

Appunti di Algebra Lineare

Antonino Salibra

January 25, 2017

Programma di Algebra Lineare (2016/17):

1. Libri di Testo:
 - A. Salibra: Appunti di Algebra Lineare, 2016.
 - M. Abate, C. de Fabritiis: Geometria analitica con elementi di algebra lineare, Seconda Edizione, McGraw-Hill, 2010.
2. Campi numerici. Esempi di campi numerici: il campo dei numeri reali; il campo dei numeri complessi; il campo dei numeri modulo 2; il campo dei numeri modulo p per un numero primo p .
3. Numeri complessi: parte reale ed immaginaria, coniugato di un numero complesso. Modulo e norma. Prodotto e somma di numeri complessi. Rappresentazione trigonometrica: piano complesso. Rappresentazione esponenziale.
4. Introduzione ai vettori. Grandezze fisiche e vettori. La definizione di spazio vettoriale con somma vettoriale e prodotto per uno scalare.
5. Prodotto interno (o scalare) di due vettori del piano o dello spazio. Proprietà del prodotto interno. Lunghezza di un vettore. Caratterizzazione della perpendicolarità con il prodotto interno. Disuguaglianza di Schwartz. Caratterizzazione del coseno dell'angolo formato da due vettori con il prodotto interno. Alcune applicazioni alla geometria euclidea.
6. Rette nel piano. Equazione lineare di una retta. Equazione parametrica di una retta. Retta passante per l'origine e perpendicolare ad un dato vettore. Retta parallela ad una retta passante per l'origine. Retta passante per un punto e parallela (alla retta determinata da) ad un dato vettore. Retta passante per due punti.
7. Rette e piani nello spazio. Equazione lineare di un piano. Equazione parametrica di un piano. Piano passante per l'origine e perpendicolare ad un dato vettore. Piano parallelo ad un piano passante per l'origine. Piano passante per tre punti non allineati.
8. Sistema Lineari. Il metodo di eliminazione di Gauss. Criteri da applicare nel metodo di eliminazione di Gauss.

9. Matrici. Matrice quadrata, simmetrica, diagonale, triangolare superiore e inferiore. Trasposta di una matrice. Lo spazio vettoriale delle matrici di tipo $m \times n$. Prodotto di matrici. Proprietà del prodotto di matrici. Moltiplicazione di matrici a blocchi.
10. Matrici e sistemi lineari. Matrici elementari, operazioni elementari e metodo di eliminazione di Gauss. Matrice a gradini e matrice ridotta. Riduzione di una matrice in forma a gradini e in forma ridotta. Applicazioni ai sistemi lineari.
11. Spazi vettoriali. Sottospazi. Vettori linearmente dipendenti e linearmente indipendenti. Basi e dimensione di uno spazio vettoriale.
12. Trasformazioni lineari. Teorema della dimensione. Matrice di una trasformazione lineare. Trasformazione lineare definita da una matrice. Sottospazio vettoriale delle colonne (righe) di una matrice. Isomorfismi e cambi di base. Sistemi lineari e trasformazioni lineari.
13. Interpretazione geometrica del determinante come area di un parallelogramma e come volume di un parallelepipedo. Prodotto vettoriale di due vettori dello spazio. Determinante di una matrice quadrata di ordine n . Determinante della matrice trasposta, del prodotto di due matrici. Regola di Cramer. Calcolo della matrice inversa con la regola di Cramer, con il metodo dei cofattori, e con il metodo di Gauss-Jordan.
14. Autovalori e autovettori. Autospazi. Polinomio caratteristico. Radici del polinomio caratteristico. Matrici simili hanno lo stesso polinomio caratteristico. Autovettori di autovalori distinti. Diagonalizzazione di una matrice. Basi di autovettori. Teorema fondamentale. Esempi: algoritmo di Google. Numeri di Fibonacci.

Chapter 1

Campi Numerici

Un campo è un insieme di numeri chiuso rispetto alle quattro operazioni: addizione, moltiplicazione, sottrazione e divisione.

Definition 1.0.1. Un *campo numerico* è una quintupla $(X, +, -, 0, \cdot, ^{-1}, 1)$, dove X è un insieme, $+$ e \cdot sono operazioni binarie, $-$, $^{-1}$ sono operazioni unarie, e $0, 1$ costanti che soddisfano le seguenti equazioni:

1. Proprietà associativa: $x + (y + z) = (x + y) + z; \quad x \cdot (y \cdot z) = (x \cdot y) \cdot z.$
2. Proprietà commutativa: $x + y = y + x; \quad x \cdot y = y \cdot x.$
3. Elemento neutro: $x + 0 = x; \quad x \cdot 1 = x.$
4. Proprietà distributiva: $x \cdot (y + z) = (x \cdot y) + (x \cdot z).$
5. Opposto: $x + (-x) = 0.$
6. Inverso: Se $x \neq 0$, allora $x \cdot x^{-1} = 1.$
7. Prodotto per 0: $x \cdot 0 = 0 = 0 \cdot x.$

Scriveremo

- xy al posto di $x \cdot y$;
- $x - y$ al posto di $x + (-y)$;
- x/y oppure $\frac{x}{y}$ per $x \cdot y^{-1}$.

Inoltre, il prodotto lega più della somma. Per esempio, $x + yz$ significa $x + (yz)$.

La quadrupla $(X, +, -, 0)$ è un *gruppo commutativo* rispetto alla somma, mentre $(X \setminus \{0\}, \cdot, ^{-1}, 1)$ è un *gruppo commutativo* rispetto al prodotto.

Example 1. I seguenti insiemi sono campi numerici:

- L'insieme dei numeri razionali \mathbb{Q} ;
- L'insieme dei numeri reali \mathbb{R} ;
- L'insieme dei numeri complessi \mathbb{C} ;
- L'insieme dei bits $\mathbb{B} = \{0, 1\}$ con le operazioni di somma e prodotto definiti come segue:

$$0 +_2 0 = 1 +_2 1 = 0; \quad 0 +_2 1 = 1 +_2 0 = 1;$$

e

$$0 \times_2 0 = 0 \times_2 1 = 1 \times_2 0 = 0; \quad 1 \times_2 1 = 1.$$

L'opposto di 1 è uno, cioè $-1 = 1$. Questo campo numerico rappresenta l'aritmetica dei numeri modulo 2.

- Sia p un numero primo. Allora l'insieme dei numeri $\{0, 1, \dots, p-1\}$ con le operazioni di addizione $+_p$ e moltiplicazione \times_p modulo p è un campo numerico. Se $x, y \leq p-1$, allora abbiamo:

$$x +_p y = \begin{cases} x + y & \text{if } x + y \leq p-1 \\ r & \text{if } x + y = p + r \text{ per un certo } r. \end{cases}$$

$$x \times_p y = \begin{cases} x \times y & \text{if } x \times y \leq p-1 \\ r & \text{if } x \times y = q \times p + r \text{ con } 0 \leq r \leq p-1. \end{cases}$$

Per esempio, se $p = 5$, abbiamo $3 +_5 2 = 0$ e $3 +_5 3 = 1$, mentre $3 \times_5 3 = 4$ e $4 \times_5 4 = 1$.

1.1 Il campo dei numeri complessi

I numeri reali negativi non ammettono radice quadrata. Questa è una limitazione. Per esempio, l'equazione $2x + 2x + 4 = 0$ non ammette soluzioni reali:

$$x = \frac{-2 \pm \sqrt{-28}}{4} = -\frac{1}{2} \pm \frac{\sqrt{7}}{2}\sqrt{-1}.$$

Le soluzioni dell'equazione $2x + 2x + 4 = 0$ sono numeri complessi perché nella formula risolutiva compare $\sqrt{-1}$.

Il nuovo numero $\sqrt{-1}$, indicato con la lettera “ i ”, soddisfa l'equazione $i^2 = -1$. In generale un numero complesso z è un numero della forma

$$z = a + bi$$

con a, b numeri reali. Il numero a è la *parte reale di z* mentre b è la *parte immaginaria di z* . Questi numeri vengono denotati rispettivamente da $\operatorname{Re}(z) = a$ e $\operatorname{Im}(z) = b$.

Proposition 1.1.1. *I numeri complessi sono un campo numerico, cioè sono chiusi rispetto alle quattro operazioni.*

Proof. I numeri complessi possono essere sommati e moltiplicati:

$$(a+bi)+(c+di) = (a+c)+(b+d)i; \quad (a+bi)(c+di) = (ac-bd)+(ad+bc)i.$$

Si noti che $(bi)(di) = bd(i^2) = bd(-1) = -bd$.

L'opposto del numero complesso $a+bi$ è il numero $-a-bi$, mentre l'inverso del numero $a+bi$ (supponendo che $a+bi \neq 0$) si calcola come segue:

$$\frac{1}{a+bi} = \left(\frac{1}{a+bi} \right) \left(\frac{a-bi}{a-bi} \right) = \frac{a-bi}{(a+bi)(a-bi)} = \frac{a-bi}{a^2+b^2} = \frac{a}{a^2+b^2} - \frac{b}{a^2+b^2}i.$$

□

Example 2. Siano $z = 3+2i$ e $w = 5+7i$ numeri complessi. Allora abbiamo

$$z+w = (3+2i)+(5+7i) = 3+2i+5+7i = (3+5)+(2+7)i = 8+9i.$$

$$zw = (3+2i)(5+7i) = 15+21i+10i+14i^2 = 15+31i+14(-1) = 15-14+31i = 1+31i.$$

Coniugato e modulo di un numero complesso

Se $z = a + ib$ è un numero complesso, il numero $\bar{z} = a - bi$ si chiama il *coniugato* di z . Si calcola facilmente che $z\bar{z} = a^2 + b^2$ è un numero reale.

Valgono le seguenti proprietà:

Proposition 1.1.2. (i) $\overline{z_1 + z_2} = \bar{z}_1 + \bar{z}_2$;

(ii) $\overline{z_1 z_2} = \bar{z}_1 \bar{z}_2$.

(iii) z è un numero reale sse $z = \bar{z}$.

Example 3. Siano $z = 1 + 2i$ e $w = 2 + 2i$. Allora abbiamo:

- $\overline{z + w} = \overline{3 + 4i} = 3 - 4i$.
- $\overline{z + w} = \overline{1 + 2i + 2 + 2i} = (1 - 2i) + (2 - 2i) = 3 - 4i = \overline{z + w}$.
- $\overline{zw} = \overline{(1 + 2i)(2 + 2i)} = \overline{-2 + 6i} = -2 - 6i$.
- $\overline{z} \overline{w} = (1 - 2i)(2 - 2i) = -2 - 6i = \overline{zw}$.

Ogni numero complesso è rappresentato in maniera unica da un punto del piano:

$$a + bi \mapsto (a, b).$$

La distanza del punto (a, b) dall'origine degli assi cartesiani è il modulo del numero complesso z

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

Proposition 1.1.3. Siano z e w numeri complessi. Allora, $|z| = \sqrt{z\bar{z}}$ and $|zw| = |z||w|$.

