativi agli autovalori della matrice $A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix}$ sono ortogonali.
4. Sia A matrice emisimmetrica. Allora A^2 è diagonalizzabile.
5. Sia A diagonalizzabile da una matrice ortogonale. Allora A è simmetrica.

8.7.2 Esercizi aperti

Esercizio 8.7.1 Sia A una matrice $n \times n$. Mostrare che AA^t è diagonalizzabile.

Esercizio 8.7.2 Sia A una matrice emisimmetrica. Mostrare che $I + A$ è non singolare, e che $P = (I - A)(I + A)^{-1}$ è ortogonale.

Suggerimento: provare che il sistema $(I + A)x = 0$ ha solo la soluzione nulla, scrivendo

$$0 = \langle I + Ax, x \rangle = \dots,$$

Poi osservare che $I + A$ e $I - A$ commutano e che $P^t P = I$.

Esercizio 8.7.3 Sia A una matrice simmetrica. Le seguenti condizioni sono equivalenti:

(a) A è ortogonale;
(b) $A^2 = I$;
(c) gli autovalori di A sono contenuti nell'insieme $\{-1, 1\}$.

Suggerimento: ricordare che $A = P^t \Lambda P$, per una matrice opportuna P e con $\Lambda \ldots$.

Esercizio 8.7.4 Sia A la matrice $\begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ h & 1 & 0 \end{pmatrix}$

1. Determinare per quali valori di h la matrice A è diagonalizzabile.
2. Stabilire se esistono valori di h in corrispondenza dei quali la matrice A è simile alla matrice $B = \begin{pmatrix} \frac{1}{2} & 1 & 2 \\ 0 & -1 & -3 \\ 0 & 0 & \frac{1}{2} \end{pmatrix}$

Esercizio 8.7.5 Data la matrice $A = \begin{pmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{pmatrix}$

1. calcolarne autovalori ed autovettori
2. determinare una matrice ortogonale che diagonalizza A.

Esercizio 8.7.6 Trovare una base ortonormale di $\mathbb{R}^3$ formata da autovettori della matrice

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

9

Coniche

9.1 La circonferenza

Come ben sappiamo, la circonferenza è il luogo geometrico dei punti del piano equidistanti da un punto prefissato, detto *centro*. La distanza dal centro si chiama il *raggio* della circonferenza. Se $P_0(x_0, y_0)$ è il centro ed r è il raggio, indicato con $P(x, y)$ un generico punto della circonferenza, l'equazione diventa: $|PP_0|^2 = r^2$, che rappresenta *l'equazione cartesiana della circonferenza*. Svolgendo i calcoli:

$$x^2 + y^2 + \alpha x + \beta y + \gamma = 0, \tag{9.1}$$

con $x_0 = -\alpha/2$, $y_0 = -\beta/2$, $r = \sqrt{x_0^2 + y_0^2 - \gamma}$. Naturalmente affinché un'equazione come la (9.1) rappresenti una circonferenza deve essere $\alpha^2/4 + \beta^2/4 - \gamma > 0$ (cioè il raggio deve essere positivo). Naturalmente, anche l'equazione $k(x^2+y^2+\alpha x+\beta y+\gamma) = 0$ rappresenta una circonferenza (se $\alpha^2/4 + \beta^2/4 - \gamma > 0$).

La tecnica dei fasci di rette e di piani può naturalmente essere estesa ai fasci di circonferenze. Siano date due circonferenze: C_1: $x^2 + y^2 + \alpha x + \beta y + \gamma = 0$ e C_2: $x^2 + y^2 + \hat{\alpha} x + \hat{\beta} y + \hat{\gamma} = 0$. Allora il fascio delle due circonferenze si ottiene facendo una loro combinazione lineare:

$$\lambda(x^2 + y^2 + \alpha x + \beta y + \gamma) + \hat{\lambda}(x^2 + y^2 + \hat{\alpha} x + \hat{\beta} y + \hat{\gamma}) = 0, \tag{9.2}$$

con λ e $\hat{\lambda}$ non contemporaneamente nulli. L' equazione (9.2) rappresenta ancora una circonferenza (infatti i coefficienti dei termini di secondo grado sono uguali), tranne nel caso in cui $\lambda = -\hat{\lambda}$. Infatti in questa situazione spariscono i termini di secondo grado e l'equazione rappresenta una retta, detta *asse radicale*. Tutte le circonferenze del fascio (compreso l'asse radicale) passano per gli eventuali punti di intersezione delle due circonferenze C_1 e C_2.

9.1.1 Problemi geometrici

Vediamo ora come risolvere alcuni problemi geometrici in cui intervengono le circonferenze. Spesso questi problemi hanno risoluzioni elementari, già viste nello studio della geometria analitica nelle scuole secondarie: senza dimenticare quel tipo di risoluzione, che a volte è quello più naturale, vogliamo qui utilizzare anche alcuni strumenti introdotti in questo corso.

Esempio 9.1.1 Consideriamo il problema di determinare il luogo geometrico dei punti del piano tale che il rapporto tra le distanze da due punti dati sia costante. Chiamati $P_1(x_1, y_1)$ e $P_2(x_2, y_2)$ i punti in questione, $P(x, y)$ un generico punto del luogo, dobbiamo allora imporre che

$$\frac{|PP_1|}{|PP_2|} = h,$$

con h costante prefissato. Osserviamo che nel caso $h = 1$ si tratta di determinare l'asse del segmento di estremi P_1 e P_2, cioè la retta perpendicolare al segmento e passante per il suo punto medio.

Svolgendo i calcoli si ottiene la famiglia di equazioni (al variare del parametro h):

$$(x^2 + y^2)(1 - h^2) - 2x(x_1 - h^2 x_2) - 2y(y_1 - h^2 y_2) + x_1^2 + y_1^2 - h^2(x_2^2 + y_2^2) = 0,$$

che rappresenta una famiglia di circonferenze, dette *cerchi di Apollonio*. Notare che, come deve essere, nel caso $h = 1$ la circonferenza "degenera" in una retta, l'asse del segmento.

Esempio 9.1.2 Dati tre punti non allineati $P_1(x_1, y_1)$, $P_2(x_2, y_2)$, $P_3(x_3, y_3)$, trovare la circonferenza che li contiene.

Vediamo come si può risolvere.

1. Primo metodo. Si trovano le equazioni degli assi dei due segmenti P_1P_2 e P_1P_3, che si incontrano in un punto P_0 (in quanto i tre punti non sono allineati) il quale, appartenendo ad entrambi gli assi, è equidistante dai tre punti dati e quindi è il centro della circonferenza. Il raggio poi si trova calcolando la distanza $|P_0P_i|$ (prendendo come i il più conveniente fra $1, 2, 3$ per fare i calcoli).
2. Secondo metodo. Si scrive l'equazione cartesiana della circonferenza: $x^2+y^2+\alpha x+\beta y + \gamma = 0$, e si impone che passi per i tre punti. Si ottiene un sistema lineare di tre equazioni nelle incognite α, β, γ, che ammette una ed una sola soluzione (verificare che il determinante della matrice dei coefficienti è non nullo, grazie alla condizione imposta che i tre punti non siano allineati);
3. Terzo metodo. Si pone uguale a zero il determinante della seguente matrice:

$$\begin{pmatrix} x^2 + y^2 & x & y & 1 \\ x_1^2 + y_1^2 & x_1 & y_1 & 1 \\ x_2^2 + y_2^2 & x_2 & y_2 & 1 \\ x_3^2 + y_3^2 & x_3 & y_3 & 1 \end{pmatrix}.$$

   Chiaramente questo metodo è operativo solo nel caso in cui si riesca a mettere un po' di zeri nella matrice precedente, il che accade di rado. Ciò mostra che l'idea è interessante soprattutto come fatto teorico. Per convincerci che questo metodo è ragionevole, basta considerare che annullare il determinante della matrice dà in effetti l'equazione di una circonferenza, e che la circonferenza passa dai tre punti richiesti (sostituendo nella prima riga a (x, y) le coordinate di uno dei tre punti il determinante si annulla perché ci sono due righe uguali);

4. Quarto metodo. Scriviamo il fascio delle circonferenze che passano per due punti, ad esempio P_1 e P_2, con questi accorgimenti: utilizziamo l'asse radicale, scrivendo l'equazione della retta per due punti, e la circonferenza che ha per diametro il segmento di estremi P_1 e P_2; questo è facile perché il centro di tale circonferenza è il punto medio del segmento, ed il raggio è metà del segmento di estremi P_1 e P_2; nell'equazione del fascio così ottenuta (equazione che dipende da un parametro), imponiamo il passaggio per il punto P_3, e troviamo la circonferenza cercata.

Vediamo su un esempio, come utilizzare i due metodi.

Esercizio 9.1.1 Scrivere l'equazione della circonferenza passante per $P_1(0, 0)$ $P_2(2, 0)$ e $P_3(1, 1)$.

1. Poiché l'asse del segmento di estremi $(0, 0)$ e $(0, 2)$ è la retta $x = 1$, che contiene il punto $(1, 1)$, si ha immediatamente che il centro è $(1, 0)$ e il raggio è 1. Quindi $(x - 1)^2 + y^2 = 1$, cioè $x^2 - 2x + y^2 = 0$;
2. Il sistema viene $\begin{cases} \gamma = 0 \\ 4 + 2\alpha + \gamma = 0 \\ 2 + \alpha + \beta + \gamma = 0 \end{cases}$ che ha come soluzione $\alpha = -2$, $\beta = \gamma = 0$;
3. Si tratta di calcolare il determinante della matrice
$$\begin{pmatrix} x^2 + y^2 & x & y & 1 \\ 0 & 0 & 0 & 1 \\ 4 & 2 & 0 & 1 \\ 2 & 1 & 1 & 1 \end{pmatrix};$$
4. L'asse radicale è $y = 0$ e la circonferenza per $(0, 0)$ e $(2, 0)$ è $x^2 + y^2 - 2x = 0$ per cui il fascio è dato da $\alpha y + \beta(x^2 + y^2 - 2x) = 0$: imponendo il passaggio per $(1, 1)$ si ottiene $\alpha = 0$.

Esempio 9.1.3 Come determinare l'equazione di una circonferenza di cui si conosce un punto P_1 per cui passa, una retta a cui è tangente e le coordinate del punto P_2 di tangenza. Ecco due metodi possibili.

- Primo metodo. Troviamo il centro, che è l'intersezione fra la retta perpendicolare alla tangente e passante per P_2, e la retta asse del segmento di estremi P_1 e P_2. Il raggio si determina calcolando la distanza del centro precedentemente trovato dal punto P_2;
- Secondo metodo. Costruiamo il fascio di circonferenze per P_1 e P_2 come nell'Esempio 9.1.2. Mettiamo a sistema retta tangente e fascio, imponiamo poi che l'equazione di secondo grado abbia una soluzione sola (o due coincidenti, che dir si voglia), ponendo il discriminante uguale a zero.

Vediamo su un esempio, come utilizzare i due metodi.

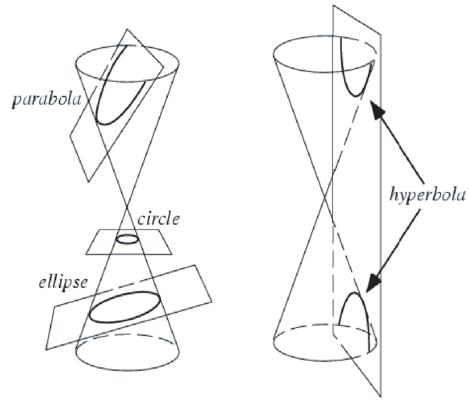
Esercizio 9.1.2 Scrivere l'equazione della circonferenza passante per $P(2,-2)$ e tangente alla retta $r : x+2y+1=0$ nel punto $A(-1,0)$.