Proof. Sia $z = a + bi$ e $w = c + di$. Allora $zw = (ac - bd) + (ad + bc)i$.

$$\begin{aligned} |zw|^2 &= (ac - bd)^2 + (ad + bc)^2 \\ &= a^2c^2 + b^2d^2 - 2abcd + a^2d^2 + b^2c^2 + 2abcd \\ &= a^2c^2 + b^2d^2 + a^2d^2 + b^2c^2 \\ &= a^2(c^2 + d^2) + b^2(c^2 + d^2) \\ &= (a^2 + b^2)(c^2 + d^2) \\ &= |z|^2|w|^2. \end{aligned}$$

Un altro metodo di prova:

$$\begin{aligned} |zw|^2 &= (zw)(\bar{zw}) \\ &= zw\bar{z}\bar{w} \quad \text{Proposizione 1.1.2(ii)} \\ &= z\bar{z}w\bar{w} \\ &= |z|^2|w|^2 \end{aligned}$$

Prendendo le radici quadrate positive si ricava la conclusione: $|zw| = |z||w|$. \square

Proposition 1.1.4. *Ogni numero complesso ha una radice quadrata.*

Proof. Sia $z = a + ib$. Dobbiamo trovare un numero $w = x + iy$ tale che $w^2 = z$. Se questo è il caso, allora $w = \sqrt{z}$.

Cerchiamo di determinare delle condizioni che ci permettano di ricavare la parte reale e la parte immaginaria di w . Da $w^2 = (x^2 - y^2) + 2xyi = a + ib$ si ricava:

$$a = x^2 - y^2; \quad b = 2xy.$$

Inoltre, da $z = w^2$ e dalla Proposizione 1.1.3 si ottiene:

$$\sqrt{a^2 + b^2} = |z| = |w^2| = |w||w| = x^2 + y^2.$$

Sommando membro a membro le due uguaglianze $x^2 + y^2 = \sqrt{a^2 + b^2}$ e $x^2 - y^2 = a$ si ha:

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

Possiamo risolvere per x

$$x = \pm \sqrt{\frac{a + \sqrt{a^2 + b^2}}{2}}.$$

Si noti che x è un numero reale! Possiamo poi ottenere y dall'equazione $b = 2xy$:

$$y = \frac{b}{2x}.$$

\square

Example 4. Sia $z = 3 + 4i$. Allora applicando le formule della prova precedente con $a = 3$ e $b = 4$, si ottiene

$$\sqrt{z} = \sqrt{\frac{3 + \sqrt{25}}{2}} + \frac{4i}{2\sqrt{\frac{3+\sqrt{25}}{2}}} = 2 + i.$$

Verifichiamo che elevando al quadrato \sqrt{z} si ottiene effettivamente z : $(2 + i)^2 = (2 + i)(2 + i) = 4 - 1 + 4i = 3 + 4i$.

Da questa proposizione deriviamo che ogni equazione di secondo grado è risolubile nel campo complesso. Vale un risultato molto più generale:

Theorem 1.1.5. (Teorema fondamentale dell'algebra) *Ogni equazione polinomiale di grado n*

$$a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 = 0$$

è risolubile nel campo complesso.

Si dice che il campo complesso è *algebricamente chiuso*.

Il piano complesso

Sia $z = a + bi$ un numero complesso. Il numero z ha una naturale interpretazione geometrica in termini della sua parte reale ed immaginaria: $Re(z) = a$ e $Im(z) = b$ come coppia di punti (a, b) del piano cartesiano xy . La retta dei punti $y = 0$ si dice *asse reale*, mentre la retta $x = 0$ si dice *asse immaginario*. Per esempio l'unità immaginaria i ha coordinate $(0, 1)$. Il piano visto come rappresentazione dei numeri complessi si dice *piano complesso*.

Torniamo al numero $z = a + ib$. I quattro punti $(0, 0), (a, 0), (a, b), (0, b)$ del piano complesso determinano un rettangolo nel piano complesso la cui diagonale principale è un segmento di lunghezza $|z| = \sqrt{a^2 + b^2}$. La diagonale forma un angolo θ con l'asse reale. L'angolo θ ed il modulo $|z|$ determinano univocamente z . Si scrive

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

Se consideriamo il numero complesso $z = a + ib$ come punto (a, b) del piano, allora la coppia $(\sqrt{a^2 + b^2}, \theta)$ determina univocamente il punto (a, b) del piano e definisce le cosiddette *coordinate polari* del punto (a, b) .

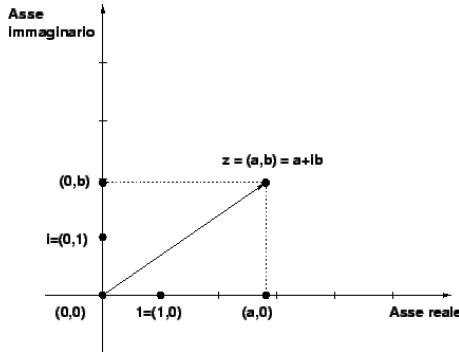
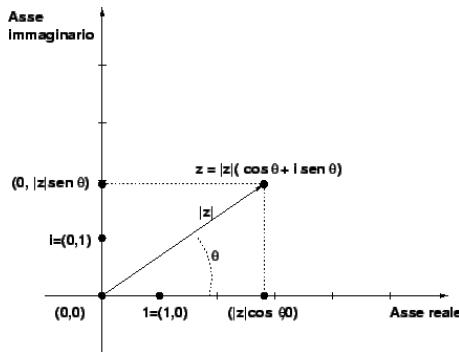


Figure 1.1: piano complesso

Figure 1.2: coordinate polari di z

Per esempio, l'angolo di 90 gradi ed il modulo 1 determinano univocamente l'unità immaginaria, mentre l'angolo di 180 gradi ed il modulo 1 determinano il numero reale -1 , etc.

Proposition 1.1.6. *Siano $z = |z|(\cos \theta + i \sin \theta)$ e $w = |w|(\cos \phi + i \sin \phi)$ due numeri complessi. Allora,*

$$zw = |zw|(\cos(\theta + \phi) + i \sin(\theta + \phi)).$$

Il prodotto di z e w si ottiene moltiplicando i moduli e sommando gli angoli. Si ricordino le formule

$$\sin(\theta + \phi) = \sin \theta \cos \phi + \cos \theta \sin \phi$$

$$\cos(\theta + \phi) = \cos \theta \cos \phi - \sin \theta \sin \phi$$

Proposition 1.1.7. *Sia $z = \cos \theta + i \sin \theta$ un numero complesso di modulo 1 ed n un numero naturale. Allora,*

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

Per esempio, $i^2 = \cos(2\frac{\pi}{2}) + i \sin(2\frac{\pi}{2}) = \cos \pi + i \sin \pi = -1 + 0i = -1$, mentre $i^4 = \cos(4\frac{\pi}{2}) + i \sin(4\frac{\pi}{2}) = \cos(2\pi) + i \sin(2\pi) = \cos 0 + i \sin 0 = 1 + 0i = 1$. L'angolo 2π corrisponde ad un giro completo della circonferenza.

La formula magica di Eulero

Concludiamo questa sezione presentando la formula di Eulero. Il logaritmo naturale $\ln(a)$ del numero reale $a > 0$ (descritto per la prima volta da Nepero) è l'area sottesa dal grafico della funzione $f(x) = \frac{1}{x}$ da $x = 1$ a $x = a$. La base del logaritmo naturale è data dal numero reale e tale che $\ln(e) = 1$. Questo numero $e = 2,71828\dots$ è la costante di Nepero. Il logaritmo è la funzione inversa dell'esponenziale:

$$e^{\ln(x)} = x.$$

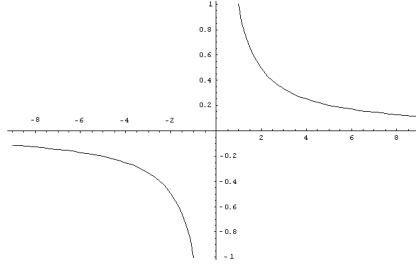


Figure 1.3: grafico di $1/x$

Le funzioni, quando possibile, si approssimano con polinomi considerando l'espansione in serie. Abbiamo:

- $e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \dots;$
- $\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \dots;$
- $\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \dots.$

Applichiamo la funzione al numero complesso $i\theta$:

$$e^{i\theta} = 1 + i\theta + \frac{(i\theta)^2}{2!} + \frac{(i\theta)^3}{3!} + \frac{(i\theta)^4}{4!} + \dots$$

Calcolando otteniamo:

$$e^{i\theta} = 1 + i\theta - \frac{\theta^2}{2!} - \frac{\theta^3}{3!}i + \frac{\theta^4}{4!} + \dots$$

Separando la parte reale dalla parte immaginaria si ha:

$$e^{i\theta} = \left(1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \frac{\theta^6}{6!} + \dots\right) + \left(\theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} + \dots\right)i.$$

Ne segue l'identità di Eulero:

Theorem 1.1.8. $e^{i\theta} = \cos \theta + i \sin \theta$.

Ne segue un'identità che lega tra loro le costanti più importanti della matematica π (lunghezza della circonferenza), e (logaritmo naturale) e i (unità immaginaria):

$$e^{i\pi} = -1.$$

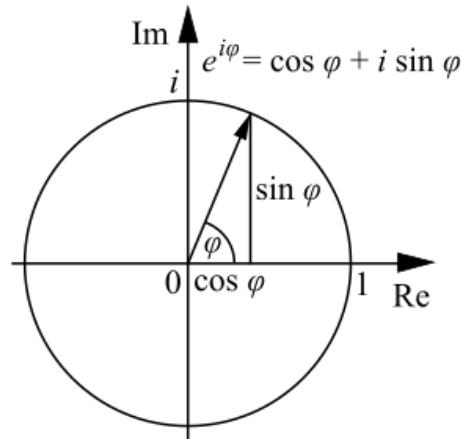


Figure 1.4: formula di Eulero per numeri complessi di modulo 1

Radici dell'unità

Un numero complesso z tale che $z^n = 1$ si chiama *radice n-esima dell'unità*. Nel seguito calcoliamo le radici dell'unità tramite l'interpretazione geometrica dei numeri complessi.

Lemma 1.1.9. *Se $z^n = 1$ allora $|z| = 1$.*

Proof. Dalla Proposizione 1.1.3 segue che $|z^n| = |z|^n = 1$. Dal fatto che $|z| \geq 0$ è un numero reale si ricava $|z| = 1$. \square

Proposition 1.1.10. *L'equazione*

$$z^n = 1$$

ammette come soluzione principale il numero complesso

$$\omega = \cos\left(\frac{2\pi}{n}\right) + i \sin\left(\frac{2\pi}{n}\right).$$

Le altre soluzioni sono le potenze di ω :

$$\omega^k = \cos\left(k \frac{2\pi}{n}\right) + i \sin\left(k \frac{2\pi}{n}\right), \quad \text{per ogni } k \geq 0.$$

Le soluzioni distinte sono $1, \omega, \omega^2, \dots, \omega^{n-1}$.