1. Scriviamo la retta s perpendicolare alla retta r in A: $s : y-2x-2=0$. Scriviamo l'equazione dell'asse del segmento $\overline{AP}$ calcolando le coordinate del punto medio M di tale segmento ($M(\frac{1}{2},-1)$) ed imponendo che la retta che stiamo cercando sia perpendicolare alla retta passante per a e P. L'asse t ha equazione: $-4y+6x-7=0$. Intersechiamo adesso le rette s e t per trovare il centro C della circonferenza: $\begin{cases} -2x+y-2=0 \\ -4y+6x-7=0 \end{cases}$, $C(\frac{-15}{2};-13)$. Il raggio della circonferenza si trova calcolando la distanza tra A e C (oppure tra P e C) $AC=\frac{13}{2}\sqrt{5}$. A questo punto possiamo scrivere l'equazione della circonferenza richiesta:
$$(x+\frac{15}{2})^2+(y+13)^2=\frac{845}{4}.$$
2. Costruiamo il fascio di circonferenze passanti per due punti scrivendo la circonferenza che ha per diametro $\overline{AP}$ e l'asse radicale(la retta AP). La circonferenza di diametro $\overline{AP}$ ha equazione: $(x-\frac{1}{2})^2+(y+1)^2=\frac{13}{4}$, la retta AP ha equazione $3y+2x+2=0$. Il fascio di circonferenze passanti per A e P ha equazione $h((x-\frac{1}{2})^2+(y+1)^2-\frac{13}{4})+k(3y+2x+2)=0$ che messo a sistema con la retta tangente ed imponendo che l'equazione di secondo grado che si ottiene abbia due soluzioni coincidenti, ci dà la condizione $k=8h$ che, sostituita nel fascio, fornisce l'equazione della circonferenza cercata.

9.2 Coniche

La circonferenza appartiene ad una famiglia di curve dette coniche perché si ottengono intersecando un piano con un cono di rotazione: una superficie illimitata ottenuta dalla rotazione di una retta attorno ad un'altra retta ad essa incidente; l'asse di rotazione si chiama *asse del cono* e le rette ottenute dalle successive posizioni della retta rotante si chiamano *generatrici del cono*, l'angolo α formato dalle generatrici con l'asse di rotazione si chiama *angolo di semiapertura del cono*. A seconda della giacitura del piano secante rispetto all'asse di rotazione, si ottengono diverse curve:

1. Ellisse. L'angolo tra asse di rotazione e piano secante è maggiore dell'angolo di semiapertura α. In particolare si ottiene la circonferenza se il piano secante è perpendicolare all'asse;
2. Parabola. L'angolo tra asse di rotazione e piano secante è uguale all' angolo di semiapertura α;
3. Iperbole. L'angolo tra asse di rotazione e piano secante è minore dell' angolo di semiapertura α.



Le sezioni coniche
Figura 9.1

Se il piano passa per il vertice del cono si hanno casi particolari, che vanno sotto il nome di *coniche degeneri*:

1. Ellisse. Si riduce ad un punto solo;
2. Parabola. Due rette coincidenti;
3. Iperbole. Due rette distinte (dette *generatrici*).

Queste definizioni hanno il pregio di visualizzare il modo in cui si ottiene una conica, ma non sono semplici da usare perché descrivono nello spazio tridimensionale oggetti che stanno in un piano. In un approfondimento vedremo qualche definizione equivalente di conica nel piano, nel prossimo paragrafo invece studiamo le coniche dal punto di vista della loro espressione algebrica come equazione di secondo grado.

9.2.1 Classificazione delle coniche in forma generale

In questo paragrafo ci occupiamo di un metodo per riconoscere, data l'equazione generale di una conica, che tipo di conica sia e se sia degenere oppure no. L'equazione generale di una conica è un'equazione di secondo grado nelle variabili x e y, a coefficienti reali:

$$a_{11}x^2 + 2a_{12}xy + a_{22}y^2 + 2a_{13}x + 2a_{23}y + a_{33} = 0. \tag{9.3}$$

Cominciamo con una considerazione di carattere generale. L'equazione (9.3) può essere riscritta con il linguaggio delle matrici e dei vettori. Infatti, posto

$$A_1 = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{pmatrix},$$

l'equazione (9.3) si riscrive come

$$(x, y, 1)A_1 \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = 0.$$

Inoltre anche la parte quadratica della conica può essere messa sotto forma di matrice: posto

$$A_2 = \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix},$$

si ha allora

$$a_{11}x^2 + 2a_{12}xy + a_{22}y^2 = (x, y)A_2 \begin{pmatrix} x \\ y \end{pmatrix}.$$

Appare dunque naturale che le due matrici simmetriche A_1, A_2 giochino un ruolo importante nella classificazione delle coniche.

Una prima importante distinzione è fra coniche degeneri oppure no. Ricordiamo che una iperbole può degenerare in una coppia di rette, una parabola in una retta, un ellisse in un punto. Come si vede questo dall'equazione (9.3)?

Cominciamo col vedere alcuni semplici esempi in cui la degenerazione delle coniche è evidente:

1. $x^2 + y^2 = 0$. L'ellisse (che poi è una circonferenza), degenera in un punto;
2. $(x + y)^2 = 0$. La parabola degenera in una retta;
3. $x^2 - y^2 = 0$. L'iperbole degenera in una coppia di rette.

La degenerazione, negli esempi sopra, appare ben chiara. Dal punto di vista algebrico, che cosa ci "segnala" la degenerazione? Negli esempi 2. e 3., è chiaro che quel che succede è che l'equazione della conica può essere scritta come prodotto di polinomi di primo grado (nelle variabili x, y).

Vediamo alcuni casi particolari in cui il polinomio (9.3) rappresenti una conica degenere.

1. Supponiamo compaia esplicitamente una variabile sola, ad esempio x. In tal caso, l'equazione (9.3) diventa allora della forma $a_{11}x^2 + 2a_{13}x + a_{33} = 0$. Se questo polinomio ammette due zeri reali, x_1, x_2, si può scomporre in $a_{11}(x - x_1)(x - x_2)$, con $x_1, x_2 \in \mathbb{R}$: si hanno allora due rette verticali se $x_1 \neq x_2$, una retta se $x_1 = x_2$;
2. Supponiamo ora che non ci siano termini diversi da quelli di secondo grado, cioè che l'equazione (9.3) assuma la forma $a_{11}x^2 + 2a_{12}xy + a_{22}y^2 = 0$. Se a_{11} oppure a_{22} sono nulli, il polinomio è scomponibile nel prodotto di x (oppure y) per un polinomio di primo grado e quindi la conica è ancora spezzata in due rette. Se invece $a_{11} \neq 0$ e $a_{22} \neq 0$, ponendo $\frac{y}{x} = t$, l'equazione diventa $a_{11} + 2a_{12}t + a_{22}t^2$. Ora questa equazione, se ammette radici reali, si scompone in $a_{22}(t - t_1)(t - t_2) = 0$ cioè la conica rappresenta le due rette $t = t_1$ e $t = t_2$ cioè $y = t_1x$ e $y = t_2x$, che naturalmente coincidono se $t_1 = t_2$.

Nel primo caso la (9.3) è stata scomposta in $a_{11}(x-x_1)(x-x_2)=0$, nel secondo caso, a seconda che $a_{11}=0$ oppure $a_{22}=0$ oppure a_{11} e a_{22} siano entrambi non nulli, il polinomio è stato comunuqe scomposto nel prodotto di due polinomi di primo grado. Dunque la degenerazione si ha, in questi casi, quando la (9.3) è riducibile nel prodotto di polinomi di primo grado. Supponiamo allora che l'equazione (9.3) possa essere riscritta nella forma:

$$(ax+by+c)(a'x+b'y+c')=0.$$

Se i coefficienti a,b,c,a',b',c' sono reali, il luogo dei punti del piano che la soddisfano è rappresentato allora dai punti che appartengono alle rette $ax+by+c=0$ e $a'x+b'y+c'=0$, che sono distinte a meno che i vettori (a,b,c) e (a',b',c') non siano proporzionali: in tal caso l'equazione può essere riscritta in forma equivalente come $(ax+by+c)^2=0$ che rappresenta il luogo dei punti appartenenti alla retta $ax+by+c=0$. In questi casi allora è chiaro che siamo di fronte ad una conica degenere.

Cerchiamo ora di individuare quali sono le condizioni alle quali devono soddisfare i coefficienti dell'equazione (9.3), perché essa rappresenti una conica degenere.
Osserviamo che, nelle considerazioni precedenti, non abbiamo incluso l'esempio di degenerazione dell'ellisse $(x^2+y^2=0)$. Questo caso rappresenta un'eccezione oppure può essere inserito nel discorso precedente? Ricordando che stiamo cercando di fattorizzare un polinomio di secondo grado, ci ricordiamo anche che queste cose si fanno, in maniera più naturale, nell'insieme dei *complessi* piuttosto che nell'insieme dei reali. E dunque, anche nel caso dell'equazione $x^2+y^2=0$ possiamo fattorizzare, scrivendo $(x+iy)(x-iy)=0$. Andando nei complessi, quel che si paga è che il risultato ottenuto, sui reali, non è evidentemente una coppia di rette.

Torniamo al problema della fattorizzazione, senza preoccuparci ora che quello che otteniamo siano numeri reali. Per cominciare, supponiamo che a_{11} e a_{22} non siano contemporaneamente nulli. Consideriamo il caso in cui $a_{11}\neq 0$, essendo l'altro del tutto analogo. In questo caso l'equazione (9.3) rappresenta una conica degenere in due rette se, pensando l'equazione di secondo grado in x o y, il suo discriminante è un quadrato: chiariamo con un esempio.

Esempio 9.2.1 Consideriamo l'equazione $2x^2+5xy+3y^2-2x-3y=0$ pensandola come equazione di secondo grado in x: $2x^2+x(5y-2)+3y^2-3y=0$. Calcoliamo il discriminante. $\Delta=(5y-2)^2-8(3y^2-3y)=y^2-4y+4$ ed essendo un quadrato, possiamo scrivere: $x=\frac{2-5y\pm\sqrt{(y-2)^2}}{4}$ che ha come soluzioni le due rette $2x=-3y+2$ e $x=-y$: l'equazione della conica rappresenta due rette.

Naturalmente, quel che può succedere in generale è che il discriminante sia anche il negativo di un quadrato: non importa, lavorando nei complessi questo non è un problema (anche se diventa meno immediata l'interpretazione geometrica).

Il discriminante dell'equazione (9.3), pensata nella variabile x, è

$$\Delta=(a_{12}y+a_{13})^2-a_{11}(a_{22}y^2+2a_{23}y+a_{33})=$$

$$= (a_{12}^2 - a_{11}a_{22})y^2 - 2y(-a_{12}a_{13} + a_{23}a_{11}) + a_{13}^2 - a_{11}a_{33}.$$

Affinché questo polinomio sia un quadrato deve essere nullo il suo discriminante cioè deve essere:

$$\Delta_1 = (-a_{12}a_{13} + a_{23}a_{11})^2 - (a_{12}^2 - a_{11}a_{22})(a_{13}^2 - a_{11}a_{33}) = 0,$$

e questa espressione può essere scritta sotto forma di determinante:

$$\Delta_1 = a_{11}\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{pmatrix} = 0.$$

Dato che abbiamo supposto $a_{11} \neq 0$, allora deve essere nullo il determinante della matrice

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{pmatrix}, \tag{9.4}$$

che è la matrice A_1 introdotta a pagina 132.

Abbiamo trovato una condizione molto sintetica ed efficace per esprimere il fatto che l'equazione di una conica si possa esprimere come prodotto di due polinomi di primo grado (a coefficienti complessi). Diremo allora che la condizione che il determinante della matrice (9.4) sia nullo esprime il fatto che la conica è *algebricamente degenere*.

Facciamo ancora le osservazioni seguenti:

Osservazione 9.2.1 Valgono i fatti seguenti:

1. Alla stessa condizione si arriva (ovviamente!) supponendo $a_{22} \neq 0$ e considerando l'equazione in y;
2. Nel caso, considerato sopra, in cui l'equazione contenga in maniera esplicita una sola variabile, il determinante della matrice (9.4) è ovviamente nullo, perché la matrice ha una riga di zeri;
3. La condizione affinché una conica sia degenere non cambia se $a_{11} = a_{22} = 0$.

Vediamo di dimostrare l'ultima affermazione, per esercizio.

Supponiamo allora di avere la conica nella forma

$$2a_{12}xy + 2a_{13}x + 2a_{23}y + a_{33} = 0. \tag{9.5}$$

La matrice (9.4) diventa allora:

$$\begin{pmatrix} 0 & a_{12} & a_{13} \\ a_{12} & 0 & a_{23} \\ a_{13} & a_{23} & a_{33} \end{pmatrix}.$$

Il suo determinante è:

$$a_{12}(2a_{13}a_{23} - a_{33}a_{12}),$$

e poiché supponiamo $a_{12} \neq 0$, annullando il determinante si ricava

$$a_{33} = \frac{2a_{13}a_{23}}{a_{12}}.$$

Sostituendo nella (9.5) si ottiene allora

$$a_{12}x(a_{12}y + a_{13}) = -a_{23}(a_{12}y + a_{13}),$$

che ha come soluzioni:

$$y = -\frac{a_{13}}{a_{12}}, \qquad x = -\frac{a_{23}}{a_{12}},$$

e cioè proprio due rette, come richiesto.

Vediamo ora il caso delle coniche non degeneri.

Supponiamo di aver verificato che la generica equazione (9.3) non rappresenti una conica degenere. Ora distinguiamo il caso in cui la conica abbia un centro di simmetria (ellisse, iperbole) oppure no (parabola). Se la conica ha un centro di simmetria, possiamo applicare ad essa una traslazione secondo il vettore (m, n) (con $m, n \in \mathbb{R}$) che trasforma il centro di simmetria della conica nell'origine $(0, 0)$.

L'equazione di questa traslazione è :

$$\begin{cases} x' = x - m \\ y' = y - n \end{cases},$$

da cui si ottiene

$$\begin{cases} x = x' + m \\ y = y' + n \end{cases}.$$

Applicando la traslazione alla conica:

$$a_{11}(x'+m)^2+2a_{12}(x'+m)(y'+n)+a_{22}(y'+n)^2+2a_{13}(x'+m)+2a_{23}(y'+n)+a_{33} = 0.$$

Svolgendo i calcoli e raccogliendo i termini di primo grado:

$$a_{11}x'^2 + 2a_{12}x'y' + a_{22}y'^2 + x'(2a_{11}m + 2a_{12}n + 2a_{13}) + y'(2a_{12}m + 2a_{22}n+$$
$$+2a_{23}) + a_{11}m^2 + a_{22}n^2 + 2a_{12}mn + 2a_{13}m + 2a_{23}n + a_{33} = 0.$$

Avendo traslato il centro della conica nell'origine, dobbiamo ottenere una curva simmetrica rispetto all'origine, quindi l'equazione che abbiamo ricavato non deve contenere i termini di primo grado (ricordiamo che l'equazione della simmetria rispetto all'origine è : $\begin{cases} x' = -x \\ y' = -y \end{cases}$). Dobbiamo perciò imporre che

$$\begin{cases} 2a_{11}m + 2a_{12}n + 2a_{13} = 0 \\ 2a_{12}m + 2a_{22}n + 2a_{23} = 0 \end{cases},$$

dove le incognite del sistema sono n, m. Questo sistema ammette soluzione unica se il determinante della matrice $A_2 = \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix}$ è diverso da zero, cioè

$$a_{11}a_{22} - a_{12}^2 \neq 0.$$

Abbiamo così ricavato la condizione affinché la conica rappresentata dall'equazione (9.3) abbia un centro di simmetria (sia cioè un'ellisse o un'iperbole). Le coordinate del centro, se esiste, si calcolano risolvendo il sistema.
La condizione che abbiamo ricavato: $a_{11}a_{22} - a_{12}^2 \neq 0$ ci dice che i termini di secondo grado $a_{11}x^2 + 2a_{12}xy + a_{22}y^2$ non rappresentano il quadrato di un binomio. Se così fosse, infatti, il discriminante dell'equazione $a_{11}x^2 + 2a_{12}xy + a_{22}y^2 = 0$ sarebbe nullo ($\Delta = a_{12}^2 - a_{11}a_{22} = 0$).

Concludendo: data la conica non degenere di equazione (9.3), essa rappresenta:

1. un'ellisse o un'iperbole se $\det A_2 \neq 0$;
2. una parabola se $\det A_2 = 0$.

Vediamo ora come caratterizzare l'ellisse e l'iperbole nel caso in cui sia verificata la condizione $\det A_2 \neq 0$.

Supponiamo di aver traslato il centro della conica nell'origine. La sua equazione è diventata:

$$a_{11}x^2 + 2a_{12}xy + a_{22}y^2 + k = 0,$$

con

$$k = a_{11}m^2 + 2a_{12}mn + a_{22}n^2 + 2a_{13}m + 2a_{13}m + 2a_{23}n + a_{33}.$$

Se la conica fosse un'iperbole (che ha equazione canonica $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$), i termini di secondo grado si potrebbero scomporre nel prodotto di due binomi: $(\frac{x}{a} - \frac{y}{b})(\frac{x}{a} + \frac{y}{b}) = 1$. Nel caso dell'equazione $a_{11}x^2 + 2a_{12}xy + a_{22}y^2 + k = 0$, il trinomio

$$a_{11}x^2 + 2a_{12}xy + a_{22}y^2$$

dovrebbe allora essere scomponibile nel prodotto di due binomi di primo grado: questa condizione è verificata se
$$a_{12}^2 - a_{11}a_{22} > 0.$$
Infatti, dall'equazione $a_{11}x^2 + 2a_{12}xy + a_{22}y^2 = 0$ si ottiene:
$$x = \frac{-a_{12}y \pm y\sqrt{a_{12}^2 - a_{11}a_{22}}}{a_{11}},$$
che rappresenta una coppia di rette distinte se e solo se $\Delta > 0$, cioè $\det A_2 > 0$.
Se $\Delta < 0$ ($\det A_2 < 0$), il trinomio di secondo grado non è scomponibile, e la conica è un'ellisse.

RIASSUMENDO:

L'equazione
$$a_{11}x^2 + 2a_{12}xy + a_{22}y^2 + 2a_{13}x + 2a_{23}y + a_{33} = 0$$

- rappresenta una conica degenere se $\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{pmatrix} = 0$;
- se $\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{pmatrix} \neq 0$ la conica è non degenere e
  - se $\det \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix} = 0$ è una parabola;
  - se $\det \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix} < 0$ è un'iperbole;
  - se se $\det \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix} > 0$ è un'ellisse.

Una volta riconosciuta la conica è possibile, con un'opportuna scelta del sistema di riferimento, portarla in forma canonica.

Esercizio 9.2.1 Studiamo la conica: $x^2 - 4y^2 - 2x + y - 1 = 0$.
Vediamo se è degenere: $\det \begin{pmatrix} 1 & 0 & -1 \\ 0 & -4 & \frac{1}{2} \\ -1 & \frac{1}{2} & -1 \end{pmatrix} = \frac{33}{4} \neq 0$.

La conica non è degenere e poiché $\det A_2 = \det \begin{pmatrix} 1 & 0 \\ 0 & -4 \end{pmatrix} = -4 < 0$ si tratta di un'iperbole. Se vogliamo determinare il centro di simmetria, dobbiamo risolvere il sistema: $\begin{cases} 2a_{11}m + 2a_{12}n + 2a_{13} = 0 \\ 2a_{12}m + 2a_{22}n + 2a_{23} = 0 \end{cases}$, cioè $\begin{cases} 2m - 2 = 0 \\ -8n + 1 = 0 \end{cases}$ che ha come soluzione $\begin{cases} m = 1 \\ n = \frac{1}{8}. \end{cases}$

Osservazione 9.2.2 Per finire, un'osservazione importante. Abbiamo visto che l'irriducibilità, nei complessi, dell'equazione che caratterizza la conica, è equivalente al fatto che la conica non sia degenere. Ma, lavorando nei complessi, non possiamo aspettarci che l'equazione definisca poi una conica nel mondo reale. Ad esempio, l'equazione $x^2 + y^2 + 1 = 0$, essendo irriducibile, definisce una conica non degenere in senso algebrico, poi con ovvi calcoli si vede che si tratta di una conica classificata come ellisse; d'altra parte nessun punto del piano soddisfa ovviamente l'equazione precedente, per cui siamo in presenza di un ellisse immaginaria.

Concludiamo questa parte con una bella figura di coniche nel mondo reale.

Coniche nel mondo

Figura 9.2

9.3 Approfondimenti 1: Definizione geometrica delle coniche nel piano

Vediamo ora in questo approfondimento di introdurre le definizioni di coniche come figure geometriche giacenti nel piano.

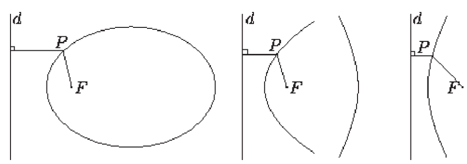
Definizione 9.3.1 *Una* conica *è il luogo dei punti del piano per i quali è costante il rapporto tra la distanza da un punto F, detto* fuoco *e da una retta r, detta* direttrice*. Questa costante è detta* eccentricità *della conica, e viene indicata con la lettera e.*

Sia P un punto generico della conica. Chiamato H il piede della perpendicolare condotta da P alla direttrice r, la precedente definizione diventa:

$$\frac{|PF|}{|PH|} = e.$$

Si verifica che:

- Se $e < 1$ si ottiene un'ellisse;
- Se $e = 1$ si ottiene parabola;
- Se $e > 1$ si ottiene un'iperbole.



L'eccencitricità delle coniche
Figura 9.3

La circonferenza è una particolare ellisse che ha il fuoco nel centro, la direttrice all'infinito ed eccentricità $e = 0$.

Si ha inoltre che:

- Ellissi ed iperboli hanno un centro di simmetria e due assi di simmetria ortogonali;
- La parabola non ha centro di simmetria ed ha un unico asse di simmetria.

Vediamo ora definizioni alternative di coniche:

Definizione 9.3.2 *L'ellisse è il luogo dei punti del piano che hanno distanze da due punti, detti* fuochi*, la cui somma costante.*

Se $F_1 = (-c, 0)$ e $F_2(c, 0)$ sono i due fuochi e P è un generico punto dell'ellisse, la precedente definizione si può scrivere:

$$|PF_1| + |PF_2| = 2c.$$

Si ha anche che

$$|F_1F_2| = 2c, \qquad a^2 = b^2 + c^2,$$

dove $(-a, 0)$ e $(a, 0)$ sono i punti intersezione dell'ellisse con l'asse x, e $(0, -b)$ e $(0, b)$ sono i punti intersezione dell'ellisse con l'asse y. a è la lunghezza del semiasse orizzontale dell'ellisse, b è la lunghezza del semiasse verticale dell'ellisse.