Proof. Applichiamo la Proposizione 1.1.7 al numero complesso $z = \cos(\theta) + i \sin(\theta)$:

$$z^n = \cos(n\theta) + i \sin(n\theta) = 1 = \cos(k2\pi) + i \sin(k2\pi),$$

perché $\cos(k2\pi) = 1$ e $\sin(k2\pi) = 0$ per ogni numero naturale $k \geq 0$. Si ricava quindi $\theta = k \frac{2\pi}{n}$. \square

Per esempio, le soluzioni dell'equazione $z^3 = 1$ corrispondono ad i numeri complessi nel cerchio unitario di angolo $120^\circ, 240^\circ, 360^\circ$. La radice principale è il numero $z = -\frac{1}{2} + i\frac{\sqrt{3}}{2}$.

Le soluzioni di $z^4 = 1$ corrispondono ad i numeri complessi nel cerchio unitario di angolo $90^\circ, 180^\circ, 270^\circ, 360^\circ$.

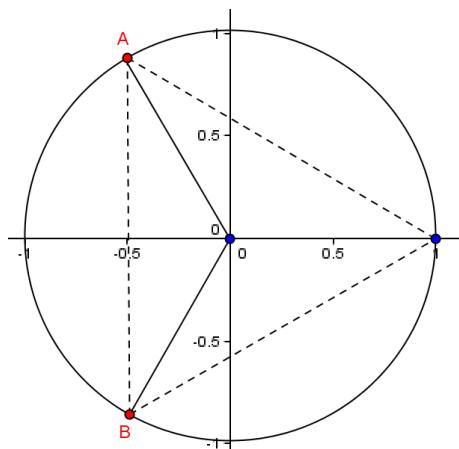


Figure 1.5: Radici terze dell'unità

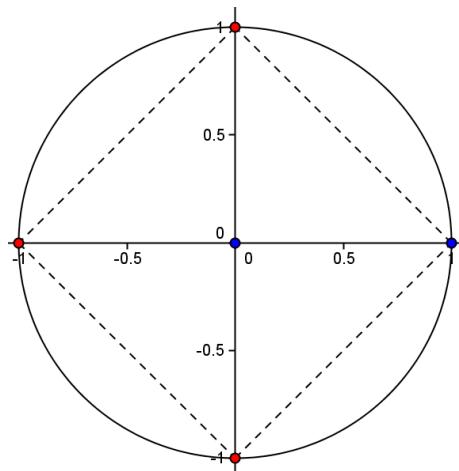


Figure 1.6: Radici quarte dell'unità

Chapter 2

Introduzione agli spazi vettoriali

2.1 Introduzione ai vettori ed al prodotto interno

In Fisica ma anche nella vita di tutti i giorni dobbiamo continuamente misurare qualcosa. Alcune di queste grandezze le possiamo misurare con un numero reale: saldo del conto corrente, altezza, età, etc. Ad altre grandezze corrisponde non solo una quantità rappresentata da un numero ma anche una direzione (con/senza verso). Per esempio,

- La forza di gravità terrestre, la cui direzione ed il cui verso vanno dal punto in cui vi trovate verso il centro della terra;
- La forza di gravità su Giove (molto maggiore della forza di gravità terrestre);
- La forza esercitata in un punto preciso. Ha una grandezza, una direzione ed un verso ben precisi;
- La velocità istantanea di un'automobile. Non conta soltanto il valore, per esempio 120 Km/ora, ma anche la direzione e verso di marcia.

I vettori sono una rappresentazione astratta delle grandezze che hanno una direzione (e talvolta verso).

È importante distinguere tra *vettore liberi* e *vettore applicati*. Se in automobile viaggiamo a velocità costante lungo una linea retta, al tempo t ci troviamo in un determinato punto P della retta, mentre al tempo successivo $t + 10$ ci troveremo in un altro punto Q . Se misuriamo la velocità istantanea nel punto P (velocità misurata + direzione) e poi nel punto Q , otterremo lo stesso risultato. Lo stesso vettore è applicato prima nel punto P e poi nel punto Q .

In generale un vettore è caratterizzato da (a) *lunghezza* (o grandezza, o modulo, o quantità) che è misurata da un valore in un campo numerico (vedi Capitolo 1); (b) *direzione*. Possiamo sempre *moltiplicare un vettore per uno scalare*, che è un elemento del campo numerico con cui misuriamo le lunghezze: Se \mathbf{a} è un vettore ed r uno scalare, allora $r\mathbf{a}$ rappresenta il vettore che ha la stessa direzione di \mathbf{a} ma lunghezza r volte la lunghezza di \mathbf{a} .

Possiamo misurare la direzione ed il verso? La direzione di un vettore non è misurabile con un numero. Possiamo soltanto sapere quando due vettori \mathbf{a} e \mathbf{b} hanno la stessa direzione:

- I vettori \mathbf{a} e \mathbf{b} *hanno la stessa direzione* se esiste uno “scalare” r del campo numerico tale che $\mathbf{a} = r\mathbf{b}$.

Se il campo numerico è totalmente ordinato, come nel caso dei numeri reali oppure i numeri razionali, possiamo dire anche se due vettori \mathbf{a} e \mathbf{b} *hanno lo stesso verso*:

- I vettori \mathbf{a} e \mathbf{b} *hanno stessa direzione e verso* se esiste uno scalare $r > 0$ tale che $\mathbf{a} = r\mathbf{b}$. I vettori \mathbf{a} e \mathbf{b} hanno stessa direzione ma verso opposto se esiste $r < 0$ tale che $\mathbf{a} = r\mathbf{b}$.

Il campo dei numeri complessi non ha un ordinamento naturale. Quindi i vettori complessi hanno una direzione, ma non un verso.

Oltre alla moltiplicazione per uno scalare, i vettori ammettono un’altra operazione, che è detta *somma o addizione di vettori*. Per spiegare la somma vettoriale, immaginiamo di effettuare il seguente esperimento. Appoggiamo una palla enorme nell’origine O del piano cartesiano xy . Immaginiamo che due persone Pinco e Pallino spingano la palla con forza. Pinco spinge lungo l’asse delle y (che corrisponde alla direzione della retta $x = 0$) dall’asse negativo delle y verso l’asse positivo delle y . Pallino spinge lungo l’asse delle x (che corrisponde alla direzione della retta $y = 0$) dall’asse negativo delle x verso l’asse positivo delle x . La palla riceve una spinta da Pinco

rappresentata dal vettore \mathbf{a} (grandezza, direzione, verso) e un'altra spinta da Pallino rappresentata dal vettore \mathbf{b} . In che direzione si muoverà la palla? Se le lunghezze dei due vettori sono uguali (cioé, pari spinta), allora la palla comincerà a muoversi lungo la retta $y = x$, che è la direzione della retta che biseca l'angolo di novanta gradi formato dalla retta $y = 0$ e la retta $x = 0$. La spinta totale che riceve la palla è rappresentata dal vettore $\mathbf{a} + \mathbf{b}$. Se la lunghezza del vettore \mathbf{a} è 1 (si scrive $\|\mathbf{a}\| = 1$) e se $\|\mathbf{b}\| = 1$, allora la lunghezza del vettore $\mathbf{a} + \mathbf{b}$ è $\sqrt{2}$ (non 2 come si potrebbe pensare).

Se Pinco e Pallino spingono con pari forza, il primo dall'asse positivo delle y verso l'asse negativo delle y ed il secondo dall'asse negativo delle y verso l'asse positivo delle y , la palla non si muoverà nonostante lo sforzo di entrambi. La somma di due vettori di pari lunghezza e direzione ma di verso opposto è nulla.

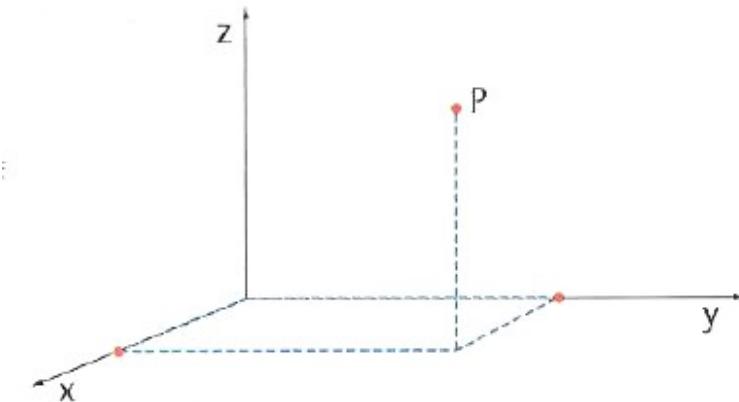


Figure 2.1: Forza e vettori

2.2 Spazi vettoriali

2.2.1 Vettori nello spazio

Fissiamo un sistema di assi cartesiani nello spazio euclideo. Questo corrisponde a scegliere un punto O , l'origine degli assi, e tre rette x, y, z passanti per O perpendicolari tra loro. Un punto P nello spazio euclideo di assi Cartesiani xyz è rappresentato da una terna $P = [p_1, p_2, p_3]$ di numeri reali, le sue coordinate Cartesiane. La prima coordinata p_1 e la seconda p_2 si ottengono proiettando perpendicolarmente il punto P nel piano $z=0$, ottenendo il punto p_{xy} . Poi si proietta questo punto p_{xy} sull'asse delle x (sull'asse delle y , rispettivamente) per ottenere p_1 (p_2 , rispettivamente). Similmente la restante coordinata p_3 si ottiene proiettando il punto P nel piano $x = 0$, ottenendo il punto p_{yz} . Poi si proietta questo punto sull'asse z e si ottiene p_3 .

Figure 2.2: Coordinate del punto P

Un punto $P = [p_1, p_2, p_3] \neq [0, 0, 0]$ dello spazio determina univocamente un vettore \overrightarrow{OP} che va dall'origine $O = [0, 0, 0]$ degli assi al punto P . Questo vettore ha come lunghezza la lunghezza $\sqrt{p_1^2 + p_2^2 + p_3^2}$ del segmento OP che unisce l'origine degli assi al punto P , come direzione l'unica retta che passa attraverso i punti O e P , e come verso quello che va dal punto O al punto P . Il vettore \overrightarrow{OP} si indicherà spesso con il punto P stesso. Quindi parleremo di vettore P intendendo il vettore \overrightarrow{OP} .

I vettori possono essere sommati coordinata per coordinata. Per esempio, se $P = [2, 3, 4]$ e $Q = [1, 2, 3]$ allora $P + Q = [3, 5, 7]$. Geometricamente il punto $P + Q$ si ottiene costruendo il parallelogramma di vertici P , O e Q . Il quarto vertice è proprio il punto $P + Q$. Si veda la figura.