Esercizio 9.3.1 Provare che la retta tangente all'ellisse in un punto P è la bisettrice degli angoli formati dalle rette PF_1 e PF_2.

Definizione 9.3.3 *L'iperbole è il luogo dei punti del piano che hanno distanze da due punti, detti* fuochi*, la cui differenza costante.*

Se $F_1 = (-c, 0)$ e $F_2 = (c, 0)$ sono i due fuochi e P è un generico punto dell'iperbole, la precedente definizione si può scrivere:

$$||PF_1| - |PF_2|| = 2c.$$

Si ha anche che $y = \pm(b/a)x$ sono i due asintoti per l'iperbole, e che

$$F_1F_2 = 2c, \quad c^2 = b^2 + a^2,$$

dove $(-a, 0)$ e $(a, 0)$ sono i punti intersezione dell'iperbole con l'asse x, e a rappresenta la lunghezza del semiasse dell'iperbole.

Ogni conica è descritta da un'equazione di secondo grado in x, y. Con un'opportuna scelta degli assi la conica può essere ridotta *in forma canonica*:

- Ellisse:
$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1,$$
- Iperbole:
$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1,$$
- Parabola:
$$y = ax^2.$$

9.4 Approfondimenti 2: Quadriche in forma canonica

Le superfici nello spazio descritte da un'equazione di secondo grado in tre variabili si chiamano *quadriche*. L'esempio più semplice di superficie quadrica è quello della sfera che ha equazione $(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2 = r^2$, dove (x_0, y_0, z_0) è il centro ed r il raggio. Intersecando una sfera con un piano si trova una circonferenza che degenera in un punto (circonferenza di raggio nullo) se il piano risulta tangente alla sfera, e ci si può divertire a dire che è circonferenza immaginaria se il piano è esterno alla sfera. In generale, l'intersezione di una quadrica con un piano è una conica che diventa degenere nel caso in cui il piano sia tangente alla quadrica.

Vediamo ora una veloce carrellata delle quadriche:

- Ellissoide. Equazione
$$\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1:$$
è un solido di rotazione se due dei tre parametri a, b, c sono uguali tra loro. Intersecando un ellissoide con un piano si ottiene un'ellisse (eventualmente immaginaria). L'ellissoide ha un centro e tre piani di simmetria. Si dice che i punti dell'ellissoide sono punti *ellittici* per indicare il fatto che intersecando un ellissoide con un piano tangente si ottiene un'ellisse degenere;
- Iperboloide: si ottiene dalla rotazione di un'iperbole attorno ai propri assi di simmetria. Abbiamo due tipi di iperboloidi a seconda che la rotazione avvenga attorno all'asse reale (detto *trasverso*) o all'asse immaginario. Se si ruota l'iperbole attorno all'asse trasverso, otteniamo l'iperboloide ellittico (a due falde) la cui equazione è:
$$\frac{x^2}{a^2}-\frac{y^2}{b^2}-\frac{z^2}{c^2}=1$$
(se $b=c$ è di rotazione attorno all'asse x). Si chiama ellittico perché tutti i suoi punti sono ellittici. Se si ruota l'iperbole attorno all'asse immaginario otteniamo invece un'iperboloide iperbolico (ad una falda), la cui equazione è
$$\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}.$$
L'intersezione dell'iperboloide con un piano tangente è un'iperbole degenere (cioè due rette), da cui il nome "iperbolico";
- Paraboloide ellittico: si ottiene da una rotazione di una parabola attorno al proprio asse di simmetria. Ha equazione:
$$z=\frac{x^2}{a^2}+\frac{y^2}{b^2}$$
(se $a=b$ la rotazione è attorno all'asse z). L'intersezione di questo paraboloide con un piano tangente dà un'ellisse degenere il che giustifica il nome "ellittico";
- Paraboloide iperbolico (o "a sella"). Ha equazione
$$z=\frac{x^2}{a^2}-\frac{y^2}{b^2}.$$
L'intersezione del paraboloide con un piano tangente è un'iperbole che degenera in due rette, e i suoi punti sono iperbolici.

Esistono (purtroppo) anche le quadriche degeneri. Questo ne è elenco:

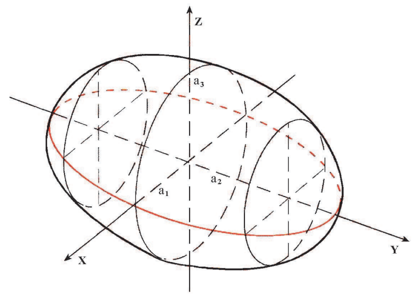
- Cono quadrico. Equazione
$$z^2=\frac{x^2}{a^2}+\frac{y^2}{b^2};$$
(se $a=b$ si ottiene il cono quadrico di rotazione attorno all'asse z);

- Cilindro quadrico. Si ottiene considerando una conica nel piano xy e tracciando delle parallele all'asse z per tutti i punti della conica. L'equazione del cilindro non contiene la variabile z, quindi sembra un'equazione in due variabili ma in realtà è un'equazione nello spazio. Si presentano i seguenti (ovvi) tre casi:
  1. Cilindro quadrico ellittico:
  $$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1;$$
  2. Cilindro quadrico iperbolico
  $$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1;$$
  3. Cilindro quadrico parabolico:
  $$y = ax^2.$$

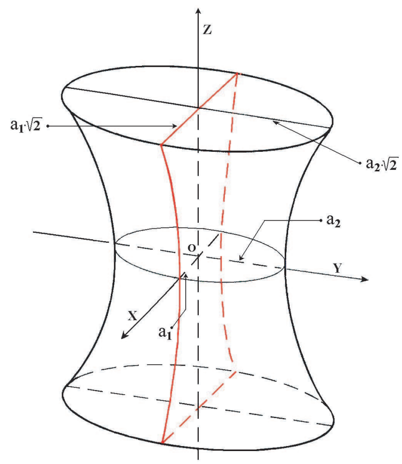
  .
- Quadriche prive di punti reali:
$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = -1.$$

Le intersezioni di un cono quadrico o di un cilindro quadrico con un piano tangente sono delle coniche che degenerano in una coppia di rette coincidenti.

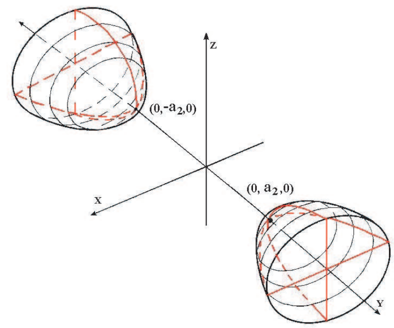
Tra le quadriche degeneri ci sono anche quelle originate da coppie di piani coincidenti, paralleli o incidenti che si spezzano nel prodotto di due fattori di primo grado (sono quadriche riducibili).



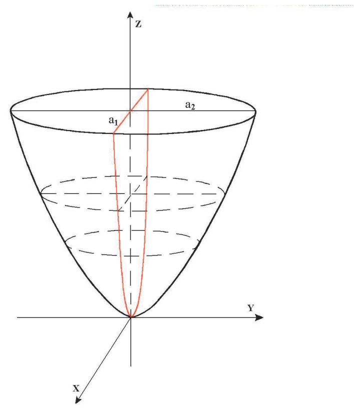
Ellissoide
Figura 9.4



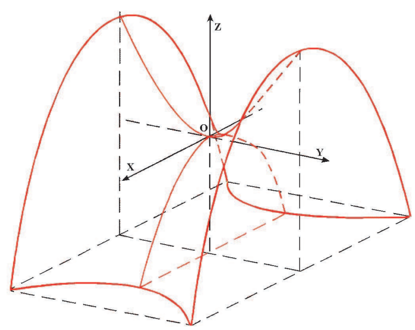
Iperboloide ad una falda
Figura 9.5



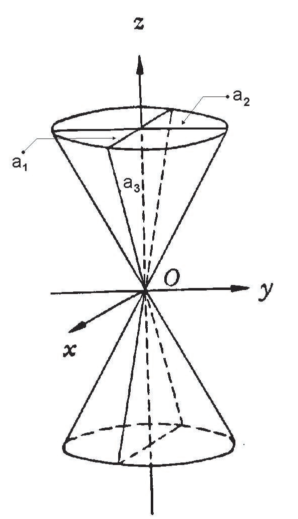
Iperboloide a due falde
Figura 9.6



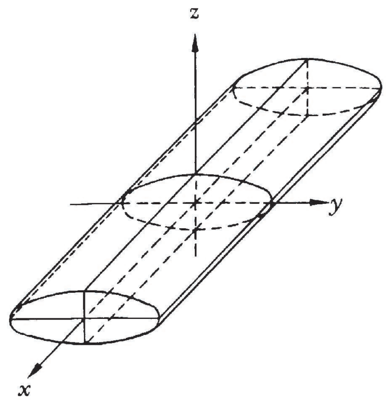
Paraboloide ellittico
Figura 9.7



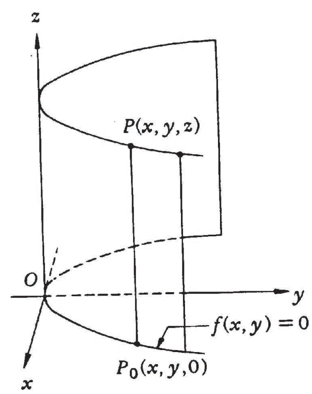
Paraboloide iperbolico (sella)
Figura 9.8



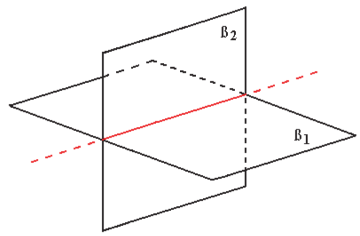
Cono
Figura 9.9



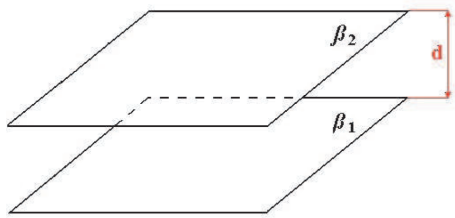
Cilindroide ellittico
Figura 9.10



Cilindroide parabolico
Figura 9.11



Piani incidenti
Figura 9.12



Piani paralleli
Figura 9.13

9.5 Esercizi

9.5.1 Quesiti a risposta chiusa

Rispondere vero o falso.

1. Il luogo dei punti del piano di equazione $x^2 + y^2 - 2x + 4y - 4 = 0$ è una circonferenza.
2. La conica di equazione $x^2 + 2y^2 + 3x - y = 0$ è un'ellisse.
3. Sia C la circonferenza di equazione $x^2 + y^2 + kx + y - 3 = 0$. Esistono due valori di $k \in \mathbb{R}$ per cui C è tangente alla retta $y = -x + 2$.
4. La conica di equazione $x^2 - y^2 - 3x - y = -2$
   a) è la coppia di rette $x + y = 1$,$x - y = 2$:
   b) è un'iperbole;
   c) è una parabola.
5. Sia C la circonferenza passante per tre punti $(-1, 0), (1, 0), (k, 1)$:
   a) C non passa per l'origine qualunque sia il valore di k;
   b) C è tangente alla retta $x + y = 1$ se e solo se $k = 1$;
   c) il centro di C appartiene all'asse y per ogni valore reale di k.