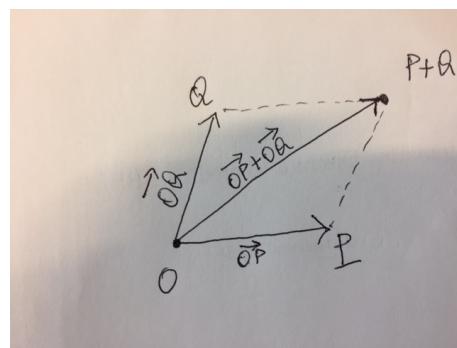


Figure 2.3: Somma di vettori

I vettori possono essere moltiplicati per uno scalare. Se \overrightarrow{OP} è un vettore nello spazio ed r è un numero reale, detto scalare, allora $r\overrightarrow{OP}$ è un altro vettore che sta sempre nella retta, passante per l'origine, determinata dal vettore \overrightarrow{OP} . Per esempio, sia $P = [3, 2, 1]$ un punto, 5 uno scalare e Q il punto di coordinate $5P = [15, 10, 5]$. Allora il vettore $5\overrightarrow{OP}$ è uguale al vettore \overrightarrow{OQ} . Nella prossima sottosezione studieremo in generale le proprietà della somma vettoriale e del prodotto per uno scalare.

2.2.2 Vettori in astratto

In generale, possiamo considerare in maniera astratta un insieme di vettori che possono essere sommati tra loro e moltiplicati per uno scalare. Uno scalare è un elemento di un fissato campo numerico \mathbb{K} . Nelle applicazioni \mathbb{K} sarà uno dei seguenti campi: il campo \mathbb{Q} dei numeri razionali, il campo \mathbb{R} dei numeri reali, il campo \mathbb{C} dei numeri complessi, oppure il campo \mathbb{Z}_p dei numeri interi modulo un numero primo p .

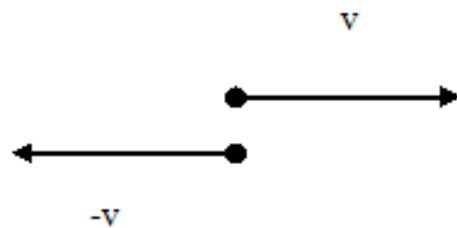


Figure 2.4: Vettori opposti.

Definition 2.2.1. Sia \mathbb{K} un campo numerico, i cui elementi sono chiamati scalari. Uno *spazio vettoriale* su \mathbb{K} è costituito da un insieme V di vettori dotati di somma vettoriale $+ : V \times V \rightarrow V$ e prodotto per uno scalare $\cdot : \mathbb{K} \times V \rightarrow V$, che soddisfano gli assiomi seguenti ($\mathbf{a}, \mathbf{b}, \mathbf{c} \in V$ sono vettori arbitrari e $r, s \in \mathbb{K}$ sono scalari arbitrari):

$$\mathbf{SV1: } \mathbf{a} + (\mathbf{b} + \mathbf{c}) = (\mathbf{a} + \mathbf{b}) + \mathbf{c};$$

$$\mathbf{SV2: } \mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a};$$

$$\mathbf{SV3: } \mathbf{0} + \mathbf{a} = \mathbf{a} = \mathbf{a} + \mathbf{0};$$

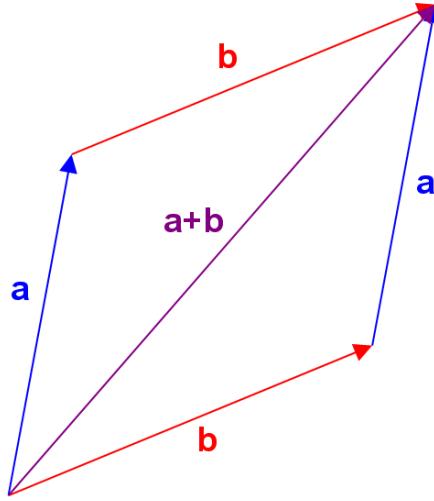


Figure 2.5: La proprietà commutativa della somma vettoriale.

SV4: $\mathbf{a} + (-\mathbf{a}) = \mathbf{0} = (-\mathbf{a}) + \mathbf{a}$;

SV5: $(r+s)\mathbf{a} = r\mathbf{a} + s\mathbf{a}$;

SV6: $(rs)\mathbf{a} = r(s\mathbf{a})$;

SV7: $r(\mathbf{a} + \mathbf{b}) = r\mathbf{a} + r\mathbf{b}$;

SV8: $0\mathbf{a} = \mathbf{0}$; $1\mathbf{a} = \mathbf{a}$; $(-1)\mathbf{a} = -\mathbf{a}$.

Si noti che il vettore nullo viene indicato con $\mathbf{0}$ e che per brevità scriviamo $r\mathbf{a}$ al posto di $r \cdot \mathbf{a}$.

Si noti anche che l'assioma (SV4) deriva da (SV5) e (SV8):

$$\mathbf{a} + (-\mathbf{a}) = 1\mathbf{a} + (-1)\mathbf{a} = (1 + (-1))\mathbf{a} = 0\mathbf{a} = \mathbf{0}.$$

Example 5. Sia \mathbb{R} il campo dei numeri reali. L'insieme \mathbb{R}^2 delle coppie di numeri reali costituisce uno spazio vettoriale reale con le seguenti operazioni:

$$[x_1, x_2] + [y_1, y_2] = [x_1 + y_1, x_2 + y_2]; \quad r[x_1, x_2] = [rx_1, rx_2]; \quad \mathbf{0} = [0, 0].$$

Per esempio, $[3, 6] + [1, 2] = [4, 8]$ e $7[1, 3] = [7, 21]$.

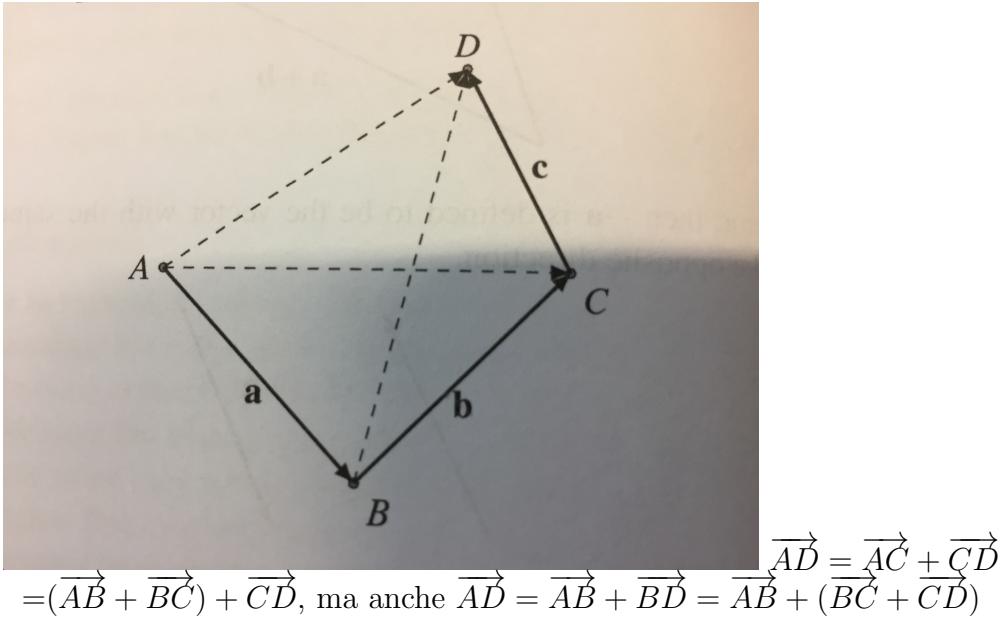


Figure 2.6: Proprietà associativa della somma vettoriale

Example 6. Sia \mathbb{R} il campo dei numeri reali. L'insieme \mathbb{C} dei numeri complessi costituisce uno spazio vettoriale reale con le seguenti operazioni (a, b, c, d, r numeri reali e i unità immaginaria):

$$(a + bi) + (c + di) = (a + c) + (b + d)i; \quad r(a + bi) = (ra) + (rb)i; \quad \mathbf{0} = 0.$$

Per esempio, $(3 + 6i) + (1 + 2i) = 4 + 8i$ e $7(1 + 3i) = 7 + 21i$.

Example 7. Sia \mathbb{C} il campo dei numeri complessi. L'insieme \mathbb{C} dei numeri complessi costituisce uno spazio vettoriale sul campo dei numeri complessi con le seguenti operazioni (a, b, c, d numeri reali, $r = r_1 + r_2i$ numero complesso e i unità immaginaria):

$$(a + bi) + (c + di) = (a + c) + (b + d)i;$$

$$r(a + bi) = (r_1 + r_2i)(a + bi) = (r_1a - r_2b) + (r_1b + r_2a)i; \quad \mathbf{0} = 0.$$

I vettori possono essere utilizzati per fornire delle prove di risultati geometrici. Un esempio è fornito dal seguente teorema.

Theorem 2.2.1. *Dato un arbitrario quadrilatero convesso $ABCD$, il quadrilatero avente per vertici i punti medi dei lati consecutivi di $ABCD$ costituisce un parallelogramma (Si veda la figura a pagina 24).*

Fissiamo in un qualsiasi quadrilatero $ABCD$ i punti mediani dei suoi lati consecutivi. Chiamiamo questi punti P, Q, R, S . Siano $\mathbf{a} = \overrightarrow{AB}$, $\mathbf{b} = \overrightarrow{BC}$, $\mathbf{c} = \overrightarrow{CD}$, $\mathbf{d} = \overrightarrow{DA}$ i vettori come in figura a pagina 24. Allora $\mathbf{a} + \mathbf{b} + \mathbf{c} + \mathbf{d} = \mathbf{0}$. Quindi $\mathbf{a} + \mathbf{b} = -(\mathbf{c} + \mathbf{d})$. Si vede sempre dalla figura che $\overrightarrow{PQ} = \frac{1}{2}\mathbf{a} + \frac{1}{2}\mathbf{b} = \frac{1}{2}(\mathbf{a} + \mathbf{b})$ e $\overrightarrow{RS} = \frac{1}{2}\mathbf{c} + \frac{1}{2}\mathbf{d} = \frac{1}{2}(\mathbf{c} + \mathbf{d})$. Quindi, applicando l'uguaglianza $\mathbf{a} + \mathbf{b} = -(\mathbf{c} + \mathbf{d})$, si ricava $\overrightarrow{PQ} = -\overrightarrow{RS}$. Segue che i vettori \overrightarrow{PQ} e \overrightarrow{RS} sono su rette parallele. In maniera simile si prova che il vettore \overrightarrow{QR} è parallelo al vettore \overrightarrow{SP} .