9.5.2 Esercizi aperti

Esercizio 9.5.1 Determinare l'equazione della circonferenza di centro $C(2, -1)$ e raggio $r = 3$.

Esercizio 9.5.2 Dire se le seguenti equazioni rappresentano una circonferenza ed in caso affermativo calcolare centro e raggio.

1. $x^2 + y^2 - 8x + 6y = 0$;
2. $x^2 + y^2 + 10x - 4y + 29 = 0$;
3. $x^2 + y^2 - 2x + y + 9 = 0$;
4. $2x^2 + 4y^2 - x + y = 0$;
5. $3x^2 + 3y^2 + 16y = 0$.

Esercizio 9.5.3 Determinare l'equazione della circonferenza che passa per $A(0, 3), B(-4, 1), C(1, 1)$.

Esercizio 9.5.4 Scrivere l'equazione delle rette passanti per $P(0, -4)$ e tangenti alla circonferenza $x^2 + y^2 = 4$.

Esercizio 9.5.5 Studiare le coniche:

1. $2x^2 + 4y^2 - 2x + 3y + 2 = 0$;
2. $-x^2 - 4y^2 - x + 3y - 2 = 0$;
3. $5x^2 + 6xy + 2y^2 - 2y + 5 = 0$;
4. $x^2 - 2xy + y^2 - x - 1 = 0$;
5. $x^2 - 4y^2 + 3xy - x + 2 = 0$;
6. $2y^2 - xy - 3x^2 = 0$.

10

Forme quadratiche

In questo capitolo vediamo brevemente come si studia una forma quadratica. Capire come si comportano le forme quadratiche è molto importante per lo studio delle funzioni di più variabili. Infatti, in prossimità di un punto critico, una funzione può essere approssimata dal suo polinomio di Taylor del secondo ordine, la cui parte quadratica è legata alla natura del punto critico della funzione. Vediamo allora di introdurre lo studio delle forme quadratiche.

Una *forma bilineare* $a(\cdot,\cdot)$ è una funzione *lineare* rispetto alle due variabili, cioè $a(x,\cdot)$ è funzione lineare per ogni x, $a(\cdot,y)$ è funzione lineare per ogni y[1]. Un primo esempio importante di forma bilineare è $a(x,y) = \overrightarrow{x}^t\overrightarrow{y} = \overrightarrow{x}\cdot\overrightarrow{y} = \sum x_i y_i$. Data una matrice $A = (a_{ij})$, una forma bilineare naturale è definita come $a(\overrightarrow{x},\overrightarrow{y}) = \sum a_{ij}x_i y_j = \overrightarrow{x}^t A\overrightarrow{y}$. Una forma bilineare si dice *simmetrica* se $a(\overrightarrow{x},\overrightarrow{y}) = a(\overrightarrow{y},\overrightarrow{x})$ per ogni $\overrightarrow{x},\overrightarrow{y}$. Nel caso della forma di sopra, generata da una matrice A, la forma è simmetrica se e solo se lo è la matrice A. Data una forma bilineare a, la funzione $\phi(\overrightarrow{x}) = a(\overrightarrow{x},\overrightarrow{x})$ è detta *forma quadratica*. Per $n = 1$, la forma quadratica diventa $\phi(x) = ax^2$, con $a > 0$, o $a = 0$, oppure $a < 0$. Per $n = 2$, data la matrice simmetrica

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

ed il vettore

$$\overrightarrow{x} = \begin{bmatrix} x \\ y \end{bmatrix},$$

la forma quadratica diventa

$$\phi(x,y) = [x\ y]A\begin{bmatrix} x \\ y \end{bmatrix} = ax^2 + 2bxy + cy^2,$$

il cui grafico rappresenta un paraboloide ellittico oppure iperbolico.

[1] La scrittura $a(x,\cdot)$ indica che consideriamo la funzione a come funzione della sola variabile y, ed x viene considerata un parametro

D'ora in avanti ci occupiamo di forme quadratiche generate da una matrice simmetrica A.

Osserviamo che una forma quadratica si annulla sempre nell'origine. Quindi quando parleremo di segno di una forma quadratica, escludiamo l'origine dalle nostre considerazioni. Consideriamo i seguenti esempi.

1. $$\phi(x, y) = x^2 + y^2.$$
   Si ha $\phi(x, y) > 0$ per ogni (x, y);
2. $$\phi(x, y) = -x^2 - y^2.$$
   Si ha $\phi(x, y) < 0$ per ogni (x, y);
3. $$\phi(x, y) = x^2 - y^2.$$
   Si ha $\phi(x, 0) > 0$ per ogni $x \neq 0$, mentre $\phi(0, y) < 0$ per ogni $y \neq 0$;
4. $$\phi(x, y) = x^2 - 2xy + y^2 = (x - y)^2.$$
   Si ha $\phi(x, y) > 0$ per ogni (x, y) tale che $y \neq x$, mentre $\phi(x, x) = 0$;
5. $$\phi(x, y) = -(x - y)^2.$$
   Si ha $\phi(x, y) < 0$ per ogni (x, y) tale che $y \neq x$, mentre $\phi(x, x) = 0$.

Gli esempi precedenti mostrano che, a parte l'origine ove si annulla necessariamente, una forma quadratica può essere sempre positiva, oppure sempre negativa, oppure può assumere segni differenti su rette differenti, oppure infine essere sempre non negativa (o non positiva) ed annullarsi lungo una retta. Dunque per studiare il segno della forma una buona idea è di vedere come si comporta sulle rette della forma $y = mx$. Si ha allora:
$$\phi(x, mx) = x^2(cm^2 + 2bm + a).$$
Dunque il segno della forma quadratica viene determinato dal fattore $(cm^2+2bm+a)$. Considerando la precedente espressione un polinomio di secondo grado nella variabile m, si ha che il suo discriminante (diviso per 4) è dato dall'espressione
$$b^2 - ac,$$
che rappresenta anche il determinante della matrice A, *con il segno cambiato*.

Riassumendo:

- Se $\det A > 0$, la forma quadratica *non cambia segno*, e viene detta *definita*;
- se $\det A = 0$ la forma è detta *semidefinita*, e si annulla lungo la retta con coefficiente angolare $-(b/c)$;
- se $\det A < 0$, la forma non ha segno costante e viene detta *indefinita*;

- Nel caso sia definita o semidefinita, per stabilire di che segno è basta vedere se $a > 0$ (nota $a = a_{11}$): in questo caso si parla di forma definita/semidefinita positiva, oppure $a < 0$: forma definita/semidefinita negativa.

C'è un modo semplice, almeno dal punto di vista teorico, per studiare il segno di una forma quadratica, nel caso la matrice simmetrica A sia di dimensione n qualunque. Consideriamo infatti un suo autovalore λ, ed $\overrightarrow{u}$ autovettore associato. Si ha allora che $A\overrightarrow{u} = \lambda\overrightarrow{u}$, e quindi

$$\phi(\overrightarrow{u}) = A\overrightarrow{u} \cdot \overrightarrow{u} = \lambda\|\overrightarrow{u}\|^2.$$

Ricordando che gli autovalori di A sono tutti reali, e considerando l'insieme di tutti gli autovalori, ne segue allora la seguente Proposizione:

Proposizione 10.0.1 *Sia ϕ la forma bilineare simmetrica associata ad una matrice A.*

- *se tutti gli autovalori di A sono positivi ϕ risulta definita positiva;*
- *se tutti gli autovalori di A sono negativi ϕ risulta definita negativa;*
- *se ci sono autovalori di segno discorde ϕ risulta indefinita;*
- *se almeno un autovalore di A è nullo e gli altri dello stesso segno ϕ è semidefinita.*

Il metodo descritto precedentemente e che riguarda l'andare a studiare il segno del determinante e dell'elemento a_{11} della matrice hessiana si estende nella maniera descritta dalla seguente proposizione. In essa si parla di sottomatrici A_k di nord-ovest e di sottomatrici principali. Le prime sono, come indica il nome, quelle matrici formate dalle prime k righe e k colonne, le seconde sono tutte le matrici simmetriche rispetto alla diagonale principale.

Proposizione 10.0.2 *La forma quadratica $\phi(x_1, \ldots, x_n)$, generata dalla matrice simmetrica A, è:*

- *definita positiva se e solo se* $\det A_k > 0$ *per* $k = 1, \ldots, n$*;*
- *definita negativa se e solo se* $(-1)^k \det A_k > 0$ *(cioè il determinante di quelli di ordine dispari deve essere negativo, quello di ordine pari deve essere positivo);*
- *semidefinita positiva se e solo se ogni sottomatrice principale ha determinante non negativo;*
- *semidefinita negativa se e solo se ogni sottomatrice principale ha determinante non negativo se è pari, non positivo se è dispari.*

Non dimostriamo la proposizione precedente; vale però la pena osservare che nel caso delle matrici semidefinite si deve verificare la condizione su una classe più ampia di sottomatrici, ma questo è naturale. Se si considera l'esempio delle due matrici seguenti:

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

si vede subito che hanno lo stesso comportamento sulle sottomatrici di Nord-Ovest, ma hanno natura molto diversa. Il fatto è che, essendo possibile che qualche elemento o determinante sia nullo, questo impedisce il controllo di altre matrici principali. Nel caso definito basta il controllo di quelle di Nord-Ovest per avere il controllo su tutte (una matrice 2×2 è definita positiva se il suo determinante è positivo e se a_{11} è positivo: questo automaticamente implica che anche a_{22} debba essere positivo).

Concludiamo osservando che le forme quadratiche in due dimensioni hanno avuto un ruolo importante nello studio delle coniche. Infatti, l'equazione di una conica si può scrivere, con le notazioni di questo paragrafo, come

$$\phi(\overrightarrow{u}) + \overrightarrow{c} \cdot \overrightarrow{u} = k, \tag{10.1}$$

con $\phi(\overrightarrow{u}) = \overrightarrow{u}^t A \overrightarrow{u}$, ed A matrice simmetrica. Abbiamo già visto che la natura della conica *non* dipende dal vettore $\overrightarrow{c}$, ma solo dalla parte quadratica ϕ: precisamente:

1. Se la forma quadratica è definita positiva o negativa, la (10.1) fornisce un ellisse;
2. Se la forma quadratica è indefinita, la (10.1) fornisce un iperbole;
3. Se la forma quadratica è semidefinita positiva o negativa, la (10.1) fornisce una parabola, o una coppia di rette, eventualmente coincidenti.

Osserviamo ancora che per certi valori di k l'equazione (10.1) potrebbe non essere soddisfatta da nessun vettore nel piano, nei casi 1. e 3., mentre non può succedere nel caso 2.

10.1 Esercizi

10.1.1 Quesiti a risposta chiusa

Rispondere vero o falso.

1. Data la forma quadratica $Q(x, y) = 2x^2 - xy + y^2$
   a) è semidefinita;
   b) è definita positiva.
2. Una forma quadratica Q su $\mathbb{R}^n$ è definita positiva:
   a) se e solo se $Q(x) > 0$ per ogni $x \in \mathbb{R}^n$;
   b) se e solo se $Q(x) \geq 0 \in \mathbb{R}^n$;
   c) se e solo se $Q(x) > 0$ per ogni $x \in \mathbb{R}^n, x \neq 0$.
3. La forma quadratica $Q(x, y) = x^2 + 4xy + y^2$ è indefinita;
4. La quadrica di equazione $-2x^2 + y^2 - z^2 + 1 = 0$
   a) è un'iperboloide a due falde;
   b) è un'iperboloide a una falda;
   c) è un paraboloide iperbolico.