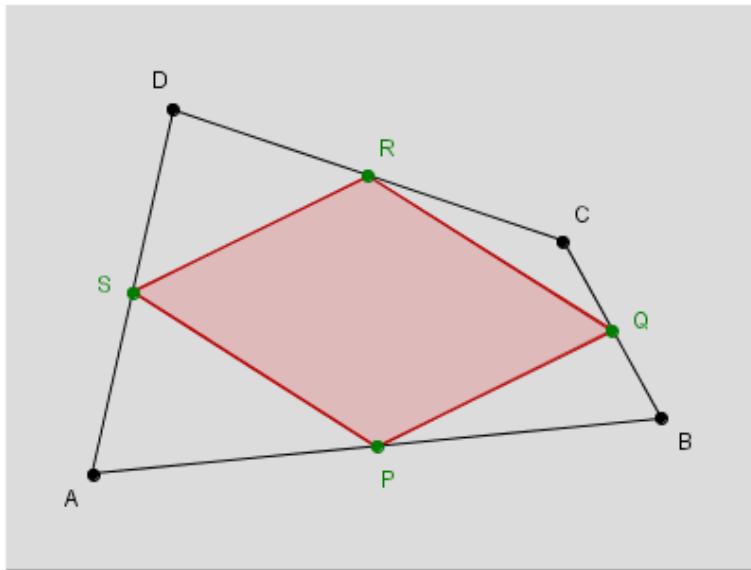


Figure 2.7: Prova geometrica con vettori

2.3 Prodotto interno (o scalare)

In questa sezione definiamo il prodotto interno di due vettori dello spazio. Analoghe definizioni possono essere date per vettori di \mathbb{R}^n con n arbitrario. Lasciamo al lettore la facile generalizzazione.

Il prodotto interno di due vettori è uno scalare, proporzionale alla lunghezza dei due vettori, che ci dà informazione sull'orientazione relativa dei due vettori tramite il coseno dell'angolo da essi formato. In particolare, due vettori sono perpendicolari tra loro se il loro prodotto interno è nullo.

Un punto A dello spazio \mathbb{R}^3 rappresenta anche il vettore \overrightarrow{OA} che va dall'origine O delle coordinate Cartesiane sino al punto A . Quindi in seguito parleremo di punto A oppure di vettore A senza alcuna distinzione.

Definiamo il prodotto interno in termini delle coordinate dei due vettori. L'interpretazione geometrica del prodotto interno verrà data nel seguito.

Definition 2.3.1. Se $A = [a_1, a_2, a_3]$ e $B = [b_1, b_2, b_3]$ sono due vettori di \mathbb{R}^3 , allora il *prodotto interno (o scalare)* di A e B è definito come segue:

$$A \cdot B = a_1 b_1 + a_2 b_2 + a_3 b_3.$$

Per esempio, se $A = [2, 3, 2]$ e $B = [-3, 4, 3]$ allora $A \cdot B = 12$.

Lemma 2.3.1. Il prodotto interno verifica le seguenti proprietà:

PS1 $A \cdot B = B \cdot A$;

PS2 $A \cdot (B + C) = A \cdot B + A \cdot C = (B + C) \cdot A$;

PS3 Se r è uno scalare $(rA) \cdot B = r(A \cdot B) = A \cdot (rB)$;

PS4 Se $\mathbf{0}$ è il vettore nullo, allora $\mathbf{0} \cdot \mathbf{0} = 0$; in ogni altro caso $A \cdot A > 0$.

Se $A = [a_1, a_2, a_3]$, allora la *lunghezza (o norma o modulo) del vettore A* è definita come

$$\|A\| = \sqrt{A \cdot A} = \sqrt{a_1^2 + a_2^2 + a_3^2}.$$

Distanza tra due punti

La lunghezza del vettore A è pari alla distanza del punto A dall'origine delle coordinate Cartesiane. Se A e B sono due punti dello spazio, la distanza tra A e B è la lunghezza del vettore applicato \overrightarrow{AB} che va da A a B . Se riportiamo questo vettore applicato nell'origine delle coordinate Cartesiane, cioè consideriamo il punto $B - A$, otteniamo che la distanza tra A e B è

$$\|B - A\|.$$

Perpendicolarità

Quando due vettori sono perpendicolari od ortogonali?

Proposition 2.3.2. *Due vettori A e B sono perpendicolari sse il loro prodotto interno $A \cdot B$ è 0.*

Proof. Due vettori A e B sono perpendicolari se e solo se la distanza tra A e B e la distanza tra A e $-B$ sono uguali (l'angolo formato dai vettori A e B è uguale all'angolo formato dai vettori A e $-B$). Questo significa che

$$\|A - B\| = \|A - (-B)\| = \|A + B\|.$$

Calcoliamo con i quadrati delle distanze:

$$\begin{aligned} 0 &= \|A - B\|^2 - \|A + B\|^2 \\ &= [(A - B) \cdot (A - B)] - [(A + B) \cdot (A + B)] \\ &= [(A \cdot A) - 2(A \cdot B) + (B \cdot B)] - [(A \cdot A) + 2(A \cdot B) + (B \cdot B)] \\ &= 4(A \cdot B) \end{aligned}$$

da cui segue $A \cdot B = 0$. \square

Nel seguito presentiamo alcune applicazioni geometriche del prodotto interno.

Theorem 2.3.3. *Se un triangolo ABC inscritto in una circonferenza ha un diametro come lato, allora il triangolo è rettangolo (si veda la Figura a pagina 27).*

Proof. Sia O il centro della circonferenza e siano A, B, C tre punti distinti della circonferenza che determinano un triangolo ABC tale che il segmento AB è un diametro della circonferenza. Allora $\overrightarrow{CO} + \overrightarrow{OB} = \overrightarrow{CB}$ ed inoltre $\overrightarrow{CO} + \overrightarrow{OA} = \overrightarrow{CA}$. Se indichiamo con $\mathbf{u} = \overrightarrow{OB}$ e $\mathbf{v} = \overrightarrow{OC}$, allora abbiamo che $\overrightarrow{OA} = -\mathbf{u}$ e

$$\overrightarrow{CB} = -\mathbf{v} + \mathbf{u}; \quad \overrightarrow{CA} = -\mathbf{v} - \mathbf{u}.$$

Calcolando il prodotto interno di \overrightarrow{CB} e \overrightarrow{CA} si ha:

$$\begin{aligned} \overrightarrow{CB} \cdot \overrightarrow{CA} &= (-\mathbf{v} + \mathbf{u}) \cdot (-\mathbf{v} - \mathbf{u}) \\ &= (\mathbf{v} \cdot \mathbf{v}) + (\mathbf{v} \cdot \mathbf{u}) - (\mathbf{u} \cdot \mathbf{v}) - (\mathbf{u} \cdot \mathbf{u}) \\ &= (\mathbf{v} \cdot \mathbf{v}) - (\mathbf{u} \cdot \mathbf{u}) \\ &= 0 \end{aligned}$$

perché $\|\mathbf{v}\|^2 = \mathbf{v} \cdot \mathbf{v}$, $\|\mathbf{u}\|^2 = \mathbf{u} \cdot \mathbf{u}$ e i vettori \mathbf{v} e \mathbf{u} hanno la stessa lunghezza uguale al raggio del cerchio. In conclusione, i vettori \overrightarrow{CB} e \overrightarrow{CA} sono perpendicolari e l'angolo tra di loro compreso è di 90 gradi. \square

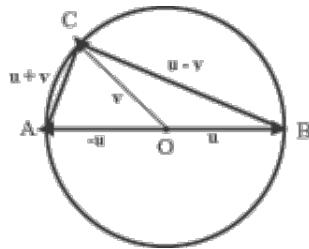


Figure 2.8: Triangolo inscritto in una circonferenza con diametro come lato

Vediamo ora come si può dimostrare il Teorema di Pitagora tramite il prodotto interno. Sia ABC un triangolo rettangolo come in figura a pagina 27. Indichiamo con $\mathbf{a} = \overrightarrow{BC}$, $\mathbf{b} = \overrightarrow{CA}$ e $\mathbf{c} = \overrightarrow{AB}$ i tre vettori dei lati

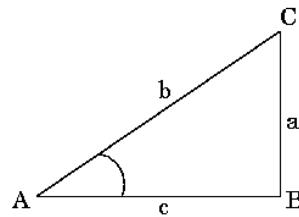


Figure 2.9: Triangolo rettangolo

consecutivi. Supponiamo che l'angolo retto sia l'angolo compreso tra i vettori \mathbf{a} e \mathbf{c} , cosicché il prodotto interno $\mathbf{a} \cdot \mathbf{c} = 0$ è nullo. Dal fatto che $\mathbf{a} + \mathbf{b} + \mathbf{c} = \mathbf{0}$, si ricava $\mathbf{b} = -(\mathbf{a} + \mathbf{c})$. Allora si ha:

$$\|\mathbf{b}\|^2 = \mathbf{b} \cdot \mathbf{b} = (-\mathbf{a} - \mathbf{c}) \cdot (-\mathbf{a} - \mathbf{c}) = (\mathbf{a} \cdot \mathbf{a}) + 2(\mathbf{a} \cdot \mathbf{c}) + (\mathbf{c} \cdot \mathbf{c}) = (\mathbf{a} \cdot \mathbf{a}) + (\mathbf{c} \cdot \mathbf{c}) = \|\mathbf{a}\|^2 + \|\mathbf{c}\|^2$$

perché $\mathbf{a} \cdot \mathbf{c} = 0$.

2.3.1 Coseno dell'angolo formato da due vettori

La parte restante di questa sezione ha lo scopo di determinare il significato geometrico del prodotto interno:

$$A \cdot B = \|A\| \|B\| \cos \theta,$$

dove θ è l'angolo formato dai vettori A e B (ovvero l'angolo \widehat{AOB}).

Lemma 2.3.4. (Disuguaglianza di Schwartz)

$$(A \cdot B)^2 \leq (A \cdot A)(B \cdot B).$$

Proof. Se A oppure B è il vettore nullo, la disuguaglianza vale. Supponiamo quindi che sia A che B sono diversi dal vettore nullo $\mathbf{0}$. Siano x e y numeri reali arbitrari. Applicando la proprietà (PS4) abbiamo:

$$0 \leq (xA + yB) \cdot (xA + yB).$$

Sviluppando otteniamo:

$$0 \leq x^2(A \cdot A) + 2xy(A \cdot B) + y^2(B \cdot B).$$

Poniamo $x = B \cdot B > 0$. Dividendo per x si ottiene:

$$0 \leq (B \cdot B)(A \cdot A) + 2y(A \cdot B) + y^2.$$

Il risultato finale si ottiene ponendo $y = -(A \cdot B)$. Si ricava

$$0 \leq (B \cdot B)(A \cdot A) - 2(A \cdot B)^2 + (A \cdot B)^2 = (B \cdot B)(A \cdot A) - (A \cdot B)^2.$$

Riportando al primo membro $(A \cdot B)^2$ si ottiene la tesi. \square

Corollary 2.3.5.

$$|A \cdot B| \leq \|A\| \|B\|.$$

Lemma 2.3.6. $\|A + B\| \leq \|A\| + \|B\|$ ed inoltre $\|xA\| = x\|A\|$,

Proof. Sviluppiamo $\|A+B\|^2 = (A+B) \cdot (A+B) = A \cdot A + 2(A \cdot B) + B \cdot B = \|A\|^2 + 2(A \cdot B) + \|B\|^2 \leq \|A\|^2 + 2\|A\| \|B\| + \|B\|^2 = (\|A\| + \|B\|)^2$. \square

Proposition 2.3.7. Siano A e B due vettori, e sia θ l'angolo da essi formato. Allora si ha:

$$A \cdot B = \|A\| \|B\| \cos \theta.$$

Proof. Sia θ l'angolo formato dai vettori A e B . Se proiettiamo il vettore A sulla retta che contiene B , otteniamo un vettore rB (per un certo $r \in \mathbb{R}$) che ha la stessa direzione di B . Dalla trigonometria elementare si ha che la lunghezza $r\|B\|$ del vettore rB è uguale a $\|A\| \cos \theta$, mentre la lunghezza di

$A - rB$ è uguale a $\|A\| \sin(\theta)$ ed è perpendicolare al vettore B . Utilizzando la Proposizione 2.3.2 si ha:

$$(A - rB) \cdot B = A \cdot B - r(B \cdot B) = 0$$

da cui

$$r = \frac{A \cdot B}{B \cdot B} = \frac{A \cdot B}{\|B\|^2}.$$

Quindi, da

$$r\|B\| = \|A\| \cos \theta$$

sostituendo a r il suo valore otteniamo:

$$\left(\frac{A \cdot B}{\|B\|^2} \right) \|B\| = \|A\| \cos \theta.$$

Semplificando ricaviamo:

$$\frac{A \cdot B}{\|B\|} = \|A\| \cos \theta.$$

da cui si ha la tesi. \square

Due vettori sono allineati se $\cos \theta = +1$ oppure $\cos \theta = -1$, da cui si ha:

$$|A \cdot B| = \|A\| \|B\|.$$

Chapter 3

Rette e piani

Una *equazione lineare* nelle incognite x_1, \dots, x_n a coefficienti nel campo numerico \mathbb{K} è un'equazione:

$$a_1x_1 + \dots + a_nx_n = b \quad (3.1)$$

con $a_1, \dots, a_n, b \in \mathbb{K}$. Gli elementi a_i sono detti coefficienti, mentre b è il termine noto. Se $b = 0$, l'equazione è detta lineare *omogenea*. Una soluzione dell'equazione lineare è una n-upla (ordinata) (r_1, \dots, r_n) di elementi di \mathbb{K} tale che

$$a_1r_1 + \dots + a_nr_n = b.$$

L'insieme di tutte le soluzioni dell'equazione lineare (3.1) è detto *iperpiano* (di dimensione $n - 1$) nello spazio vettoriale \mathbb{K}^n di dimensione n .

In questa sezione utilizzeremo come campo numerico l'insieme \mathbb{R} dei numeri reali e ci limiteremo alle dimensioni $n = 2$ e $n = 3$. In tal caso, l'equazione lineare $ax + by = c$ ($a, b, c \in \mathbb{R}$) definisce una retta nel piano, mentre l'equazione lineare $ax + by + cz = d$ ($a, b, c, d \in \mathbb{R}$) definisce un piano nello spazio.

Un'equazione lineare può essere espressa tramite il prodotto interno. Ci limitiamo alla dimensione tre. Il lettore può facilmente generalizzare il lemma a dimensione 2 oppure a dimensione n arbitraria.

Lemma 3.0.1. *Sia $ax_1 + bx_2 + cx_3 = d$ un'equazione lineare in tre incognite, $\mathbf{a} = [a, b, c]$ il vettore dei coefficienti e $\mathbf{x} = [x_1, x_2, x_3]$ il vettore delle incognite. Allora l'equazione lineare si può scrivere, utilizzando il prodotto interno, come segue:*

$$\mathbf{a} \cdot \mathbf{x} = d$$

Essa esprime il luogo di tutti vettori \mathbf{x} il cui prodotto interno con il vettore costante \mathbf{a} dà come risultato lo scalare d . Se $d = 0$, l'equazione $\mathbf{a} \cdot \mathbf{x} = 0$ descrive il piano dei vettori perpendicolari al vettore \mathbf{a} .

3.1 Rette nel piano

Cominciamo questa sezione con lo studio delle rette del piano passanti per l'origine. Esse sono definite da equazioni lineari omogenee in due incognite.

- **Retta di equazione $ax + by = 0$**

Utilizzando il prodotto interno di vettori, l'equazione lineare $3x + 2y = 0$ (x, y variabili reali) si può scrivere nel modo seguente: $[3, 2] \cdot [x, y] = 0$. La retta $3x + 2y = 0$ descrive quindi l'insieme dei vettori $[x, y]$ che sono perpendicolari al vettore $[3, 2]$. Il luogo dei punti

$$\{[x, y] : 3x + 2y = 0\}$$

è la retta passante per l'origine e perpendolare al vettore $[3, 2]$. Se conosciamo un punto sulla retta, per esempio il punto $[2, -3]$, allora tutti gli altri possono essere ottenuti moltiplicando il vettore $[2, -3]$ per lo scalare t numero reale arbitrario (equazione parametrica della retta):

$$x = 2t; \quad y = -3t, \text{ al variare del reale } t.$$

In generale, se $[a, b]$ è un vettore, allora $ax + by = 0$ è l'equazione della retta passante per l'origine i cui punti $[x, y]$ sono vettori perpendicolari al vettore $[a, b]$.

Consideriamo ora le rette del piano che non passano per l'origine.

- **Retta di equazione $ax + by = c$ con $c \neq 0$**

Utilizzando il prodotto interno, l'equazione $3x + 2y = 5$ si può scrivere come segue $[3, 2] \cdot [x, y] = 5$. Per capire il rapporto esistente tra l'equazione $3x + 2y = 5$ e l'equazione omogenea $3x + 2y = 0$ fissiamo un punto, per esempio $P = [1, 1]$, che appartiene alla retta $3x + 2y = 5$. Esiste una corrispondenza bigettiva tra i punti nella retta $3x + 2y = 5$ ed i punti nella retta $3x + 2y = 0$. Se $Q = [q_0, q_1]$ soddisfa l'equazione $3q_0 + 2q_1 = 0$ allora $P + Q = [1 + q_0, 1 + q_1]$ soddisfa $3(1+q_0) + 2(1+q_1) = 5$. Viceversa, se $R = [r_0, r_1]$ soddisfa $3r_0 + 2r_1 = 5$

allora $R - P = [r_0 - 1, r_1 - 1]$ soddisfa $3(r_0 - 1) + 2(r_1 - 1) = 0$. Le due funzioni sono l'una inversa dell'altra.

Siccome i punti della retta $3x + 2y = 0$ sono descrivibili parametricamente come tQ con t scalare arbitrario e Q punto nella retta (per esempio, $Q = [2, -3]$), allora la retta $3x + 2y = 5$ si descrive parametricamente come $P + tQ$. Quindi, l'insieme dei punti $[1, 1] + t[2, -3] = [1 + 2t, 1 - 3t]$ sono tutti e soli i punti che verificano l'equazione $3x + 2y = 5$:

$$[3, 2] \cdot [1 + 2t, 1 - 3t] = 3(1 + 2t) + 2(1 - 3t) = 5.$$

Quindi la seguente equazione parametrica descrive la retta:

$$x = 1 + 2t; \quad y = 1 - 3t, \text{ al variare del numero reale } t.$$

La retta data è parallela alla retta di equazione $3x + 2y = 0$ e passa per il punto $[1, 1]$.

Lasciamo al lettore la facile generalizzazione al caso generale $ax + by = c$ con $c \neq 0$.

- **Dato un vettore A ed un punto P nel piano, determinare l'equazione della retta passante per P e parallela alla retta che include il vettore A**

Spieghiamo il procedimento con un esempio. Sia $A = [3, 2]$ e $P = [1, -1]$. Prima determiniamo i valori a e b della retta $ax + by = 0$ passante per l'origine e per il punto A . Siccome A sta nella retta abbiamo

$$3a + 2b = 0,$$

da cui $b = -(3/2)a$. Se scegliamo $a = 2$, si ricava $b = -3$. Allora l'equazione della retta che contiene il vettore A è

$$2x - 3y = 0.$$

La nuova retta avrà equazione

$$2x - 3y = c$$

per un certo $c \neq 0$. Il punto P sta in questa retta, quindi $c = 2 - 3(-1) = 5$. Quindi l'equazione della retta passante per P e parallela alla retta di direzione il vettore A è:

$$2x - 3y = 5.$$

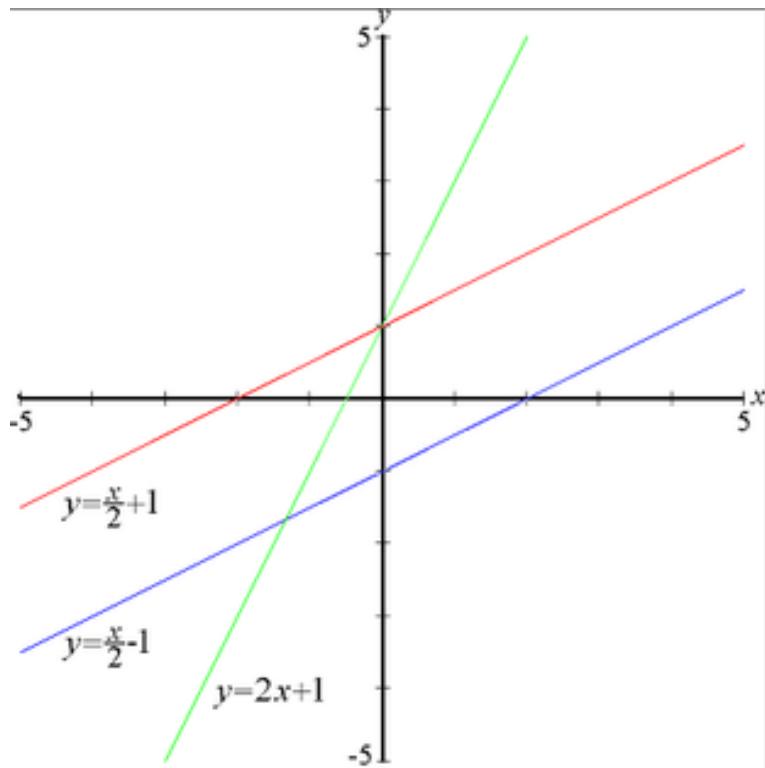


Figure 3.1: Rette nel piano

- **Dati due punti distinti A e B nel piano, determinare l'equazione della retta passante per A e B**

Spieghiamo il procedimento con un esempio. Sia $A = [2, 2]$ e $B = [-2, 6]$. La retta passante per A e B conterrà il vettore applicato \vec{AB} . Riportiamo tale vettore nell'origine tramite il punto $B - A = [-4, 4]$. Quindi il vettore \vec{AB} avrà la stessa lunghezza e direzione del vettore $B - A$.

Equazione della retta contenente il vettore $B - A = [-4, 4]$: dall'equazione generica $ax + by = 0$ e dal fatto che il punto $B - A$ appartiene alla retta si ricava che $-4a + 4b = 0$ da cui $a = b$. Scegliamo $a = b = 1$. Quindi l'equazione è $x + y = 0$.

Equazione della retta parallela alla retta di equazione $x + y = 0$ e contenente il punto $B = [-2, 6]$: Da $x + y = c$ sostituendo le coordinate del punto B si ricava $c = -2 + 6 = 4$.