10.1.2 Esercizi aperti

Esercizio 10.1.1 Determinare, al variare di k, il segno della forma quadratica $kx^2 + 4xy + (k-3)y^2$ con $k \in \mathbb{R}$.

Esercizio 10.1.2 Dimostrare la Proposizione 10.0.1

11

Approfondimenti ed esercizi di ricapitolazione

11.1 Strutture

Nei capitoli precedenti abbiamo studiato un certo numero di insiemi di oggetti: dai numeri complessi alle matrici, dai vettori alle trasformazioni lineari, tanto per fare alcuni esempi. In particolare, abbiamo descritto delle operazioni fra questi oggetti, e le loro proprietà fondamentali. Alcune di queste si ripetono (ad esempio la formula $a + b = b + a$ è vera sia che a e b siano vettori, oppure matrici, oppure ancora numeri complessi, anche se il simbolo "+" assume di volta in volta significati diversi, pur se collegati ad un'idea di somma). È una questione estremamente importante quella di sapere individuare classi di oggetti che, pur essendo molto diversi, hanno proprietà che si ripetono in maniera costante. Sebbene con una conoscenza non approfondita della matematica fare numerosi esempi significativi sia piuttosto difficile, può essere interessante, per approfondire le conoscenze teoriche, vedere alcune di queste strutture. È quanto facciamo in questo paragrafo, fornendo anche alcuni primi esempi semplici.
Una prima struttura importante è già stata vista: si tratta di quella di *spazio vettoriale*. L'idea è che, dato un insieme di enti, si possono definire su di esso certe operazioni, che hanno certe peculiarità, ed in tal caso il dato dell'insieme, munito delle operazioni definite, fornisce una struttura cui diamo un nome particolare.
Nel caso di spazio vettoriale, le operazioni sono una somma, operazione interna all'insieme (cioè si "sommano", secondo l'operazione definita, due elementi dell'insieme, ed il risultato è ancora un elemento dell'insieme) ed il prodotto con uno scalare: si tratta di moltiplicare un elemento dell'insieme per un numero reale, ed il risultato è ancora un elemento dell'insieme stesso. Queste due operazioni poi godono di certe proprietà; il tutto concorre a definire una struttura ben precisa, che in questo caso chiamiamo appunto spazio vettoriale.
Questa definizione ha senso ed è utile perché sono tanti gli esempi di insiemi che hanno questa struttura. Ad esempio $\mathbb{R}^n$ (con n qualunque), ma anche $C[a, b]$, l'insieme delle funzioni definite sull'intervallo $[a, b]$, a valori reali e continue, o l'insieme di tutti i polinomi nella variabile x.

L'utilità nell' individuare queste strutture sta nel fatto che, a partire da *poche* proprietà che definiscono la struttura, se ne possono ricavare *molte altre*, comuni a tutti gli insiemi che hanno quella struttura.
Ora vediamo qualche altro esempio di struttura interessante.

Definizione 11.1.1 *Un insieme non vuoto G è detto* gruppo *se in G è definita un'operazione, chiamata* prodotto*, indicata con $\spadesuit$ e che ha le seguenti proprietà:*

1. *$a, b \in G$ implica $a \spadesuit b \in G$ (proprietà di* chiusura*);*
2. *$(a \spadesuit b) \spadesuit c = a \spadesuit (b \spadesuit c)$ (proprietà* associativa*);*
3. *Esiste un elemento $e \in G$ tale che $a \spadesuit e = e \spadesuit a = a$ per ogni $a \in G$ (esistenza dell'elemento* neutro*);*
4. *Per ogni $a \in G$ esiste un elemento $a^{-1} \in G$ tale che $a \spadesuit a^{-1} = a^{-1} \spadesuit a = e$ (esistenza degli* inversi*).*

Il gruppo G si dice commutativo *se $a \spadesuit b = b \spadesuit a$ per ogni $a, b \in G$.*

Esempio 11.1.1 L'insieme $G = \{1, -1\}$, con l'operazione $\spadesuit$ definita come prodotto, è un gruppo commutativo (un po' banale). L'insieme $\mathbb{Z}$ degli interi è un gruppo commutativo con l'operazione $\spadesuit$ definita come la somma, cioè $a \spadesuit b = a + b$.

Esempio 11.1.2 L'insieme formato da tutte le matrici 2×2 con determinante diverso da zero, forma un gruppo se si considera come operazione $\spadesuit$ l'usuale moltiplicazione fra matrici. Osservare che G non è commutativo.

Esempio 11.1.3 L'insieme formato da tutte le matrici 2×2 di rotazione R_α (vedi pagina 117), con l'operazione usuale di prodotto fra matrici, forma un gruppo *commutativo*.

Poniamo un po' di notazioni. Dato un gruppo G ed un elemento $a \in G$, definiamo $a^0 = e$ (e, ricordate, è l'elemento neutro rispetto a $\spadesuit$), $a^1 = a$, $a^2 = a \spadesuit a$, $\ldots a^k = a \spadesuit a^{k-1}$, $a^{-2} = a^{-1} \spadesuit a^{-1}$, $\ldots a^{-k} = (a^{-1})^k$. Si verifica facilmente che valgono le usuali regole delle potenze, ad esempio $a^m \spadesuit a^n = a^{m+n}$.

Esempio 11.1.4 Ecco un gruppo un po' più complicato. G consiste di tutti i simboli della forma $a^0 = a^n = e$, e per $i, j \leq n$ poniamo $a^i \spadesuit a^j = a^{i+j}$ se $i + j \leq n$, $a^i \spadesuit a^j = a^{i+j-n}$ se $i + j > n$. G allora è un gruppo. Per "vederlo" geometricamente, si può immaginare sul piano la circonferenza unitaria, a rappresenta una rotazione di $\frac{2\pi}{n}$, a^2 un' altra rotazione dello stesso angolo, che si aggiunge alla prima, etc. È chiaro che dopo n rotazioni si ritorna alla situazione di partenza.

Osservazione 11.1.1 È chiaro che l'uso del simbolo $\spadesuit$ è poco usuale in un testo di matematica. I testi in genere usano il simbolo $\cdot$, che ricorda il prodotto. Il motivo per cui lo usiamo qui è di richiamare l'attenzione sul fatto che questa operazione, che abbiamo chiamato prodotto, non è necessariamente legata al prodotto, nel senso usuale del termine (prodotto di due numeri). Infatti abbiamo visto nell'Esempio 11.1.1 che $\spadesuit$ ha il significato di somma. Se consideriamo, sull'insieme $\mathbb{Z}$ degli interi, l'usuale operazione di prodotto, allora $\mathbb{Z}$ non avrebbe la struttura di gruppo. Osserviamo ancora esplicitamente che nel gruppo $\mathbb{Z}$ con l'operazione di somma, l'elemento a^2 indica $a + a$, cioè in notazione più convenzionale, $2a$.

Vediamo ora, solo come campione, alcune proprietà che ha ogni gruppo. Il fatto di avere una lista di tali proprietà fa sì che, ogni volta che riconosciamo che un insieme A su cui abbiamo definito una certa operazione, è un gruppo, queste proprietà allora valgono automaticamente, e non c'è bisogno di dimostrarle ogni volta.

Proposizione 11.1.1 *Se G è un gruppo:*

- *L'elemento e è unico;*
- *Ogni elemento ha un unico inverso;*
- *Per ogni $a \in G$, $(a^{-1})^{-1} = a$;*
- *Per ogni $a, b \in G$, $(a \spadesuit b)^{-1} = b^{-1} \spadesuit a^{-1}$;*
- *$a \spadesuit u = a \spadesuit v$ implica $u = v$.*

Osservare l'ordine del prodotto nella quarta proprietà della proposizione: in particolare non è detto che valga: $(a \spadesuit b)^{-1} = a^{-1} \spadesuit b^{-1}$, a meno che, naturalmente, G non sia commutativo. Osserviamo inoltre che l'ultima proprietà dà una cosiddetta *legge di cancellazione* (infatti possiamo cancellare a da entrambi i membri dell'equazione).

Ecco un'altra struttura importante.

Definizione 11.1.2 *Un insieme non vuoto A è detto* anello *se in A sono definite due operazioni, indicate con $\heartsuit$ e $\diamondsuit$ con le seguenti proprietà:*

1. *$a, b \in A$ implica $a \heartsuit b \in A$ (proprietà di* chiusura *rispetto a $\heartsuit$);*
2. *$a \heartsuit b = b \heartsuit a$ (*commutatività *rispetto a $\heartsuit$);*
3. *$(a \heartsuit b) \heartsuit c = a \heartsuit (b \heartsuit c)$ (proprietà* associativa *rispetto a $\heartsuit$);*
4. *Esiste un elemento $0 \in A$ tale che $a \heartsuit 0 = a$ per ogni $a \in A$ (esistenza dell'elemento* neutro *rispetto a $\heartsuit$);*
5. *Per ogni $a \in A$ esiste un elemento $-a \in A$ tale che $a \heartsuit -a = 0$ (esistenza degli* inversi *rispetto a $\heartsuit$);*
6. *$a, b \in A$ implica $a \diamondsuit b \in A$ (proprietà di* chiusura *rispetto a $\diamondsuit$);*
7. *$(a \diamondsuit b) \diamondsuit c = a \diamondsuit (b \diamondsuit c)$ (proprietà* associativa *rispetto a $\diamondsuit$);*
8. *$a \diamondsuit (b \heartsuit c) = a \diamondsuit b \heartsuit a \diamondsuit c$ e $(b \heartsuit c) \diamondsuit a = b \diamondsuit a \heartsuit c \diamondsuit a$.*

Osservazione 11.1.2 1. . . 5. stabiliscono che A è un gruppo commutativo rispetto all'operazione $\heartsuit$, mentre 6. e 7. stabiliscono che A è chiuso rispetto a $\diamondsuit$, che è un'operazione associativa. La proprietà 8 lega le due operazioni $\heartsuit$ e $\diamondsuit$. A questo proposito c'è da notare che, ogniqualvolta in una definizione parliamo di due operazioni allo stesso tempo, allora ci deve essere una proprietà che le lega. Altrimenti non ha senso raggruppare le due operazioni e le loro proprietà fondamentali in un'unica definizione. Osserviamo ancora che, per quanto riguarda l'operazione $\diamondsuit$, non vengono richieste l'esistenza di un elemento neutro, e di un inverso. Se poi si verifica la proprietà che

$a\diamondsuit b = b\diamondsuit a$ per ogni $a, b \in A$, l'anello si chiama *commutativo*. Ancora, l'elemento neutro rispetto a $\heartsuit$ si indica con 0 in quanto $\heartsuit$ ha molto spesso il significato di somma.

Esempio 11.1.5 $\mathbb{Z}$ è un anello commutativo, con $\heartsuit = +$, $\diamondsuit = \cdot$. Anche l'insieme di tutti gli interi pari è un anello commutativo, con le operazioni precedenti. Questo esempio è molto naturale, e giustifica l'uso del simbolo 0per indicare l'elemento neutro rispetto a $\heartsuit$.

Esempio 11.1.6 L'insieme delle matrici 2×2 con le usuali operazioni di somma e prodotto fra matrici è un anello *non* commutativo.

Esempio 11.1.7 Indichiamo con $\mathbb{Z}_p$ l'insieme del *simboli* $\{0, 1, 2, \dots, p\}$, munito delle operazioni seguenti: $\heartsuit$ viene indicato con $+$ ed agisce così: $a+b$ è il resto della divisione di $a + b$ per p (ad esempio, se $p = 12$, $8 + 7 = 3$ perché $8 + 7 = 15 = 12 \cdot 1 + 3$); $\diamondsuit$ viene indicato con $\cdot$ ed agisce così: $a \cdot b$ è il resto della divisione di ab per p (ad esempio, se $p = 12$, $8 \cdot 7 = 8$ perché $8 \cdot 7 = 56 = 12 \cdot 4 + 8$). È facile verificare che $\mathbb{Z}_p$ è un anello commutativo, e forse la scelta di $p = 12$ dovrebbe dare un minimo di significato all'introduzione di $\mathbb{Z}_p$, visto che se adesso il nostro orologio segna le 6 tra 15 ore ci aspettiamo che segni le 9 ...

Osserviamo esplicitamente che in un anello la condizione $a\diamondsuit b = 0$ *non* implica che $a = 0$ oppure $b = 0$ (ricordiamo ancora che 0 è l'elemento neutro rispetto all'operazione $\heartsuit$, ad esempio nell'anello delle matrici la matrice nulla). Inoltre *non* vale la legge di cancellazione: se $a\diamondsuit c = b\diamondsuit c$ non è detto che $a = b$.

Definizione 11.1.3 *Un anello F con le operazioni $\diamondsuit$ e $\heartsuit$ definite come sopra e tale che è commutativo rispetto a $\diamondsuit$ e tale che l'insieme dei suoi elementi non nulli forma un gruppo rispetto a $\diamondsuit$, si dice* campo.

Esempio 11.1.8 Gli insiemi $\mathbb{Q}$ dei numeri razionali, $\mathbb{R}$ e $\mathbb{C}$, con le usuali operazioni, sono campi. Osservare che in $\mathbb{R}$ è definita anche una relazione $<$ con le seguenti proprietà:

1. per ogni $x, y \in \mathbb{R}$ una ed una sola delle seguenti tre relazioni è vera: o $x < y$, o $x = y$ o $y < x$;
2. se $x, y, z \in \mathbb{R}$ se $x < y$ e $y < z$, allora $x < z$.

Un *campo ordinato* è un insieme munito delle operazioni di campo e di un ordine $<$, che abbia le proprietà:

1. $x + y < x + z$ implica $y < z$;
2. $x > 0, y > 0$ implica $x \cdot y > 0$.

Osserviamo esplicitamente che le 1. e 2. sono le condizioni che legano l'operazione d'ordine con lwe due operazioni di campo.

Osserviamo inoltre che anche $\mathbb{Q}$ è un campo ordinato. La differenza fondamentale fra $\mathbb{R}$ e $\mathbb{Q}$ non è di tipo algebrico, cioè non riguarda questioni di struttura come quelle che stiamo discutendo ora, ma sta in una proprietà di completezza, di cui si parla nei corsi di Analisi, perché permette di sviluppare il calcolo infinitesimale in $\mathbb{R}$, cosa impossibile in $\mathbb{Q}$.

Esempio 11.1.9 Tornando all'Esempio 11.1.7, si può dimostrare che $\mathbb{Z}_p$ è un campo se e solo se p è primo. È facile vedere che $\mathbb{Z}_{12}$ non è un campo: in esso $3 \cdot 4 = 0$, cioè il prodotto di due numeri non nulli dà 0, il che non è possibile in un campo.

Si può dimostrare che, in ogni campo ordinato $x^2 = x \diamondsuit x > 0$ (è chiaro che cosa il simbolo 0 rappresenta qui?). Questo implica immediatamente che $\mathbb{C}$ non può essere un campo ordinato (per nessuna operazione $<$), perché $i^2 = -1$. Questo naturalmente *non* significa che non si possa dare a $\mathbb{C}$ una struttura d'ordine. Significa solo che, qualunque $<$ definiamo su $\mathbb{C}$, $\mathbb{C}$ non sarà un campo ordinato. Sarà un campo, ed anche un insieme ordinato. Ma non un campo ordinato.

Gli assiomi di campo implicano ad esempio le proprietà seguenti:

1. $x \neq 0, y \neq 0$ implica $x \diamondsuit y \neq 0$ (questo dice che l'anello delle matrici non può essere un campo);
2. $(-x) \diamondsuit y = -(x \diamondsuit y)$;
3. $(-x) \diamondsuit (-y) = x \diamondsuit y$;
4. $1 > 0$.

Da queste proprietà non è difficile ricavare, ad esempio, che in un campo ordinato, se $x > 0$ allora $-x < 0$, e $x^2 = x \diamondsuit x > 0$ per ogni $x \neq 0$. Osserviamo anche come va correttamente letta la proprietà 2. di sopra: se faccio l'operazione $\diamondsuit$ fra l'inverso di x (inverso rispetto all'operazione $\heartsuit$) e y ottengo l'inverso (rispetto all'operazione $\heartsuit$) dell'elemento $x \diamondsuit y$. È chiaro che questo sembra uno scioglilingua, è chiaro che tutto questo ci sembra evidente se al posto di $\heartsuit$ e $\diamondsuit$ mettiamo $+$ e $\cdot$. Il punto essenziale è capire che queste proprietà, che ci sembrano così naturali per la somma ed il prodotto in $\mathbb{R}$, sono operazioni che si possono applicare, con gli stessi risultati, anche a strutture più complesse, perché non dipendono dalla natura specifica dell'insieme $\mathbb{R}$, ma dalle *poche* proprietà, che abbiamo elencato nelle varie definizioni. Ancora una parola sulla relazione scritta sopra, $1 > 0$. Chi è 1? Dobbiamo specificarlo, ma dovrebbe essere chiaro: 1 rappresenta l'elemento neutro rispetto all'operazione $\diamondsuit$. Di solito $\diamondsuit$ indica un prodotto, e quindi è naturale indicare con 1 il suo elemento neutro. E poi, la relazione $1 > 0$ vuol dire che *in ogni campo ordinato* l'elemento neutro rispetto all'operazione $\diamondsuit$ è maggiore dell'elemento neutro rispetto all'operazione $\heartsuit$.

11.2 Esercizi di ricapitolazione

In questo paragrafo finale sono proposti vari esercizi di ripasso, che non sono né in ordine di argomento né in ordine di difficoltà. Ringraziamo i Prof. Grasselli e Verri di aver permesso di riportare alcune loro prove proposte a Ingegneria Fisica e Matematica negli anni passati.

11.2.1 Quesiti a risposta chiusa

Rispondere vero o falso.

1. Re $[z(u+w)] =$ Re zRe $(u+w) -$Im zIm $(u+w)$;
2. zu è reale se e solo se u è il coniugato di z;
3. Arg $z(u+w) =$ Arg $z+$Arg u solo se $w \in \mathbb{R}$;
4. Se $z^2 + w^2 = 0$, allora $z = w = 0$.
5. Dato il numero complesso $z = a + ib$, tale che $0 <$ Arg $z < \frac{\pi}{2}$:
   a) iz sta nel secondo quadrante nel piano di Argand-Gauss;
   b) Arg $z = \arctan \frac{b}{a}$;
   c) Esiste $n \neq 0$ naturale per cui z^n è un numero reale.
6. L'equazione $z^3 = 4|z|$: ha solo radici reali.
7. Siano $\overrightarrow{u} = 3\overrightarrow{i} + a\overrightarrow{j}$, $\overrightarrow{v} = b\overrightarrow{i} - \overrightarrow{j}$.
   a) Sono perpendicolari se $a = 3b$;
   b) L'insieme delle coppie (a, b) del piano per cui $|\overrightarrow{u}| = |\overrightarrow{v}|$ è una circonferenza.
8. Dati i vettori $\overrightarrow{v_1} = a\overrightarrow{i} + 3\overrightarrow{j}$ e $\overrightarrow{v_2} = -4\overrightarrow{i} + 2\overrightarrow{j}$
   esiste un valore di a per cui formano un angolo di 30°
9. Siano r la retta di equazione $2x - 3y + 4 = 0$ ed s la retta $\begin{cases} x = 2 - 3t \\ y = 1 - 2t \end{cases}$
   a) r e s non hanno punti in comune;
   b) un vettore direzione di s è $(3, 2)$;
   c) tutti i vettori perpendicolari ad r sono della forma $(a, -\frac{3a}{2})$, con $a \in \mathbb{R}$;
10. Dati i punti $A(-1, -2, 3)$, $B(4, 5, 1)$, $C(5, 7, -2)$:
    a) A appartiene al piano $3x - 7y + 2z - 1 = 0$;
    b) esiste un unico piano che li contiene;
    c) $\overrightarrow{BC}$ è perpendicolare al piano $3x + 6y - 9z - 1 = 0$.
11. Data la retta, in forma parametrica:
    $$\begin{cases} x = 1 + 2t \\ y = t \\ z = -bt \end{cases} \qquad :$$
    a) giace nel piano $x + y + z = 1$ per $b = 3$;
    b) è perpendicolare al piano $ax + y + z = 1$ se e solo se $a = 2$, $b = -1$;

c) esiste valore di b per cui è sghemba rispetto alla retta:

$$r : \begin{cases} x + y + z = 1 \\ 2x - y + 3z = 2 \end{cases} .$$

12. I vettori $(k, 1, 0)$,$(4, k, 1)$,$(-2, 1, 1)$ sono linearmente indipendenti per $k \neq -2$ e $k \neq 3$.

13. Sia $E = \{(x, y, z, w) \in \mathbb{R}^4 : x - y + z + 2w = 0\}$
    a) E è un sottospazio vettoriale di $\mathbb{R}^4$;
    b) I vettori $(0, 2, 1, 1)$,$(1, 1, 0, 1)$,$(0, 0, 1, 0)$,$(1, 1, 1, 1)$ sono una base per E;
    c) I vettori $(1, 1, -2, 1)$,$(0, 1, 1, 0)$,$(1, 2, -1, 1)$ sono una base per E.
14. Date la matrice $A = \begin{pmatrix} 2 & -3 & 1 \\ 4 & 5 & 2 \\ -1 & 3 & -1 \end{pmatrix}$ ed una qualunque matrice B di dimensione 3×3, esiste una matrice C di dimensione 3×3 tale che $AC = B$.
15. L'inversa della matrice $A = \begin{pmatrix} 2 & 1 & 2 \\ 2 & 1 & 0 \\ 0 & 1 & -2 \end{pmatrix}$ è la matrice $B = \begin{pmatrix} -\frac{1}{2} & 1 & -\frac{1}{2} \\ 1 & -1 & 1 \\ \frac{1}{2} & -\frac{1}{2} & 0 \end{pmatrix}$.
16. Siano A, B due matrici quadrate e 0 la matrice nulla della loro dimensione. Se B è invertibile e $AB = 0$ allora $A = 0$.
17. Dato il sistema
$$\begin{cases} (k+1)x + 2y - z = 0 \\ x + ky - kz = 2 \\ x + z = 1 \end{cases} :$$
    a) ha una ed una sola soluzione per $k \neq \pm\sqrt{2}$;
    b) ha soluzioni per ogni $k \in \mathbb{R}$.
18. Il sistema
$$\begin{cases} 2x - y + 5z = 7 \\ x + 3y + 2z = 4 \\ x + 10y + z = h \end{cases} ;$$
per $h = 5$ ha come soluzione una retta.
19. Le soluzioni del sistema omogeneo
$$\begin{cases} 2x - y + 5z = 0 \\ x + 3y + 2z = 0 \\ x + 10y + z = 0 \end{cases}$$

formano un sottospazio di R^3 di dimensione 2.