Quindi l'equazione della retta è $x + y = 4$. Anche il punto A sta nella retta in quanto $2 + 2 = 4$.

- **Come passare dall'equazione parametrica di una retta all'equazione $ax + by = c$**

Se descriviamo una retta con un'equazione parametrica, per esempio

$$x = 5 + 2t; \quad y = 2 + t$$

ricaviamo t da una delle due equazioni e lo sostituiamo nell'altra. Da $t = y - 2$ si ottiene $x = 5 + 2(y - 2)$ da cui $x - 2y = 1$.

3.2 Rette e piani nello spazio

Cominciamo con lo studio dei piani passanti per l'origine. Essi sono definiti da equazioni lineari omogenee in tre incognite.

- **Piano passante per l'origine e perpendicolare ad un dato vettore**

Nello spazio abbiamo tre coordinate. Consideriamo tre variabili x, y, z . Cosa descrive l'equazione $3x + 2y + z = 0$? Un piano passante per l'origine. Consideriamo il vettore $[3, 2, 1]$ ed il vettore $[x, y, z]$. L'equazione $3x + 2y + z = 0$ si può scrivere utilizzando il prodotto interno: $[3, 2, 1] \cdot [x, y, z] = 0$. Ossia il vettore $[x, y, z]$ è perpendicolare al vettore $[3, 2, 1]$. Quindi l'equazione descrive l'insieme di tutti i vettori perpendicolari al vettore $[3, 2, 1]$. Il luogo dei punti

$$\{[x, y, z] : 3x + 2y + z = 0\}$$

è il piano passante per l'origine e perpendicolare al vettore $[3, 2, 1]$.

L'equazione parametrica del piano si ottiene come segue. Consideriamo due punti P e Q nel piano $3x + 2y + z = 0$, che non siano allineati rispetto alle rette del piano passanti per l'origine degli assi (non possiamo considerare soltanto un punto perché il piano è bidimensionale). Per esempio, $P = [1, -1, -1]$ e $Q = [-1, 0, 3]$. Allora tutti i punti del piano $3x + 2y + z = 0$ si ottengono come combinazione lineare di P e Q :

$$tP + rQ, \text{ al variare di } r, t \text{ numeri reali,}$$

In altri termini tutti i punti del tipo

$$x = t - r; \quad y = -t; \quad z = -t + 3r. \quad (3.2)$$

- **Piano di equazione** $ax + by + cz = d$ con $d \neq 0$

L'equazione $3x + 2y + z = 4$ descrive un piano parallelo al piano $3x + 2y + z = 0$. Se $R = [1, 1, -1]$ è un punto del piano $3x + 2y + z = 4$, allora i punti del nuovo piano si ottengono parametricamente da (3.2) come segue:

$$x' = 1 + t - r; \quad y = 1 - t; \quad z = -1 - t + 3r.$$

- **Piano passante per tre punti non allineati**

Siano P, Q, R tre punti non allineati (cioé, per i tre punti non passa una retta). Consideriamo i vettori applicati \overrightarrow{PQ} e \overrightarrow{PR} . Li riportiamo all'origine degli assi, considerando il vettore $Q - P$ ed il vettore $R - P$. Questi due vettori non sono allineati per l'ipotesi iniziale che i tre punti non sono allineati. Il piano passante per l'origine e contenente i vettori $Q - P$ e $R - P$ ha la seguente equazione parametrica:

$$t(Q - P) + r(R - P), \text{ al variare di } r, t \text{ numeri reali,}$$

Allora il piano passante per P, Q, R è descritto da

$$P + t(Q - P) + r(R - P), \text{ al variare di } r, t \text{ numeri reali,}$$

Infatti, ponendo $r = t = 0$ si ottiene il punto P , con $t = 1$ e $r = 0$ si ottiene Q , ed infine con $t = 0$ e $r = 1$ si ha il punto R .

Esempio: $P = [1, 1, 1]$, $Q = [1, 2, 1]$ ed $R = [5, 0, 7]$. Il piano passante per l'origine ha equazione parametrica:

$$x = 4r; \quad y = t - r; \quad z = 6r.$$

Il piano passante per i tre punti ha equazione parametrica:

$$x = 1 + 4r; \quad y = 1 + t - r; \quad z = 1 + 6r.$$

Trasformiamo l'equazione parametrica in una non parametrica: $r = (x - 1)/4$, da cui $z = 1 + 6(x - 1)/4 = 1 + 3(x - 1)/2$. Infine si ha $2z = 3x - 3 + 2 = 3x - 1$:

$$-3x + 2z = -1$$

3.2.1 Fasci di rette

L'insieme di tutte le rette del piano che passano per un punto C del piano prende il nome di fascio di rette che ha come centro il punto C .

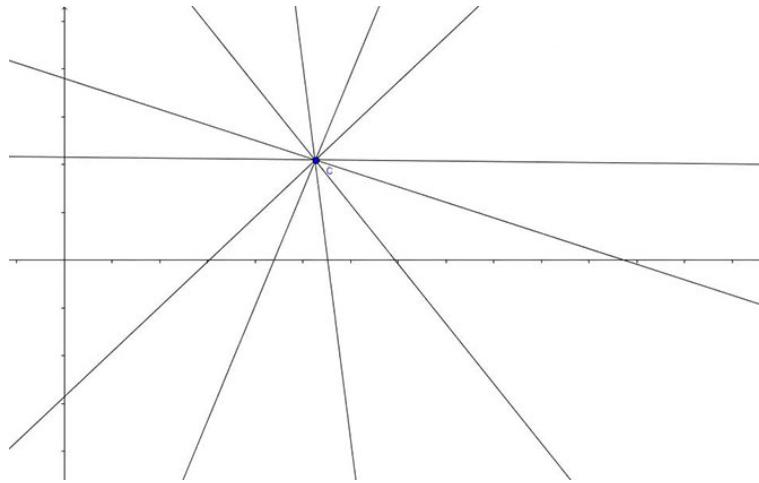


Figure 3.2: Fascio di rette di centro C

Se conosciamo le equazioni $ax + by + c = 0$ e $dx + ey + f = 0$ di due rette distinte passanti per C allora ogni altra retta nel fascio si scrive come combinazione lineare di queste due rette. L'equazione lineare

$$r(ax + by + c) + s(dx + ey + f) = 0, \quad (r, s \in \mathbb{R})$$

rappresenta tutte e sole le rette che passano per C se gli scalari r ed s non sono entrambi nulli. Se $s \neq 0$, possiamo dividere per s ed ottenere ponendo $t = r/s$:

$$t(ax + by + c) + (dx + ey + f) = 0.$$

Quest'ultima equazione descrive tutte le rette passanti per C tranne la retta $ax + by + c = 0$. Quest'ultima forma di descrizione del fascio è utile nelle applicazioni come il seguente esempio spiega.

Example 8. Vogliamo scrivere l'equazione della retta passante per i punti $P = [1, 1]$ e $Q = [2, 5]$. Consideriamo due rette distinte passanti per P , per esempio $y - x = 0$ e $x - 1 = 0$. Allora, il fascio $(y - x) + t(x - 1) = 0$ comprende tutte le rette passanti per P tranne la retta $x - 1 = 0$. Sostituendo le coordinate di Q al posto di x ed y otteniamo: $3 + t = 0$ da cui $t = -3$.

Quindi l'equazione $(y - x) - 3(x - 1) = 0$, che si semplifica a $y - 4x + 3 = 0$, descrive la retta passante per P e Q .

Chapter 4

Sistemi lineari

Un sistema lineare di m equazioni nelle incognite x_1, \dots, x_n è un insieme di equazioni lineari con coefficienti in un campo numerico \mathbb{K} :

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + \cdots + a_{2n}x_n &= b_2 \\ \cdots &\quad \cdots \quad \cdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

Il sistema lineare è omogeneo se $b_i = 0$ per ogni i . Una soluzione del sistema è un vettore $[r_1, \dots, r_n]$ di elementi del campo numerico tali che

$$\begin{aligned} a_{11}r_1 + \cdots + a_{1n}r_n &= b_1 \\ a_{21}r_1 + \cdots + a_{2n}r_n &= b_2 \\ \cdots &\quad \cdots \quad \cdots \\ a_{m1}r_1 + \cdots + a_{mn}r_n &= b_m \end{aligned}$$

Ciascuna equazione lineare del sistema descrive un iperpiano nello spazio \mathbb{K}^n . L'insieme delle soluzioni del sistema lineare descrive l'insieme dei punti che costituiscono l'intersezione di m iperpiani nello spazio \mathbb{K}^n . Ciascun iperpiano è descritto dall'equazione $a_{i1}x_1 + \cdots + a_{in}x_n = b_i$ ($i = 1, \dots, m$).

In dimensione 2 con $\mathbb{K} = \mathbb{R}$ le soluzioni del sistema

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= b_1 \\ a_{21}x_1 + a_{22}x_2 &= b_2 \end{aligned}$$

sono i punti $[x_1, x_2]$ del piano che costituiscono l'intersezione delle due rette di equazione rispettivamente $a_{11}x_1 + a_{12}x_2 = b_1$ e $a_{21}x_1 + a_{22}x_2 = b_2$. In

dimensione 3 le soluzioni del sistema

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 &= b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 &= b_3 \end{aligned}$$

sono i punti $[x_1, x_2, x_3]$ dello spazio che costituiscono l'intersezione dei tre piani.

4.1 Metodo di eliminazione di Gauss

Gauss, nato a Brunswick nel 1777 e morto a Göttingen nel 1855, introdusse un algoritmo per semplificare un sistema lineare trasformandolo in un sistema lineare più semplice con le stesse soluzioni del sistema iniziale.

Theorem 4.1.1. (Metodo di eliminazione di Gauss) *Se un sistema lineare è ottenuto da un altro con una delle seguenti operazioni:*

- (i) *Scambio di due equazioni;*
- (ii) *Sostituzione di un'equazione con un multiplo scalare non nullo dell'equazione stessa;*
- (iii) *Sostituzione di un'equazione con la somma dell'equazione stessa con un'altra moltiplicata per uno scalare d ,*

allora il sistema di partenza e quello di arrivo hanno le stesse soluzioni.

Per esempio, se $ax + by = c$ è l'equazione di una retta, moltiplicando per una stessa costante $d \neq 0$ entrambi i membri dell'uguaglianza si ottiene un'equazione

$$d(ax + by) = dc$$

che descrive la stessa retta. In termini di prodotto interno, se $[a, b] \cdot [x, y] = c$ allora $[da, db] \cdot [x, y] = dc$. Il vettore $[da, db]$ sta nella stessa retta contenente il vettore $[a, b]$. Quindi in un sistema lineare possiamo sostituire l'equazione $ax + by = c$ con l'equazione $dax + dby = dc$ senza che cambino le soluzioni del sistema.