20. La matrice
$$\begin{pmatrix} 1 & 0 & k \\ k & -1 & 0 \\ 0 & 1 & k \end{pmatrix}:$$
    a) Ha rango 3 per $k \neq 0$ e $k \neq 1$;
    b) per $k = 0$ ha rango 1.
21. Il vettore $\overrightarrow{v} = (1, 3, 2)$ nella base canonica è il vettore $\overrightarrow{v}_B = (\frac{13}{2}, 6, \frac{3}{2})$ nella base $B = \{ \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} \}$.
22. Dati i vettori $\overrightarrow{u} = \begin{pmatrix} 1 \\ -2 \\ 2 \end{pmatrix}$ $\overrightarrow{v} = \begin{pmatrix} 2 \\ 0 \\ -1 \end{pmatrix}$ $\overrightarrow{w} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$:
    a) sono complanari;
    b) $\overrightarrow{t} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \in L < \overrightarrow{u}, \overrightarrow{v}, \overrightarrow{w} >$ (lo spazio generato dai tre vettori).
23. La matrice $\begin{pmatrix} k & k \\ 0 & k \end{pmatrix}$ è diagonalizzabile per $k \neq 0$.
24. Gli autovalori di una matrice quadrata triangolare superiore sono gli elementi della diagonale principale.
25. Data la matrice
$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 2 \end{pmatrix}:$$
    a) $\lambda = 1$ è un autovalore;
    b) è diagonalizzabile;
    c) una base di autovettori è $B = \{(-1, 01)\ (-1, 1, 0)\ (0, 0, 1)\}$.
26. Data la trasformazione $T : R^3 \to R^3$ $T(x, y, z) = (3x+y+z, 2x+4y+2z, x+y+3z)$;
    a) l'immagine di $P(-1, 3, 2)$ è $P'(2, 18, 8)$;
    b) la controimmagine di $(4, 6, 2)$ è $(1, 1, 0)$;
    c) è lineare;
27. La matrice $A = \begin{pmatrix} 3 & 1 & 1 \\ 2 & 4 & 2 \\ 1 & 1 & 3 \end{pmatrix}$ associata alla trasformazione T rispetto alla base canonica, ha tre autovalori distinti;
28. i vettori $(3, 2, 1)$; $(1, 4, 1)$; $(1, 2, 3)$ formano una base di R^3.
29. Dato il sistema:
$$\begin{cases} x - y + z = 0 \\ ax - 2y + 2z = 0 \\ (1-a)x + y - z = 0 \end{cases} :$$

a) esiste un valore di $a \in \mathbb{R}$ per cui ammette solo la soluzione nulla;
b) Per $a \neq 2$ ammette ∞^1 soluzioni;
c) Per $a = 2$ il vettore $(-2, 2, -2)$ è ortogonale a tutte le soluzioni;
d) per $a \neq 2$ i vettori $(0, 1, 1)$ e $(0, -4, -4)$ generano lo spazio delle soluzioni.

30. Dato il sistema (stessa matrice dei coefficienti del precedente):

$$\begin{cases} x - y + z = 1 \\ ax - 2y + 2z = 2 \\ (1-a)x + y - z = -1 \end{cases} \quad :$$

a) ha soluzione qualunque $a \in \mathbb{R}$;
b) Per ogni a l'insieme delle soluzioni rappresenta un piano nello spazio tridimensionale;
c) per $a = 2$ ogni vettore ortogonale allo spazio delle soluzioni è della forma $(b, -b, b)$ con $b \in \mathbb{R}$.

31. Data una matrice A 3×3, l'applicazione definita come $T(\overrightarrow{x}) = A^2 \overrightarrow{x}$:
a) è lineare qualunque sia A;
b) L'equazione $T(\overrightarrow{x}) = \overrightarrow{0}$ ha unica soluzione qualunque sia A.

11.2.2 Esercizi aperti

Esercizio 11.2.1 Data una matrice B $n \times n$, ed una matrice quadrata A $n \times n$, invertibile, mostrare che il rango di AB è uguale al rango di B.

Esercizio 11.2.2 Mostrare che, se A e B sono matrici simili, allora hanno lo stesso rango e gli stessi autovalori. Mostrare che A e A^t hanno gli stessi autovalori.

Esercizio 11.2.3 Mostrare che se $B = P^{-1}AP$ e λ è autovalore per A con autovettore $\overrightarrow{x}$, allora $P^{-1}\overrightarrow{x}$ è autovettore per B relativo all'autovalore λ. Mostrare che A è diagonalizzabile se e solo se B è diagonalizzabile.

Esercizio 11.2.4 Mostrare che se A è diagonalizzabile, il rango di A è il numero di autovalori non nulli di A.

Esercizio 11.2.5 Dire se le seguenti matrici sono diagonalizzabili ed in tal caso trovare una matrice P che permette la diagonalizzazione: $\begin{pmatrix} 0 & 0 & -2 \\ 1 & 2 & 1 \\ 1 & 0 & 3 \end{pmatrix}$, $\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 4 & -17 & 8 \end{pmatrix}$, $\begin{pmatrix} 2 & 1 & 1 \\ 2 & 1 & -2 \\ -1 & 0 & -2 \end{pmatrix}$,

Esercizio 11.2.6 La conica $a_{11}x^2 + 2a_{12}xy + a_{22}y^2 + 2a_{13}x + 2a_{23}y + a_{33} = \langle A\overrightarrow{x}, \overrightarrow{x}\rangle + \langle \overrightarrow{c}, \overrightarrow{x}\rangle = 0$ è una parabola. Supponiamo che l'equazione $\langle A\overrightarrow{x}, \overrightarrow{x}\rangle + \langle \overrightarrow{c}, \overrightarrow{x}\rangle = k$ *non* abbia soluzioni per ogni k. Mostrare che in questo caso la conica è degenere.

Suggerimento. Supponiamo la matrice A sia semidefinita positiva, e $a_{11} > 0$. Mostrare che un autovettore relativo all'autovalore nullo è della forma $(-a_{12}, a_{11})$. Allora $\overrightarrow{c} = (a_{13}, a_{23})$ è proporzionale al vettore (a_{11}, a_{12}). Scrivere la matrice 3×3 relativa alla conica e mostrare che il suo determinante è zero (per esempio sviluppandolo rispetto alla terza colonna).

Esercizio 11.2.7 Nello spazio $\mathbb{R}^3$, si considerino i tre piani di equazioni $kx - 2y + kz = 0$, $(k+2)x - 3y - 2kz = 0$, $2x - y - 3z = k - 1$, con $k \in \mathbb{R}$.

1. Determinare per quali valori di k i tre piani si intersecano lungo una retta r.
2. Stabilire se r è un sottospazio vettoriale di $\mathbb{R}^3$.
3. Determinare il sottospazio vettoriale di $\mathbb{R}^3$ ortogonale ad r.

Esercizio 11.2.8 Sia $A = \begin{pmatrix} t & t & 2 \\ 0 & t & 1 \\ 0 & 0 & 1 \end{pmatrix}$

1. Determinare i valori di t tali che A è diagonalizzabile.
2. Nel caso $t = 0$, stabilire se esiste una matrice X non singolare tale che $AX = XD$ con D matrice diagonale. In caso affermativo, determinare una possibile matrice X.
3. Stabilire se esistono valori di t per cui $v = [0, 1, -1]^t$ è autovettore della matrice A e, in caso affermativo, determinarli.

Esercizio 11.2.9 Siano $A = \{z \in \mathbb{C} : 0 < Rez < \frac{\pi}{2}, -\ln 2 < Imz \leq 0\}$,$B = \{w \in \mathbb{C} : w = e^{iz}, z \in A\}$.

1. Disegnare nel piano complesso le immagini dei punti appartenenti,rispettivamente, agli insiemi A e B.
2. Stabilire se esistono soluzioni $z \in \mathbb{C}$ della disequazione
$$z^2 + z\overline{z} - z - 1 > 0$$
appartenenti all'insieme B.

Esercizio 11.2.10 Nello spazio si consideri la retta r di equazioni:

$$\begin{cases} y = 3(x-1) \\ z = 2(x-1) \end{cases} .$$

Scrivere l'equazione della retta perpendicolare e incidente r e l'asse z.

Esercizio 11.2.11 Discutere la risolubilità del sistema di equazioni lineari (a parametro reale)

$$\begin{cases} (a-2)x + 2z = 0 \\ 4x + (2-a)y - 4z = 1 \\ 4x - (a+4)z = 0 \end{cases} .$$

Esercizio 11.2.12 Risolvere l'equazione nell'incognita complessa z

$$z^2 + 4\overline{z} + 2 = 0$$

e rappresentare le soluzioni nel piano di Gauss.

Esercizio 11.2.13 Determinare tutte le matrici di ordine 2×2 che commutano con la propria trasposta.

Esercizio 11.2.14 Stabilire la dimensione del sottospazio generato dai vettori
$x_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$; $x_2 = \begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix}$; $x_3 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$; $x_4 = \begin{pmatrix} -1 \\ 1 \\ 2 \end{pmatrix}$.
Quindi determinarne una base.

Esercizio 11.2.15 Sia

$$z = \frac{i^7(1+i)^4}{(1-i)^5}.$$

1. Calcolare $|z|$, Re z, Im z, Arg z.
2. Calcolare e rappresentare nel piano di Gauss $\sqrt[5]{z}$.

Esercizio 11.2.16 Determinare le condizioni sul numero reale a in modo che i tre vettori seguenti siano linearmente indipendenti:
$x_1 = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}$; $x_2 = \begin{pmatrix} 1 \\ 3 \\ a \end{pmatrix}$; $x_3 = \begin{pmatrix} 0 \\ a \\ 2 \end{pmatrix}$.
In tal caso indicare una base del sottospazio generato dai tre vettori dati.

Esercizio 11.2.17 Si consideri l'equazione nell'incognita complessa z

$$z^4 + 2z^2 + k = 0$$

1. Per quali valori del parametro reale k tutte le radici delle'equazione sono *non reali?*
2. Determinare le radici per $k = 4$ e rappresentarle nel piano di Gauss.

Esercizio 11.2.18 Scrivere l'equazione della retta passante per $P(0, 0, 1)$ e parallela alla retta s di equazione:

$$\begin{cases} x = 1 + u \\ y = 1 - u \\ z = \phantom{1+{}} 2u \end{cases}.$$

Scrivere l'equazione della retta t passante per P e perpendicolare ed incidente ad s. Scrivere l'equazione del piano π che contiene r e t. π contiene anche s?

Esercizio 11.2.19 Siano

$$A = \begin{pmatrix} 1 & 1 & h \\ 0 & h & 6 \\ 1 & 0 & 1 \end{pmatrix}, \qquad \overrightarrow{b}^T = (2, 3, 1).$$

Discutere, al variare del parametro h, l'esistenza di soluzioni del sistema $A\overrightarrow{x} = \overrightarrow{b}$, e determinarle, quando possibile.

Esercizio 11.2.20 Trovare nel piano complesso le soluzioni di

$$|z + 2i| = |z - 2|,$$

e disegnare il luogo dei punti soluzione.

Esercizio 11.2.21 Trovare, per ogni matrice seguente, un polinomio di cui la matrice stessa sia radice:

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 2 & 3 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix}, \quad D = \begin{pmatrix} 0 & 1 & 2 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}$$

Indice analitico

Indice