Come ulteriore esempio, consideriamo il sistema lineare:

$$\begin{aligned} a_1x + b_1y &= c_1 \\ a_2x + b_2y &= c_2 \end{aligned} \tag{4.1}$$

Possiamo sostituire la prima equazione con la somma dell'equazione stessa con la seconda moltiplicata per una costante d senza modificare l'insieme delle soluzioni del sistema. In altre parole, sostituiamo l'equazione $a_1x + b_1y = c_1$ con l'equazione $(a_1x + b_1y) + d(a_2x + b_2y) = c_1 + dc_2$, ottenendo il nuovo sistema lineare:

$$\begin{aligned} (a_1 + da_2)x + (b_1 + db_2)y &= c_1 + dc_2 \\ a_2x + b_2y &= c_2 \end{aligned} \quad (4.2)$$

Proviamo che i due sistemi (4.1) e (4.2) hanno le stesse soluzioni. Sia $[p, q]$ una soluzione del sistema (4.2). Allora abbiamo direttamente da (4.2) che $a_2p + b_2q = c_2$. Resta da verificare che $a_1p + b_1q = c_1$. Da $(a_1 + da_2)p + (b_1 + db_2)q = c_1 + dc_2$ si ricava che $a_1p + b_1q = c_1 - d(a_2p + b_2q - c_2) = c_1 - d(c_2 - c_2) = c_1 + 0 = c_1$. È semplice verificare che una soluzione di (4.1) è anche una soluzione di (4.2).

4.2 Due equazioni in due variabili

- Consideriamo il sistema seguente:

$$\begin{aligned} 3x + 2y &= 5 \\ x + 4y &= 3 \end{aligned}$$

La prima è l'equazione di una retta parallela alla retta di equazione $3x + 2y = 0$, passante per l'origine e perpendicolare al vettore $A = [3, 2]$. La seconda è l'equazione di una retta parallela alla retta di equazione $x + 4y = 0$, passante per l'origine e perpendicolare al vettore $B = [1, 4]$.

Cerchiamo le soluzioni comuni alle due equazioni: un vettore $[x, y]$ il cui prodotto interno con $[3, 2]$ dà come risultato 5; e con $[1, 4]$ dà come risultato 3.

- **Numero di possibili soluzioni in generale.** Il sistema ammette un'unica soluzione se le due rette non sono parallele. Se le rette sono parallele, allora non abbiamo soluzioni se le rette sono distinte, mentre abbiamo infinite soluzioni se le due rette coincidono.

- **Numero delle soluzioni nel caso particolare.**

Per sapere se le rette sono parallele oppure no, calcoliamo il prodotto interno di A e B (si veda Sezione 2.3):

$$A \cdot B = [3, 2] \cdot [1, 4] = 3 + 8 = 11.$$

Siccome $\|A\| = \sqrt{13}$ e $\|B\| = \sqrt{17}$, abbiamo per l'angolo tra i vettori A e B :

$$\cos \theta = \frac{A \cdot B}{\sqrt{13}\sqrt{17}} = \frac{11}{\sqrt{221}} = \sqrt{\frac{121}{221}} < 1.$$

Quindi le rette non sono parallele. Se lo fossero $\cos \theta$ sarebbe uguale a +1 oppure -1.

- ***Calcolo dell'unica soluzione con il metodo di Gauss.***

Calcoliamo l'unica soluzione del sistema

$$\begin{aligned} 3x + 2y &= 5 \\ x + 4y &= 3 \end{aligned}$$

L'idea è di trasformare il sistema lineare dato in uno equivalente in cui la variabile x non compare nella seconda equazione.

Moltiplichiamo la prima equazione per $1/3$ ed otteniamo il nuovo sistema

$$\begin{aligned} x + \frac{2}{3}y &= \frac{5}{3} \\ x + 4y &= 3 \end{aligned}$$

La retta descritta dall'equazione $x + \frac{2}{3}y = \frac{5}{3}$ è la stessa descritta dall'equazione $3x + 2y = 5$. A questo punto la seconda equazione $x + 4y = 3$ e la nuova equazione $x + \frac{2}{3}y = \frac{5}{3}$ hanno lo stesso termine in x . Sottraiamo la prima equazione dalla seconda per ottenere il sistema

$$x + \frac{2}{3}y = \frac{5}{3}$$

$$\frac{10}{3}y = \frac{4}{3}$$

da cui $y = \frac{2}{5}$. Sostituendo nella prima equazione si ottiene $x = \frac{7}{5}$. La soluzione è quindi

$$x = \frac{7}{5}; \quad y = \frac{2}{5}.$$

- ***Un altro esempio con il metodo di eliminazione di Gauss.***

Consideriamo un altro esempio:

$$\begin{aligned} x + 2y &= 3 \\ 4x + 5y &= 6 \end{aligned}$$

Calcoliamo il numero delle soluzioni. Sia $A = [1, 2]$ e $B = [4, 5]$. Allora abbiamo:

$$\cos \theta = \frac{A \cdot B}{\|A\| \|B\|} = \frac{14}{\sqrt{5}\sqrt{41}} = \frac{14}{\sqrt{205}} = \sqrt{\frac{196}{221}} < 1.$$

Quindi le due rette non sono parallele ed esiste un'unica soluzione.

Calcoliamo la soluzione. Moltiplichiamo la prima equazione per 4 e la sottraiamo alla seconda. Si ottiene il sistema:

$$\begin{aligned} x + 2y &= 3 \\ -3y &= -6 \end{aligned}$$

da cui si ricava $y = 2$ e, sostituendo 2 per y nella prima equazione, si ottiene $x = -1$.

- ***Un esempio di sistema con infinite soluzioni.***

Consideriamo il seguente sistema lineare:

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

Calcoliamo il numero delle soluzioni. Sia $A = [1, 2]$ e $B = [2, 4]$. Allora abbiamo:

$$\cos \theta = \frac{A \cdot B}{\|A\| \|B\|} = \frac{10}{\sqrt{5}\sqrt{20}} = \frac{10}{\sqrt{100}} = \frac{10}{10} = 1.$$

Quindi le due rette sono parallele, anzi sono uguali. La seconda equazione si scrive come $2(x + 2y) = 2 \times 3$. Quindi il sistema lineare ammette infinite soluzioni.

- ***Un esempio di sistema con nessuna soluzione.*** Consideriamo il seguente sistema lineare:

$$\begin{aligned} x + 2y &= 3 \\ 2x + 4y &= 12 \end{aligned}$$

Calcoliamo il numero delle soluzioni. Sia $A = [1, 2]$ e $B = [2, 4]$. Allora come prima abbiamo $\cos \theta = 1$. Quindi le due rette sono parallele, ma non uguali. La seconda equazione si scrive come $2(x + 2y) = 2 \times 6$ e corrisponde alla retta $x + 2y = 6$ che è parallela alla retta della prima equazione. Quindi il sistema lineare non ammette soluzione.

4.3 Due o tre equazioni in tre variabili

1. Consideriamo il sistema lineare:

$$\begin{aligned} 3x + 2y + z &= 0 \\ x + 4y + 3z &= 0 \\ x + y + z &= 0 \end{aligned}$$

Il sistema è omogeneo perché il vettore dei termini noti è il vettore nullo.

La prima è l'equazione di un piano passante per l'origine e perpendicolare al vettore $A = [3, 2, 1]$. La seconda è l'equazione di un piano passante per l'origine e perpendicolare al vettore $B = [1, 4, 3]$. La terza è l'equazione di un piano passante per l'origine e perpendicolare al vettore $C = [1, 1, 1]$. Quindi una soluzione comune $[x, y, z]$ delle tre equazioni è un vettore contemporaneamente perpendicolare ad $A = [3, 2, 1]$, $B = [1, 4, 3]$ e $C = [1, 1, 1]$.

- (i) L'intersezione dei tre piani è un piano sse i tre piani sono coincidenti (lo stesso piano) sse i tre vettori A , B e C sono collineari (nella stessa retta per l'origine);
- (ii) L'intersezione dei tre piani è una retta sse i tre vettori A , B e C non sono collineari ma si trovano in uno stesso piano passante per l'origine;
- (iii) L'intersezione dei tre piani è un punto sse (1) i vettori A e B non si trovano nella stessa retta; (2) il vettore C non si trova nel piano generato dai vettori A e B .

Ritorniamo al sistema lineare omogeneo precedente. Calcoliamo se i piani sono paralleli. Siano $A = [3, 2, 1]$, $B = [1, 4, 3]$ e $C = [1, 1, 1]$. Allora abbiamo:

$$\cos(\theta_{AB}) = (A \cdot B) / \|A\| \|B\| = 14 / \sqrt{14} \sqrt{26} = 14 / \sqrt{364} < 1.$$

Quindi A e B non sono collineari. Se C fosse nel piano generato da A e B , allora esisterebbero due scalari c e d tali che $c[3, 2, 1] + d[1, 4, 3] = [1, 1, 1]$, da cui si ha $3c + d = 1$, $2c + 4d = 1$ e $c + 3d = 1$. Sottraendo la seconda dalla prima si ricava: $c = 3d$. Sostituendo nella terza si

ottiene $3d + 3d = 0$, cioè $d = 0$ e quindi $c = 0$. Si ottiene così che C non si trova nel piano generato da A e B . Quindi l'unica soluzione del sistema lineare iniziale è il vettore nullo.

2. Consideriamo il sistema seguente:

$$\begin{aligned} 3x + 2y + z &= 5 \\ x + 4y + 3z &= 3 \\ x + y + z &= 0 \end{aligned}$$

La prima è l'equazione di un piano parallelo al piano di equazione $3x + 2y + z = 0$ passante per l'origine e perpendicolare al vettore $A = [3, 2, 1]$. La seconda è l'equazione di un piano parallelo al piano di equazione $x + 4y + 3z = 0$ passante per l'origine e perpendicolare al vettore $B = [1, 4, 3]$. La terza è l'equazione del piano passante per l'origine e perpendicolare al vettore $C = [1, 1, 1]$.

L'intersezione dei tre piani è un punto se (1) i vettori A e B non si trovano nella stessa retta; (2) il vettore C non si trova nel piano generato dai vettori A e B . Dai fatti calcoli fatti nel punto precedente per il sistema omogeno si ottiene che C non si trova nel piano generato da A e B . Quindi la soluzione è un punto. Applichiamo il metodo di eliminazione di Gauss.

Scambiamo la prima equazione con la terza. Si ottiene il sistema:

$$\begin{aligned} x + y + z &= 0 \\ 3x + 2y + z &= 5 \\ x + 4y + 3z &= 3 \end{aligned}$$

Facciamo scomparire x dalla seconda e terza equazione.

- Sottraiamo 3 volte la prima equazione dalla seconda;
- Sottraiamo la prima equazione dalla terza;

Si ottiene:

$$\begin{aligned} x &+ y &+ z &= 0 \\ -y &- 2z &= 5 \\ 3y &+ 2z &= 3 \end{aligned}$$

Facciamo scomparire y dalla terza equazione sommando tre volte la seconda:

$$\begin{array}{rcl} x & + & y & + & z & = & 0 \\ & -y & - & 2z & = & 5 \\ & & & -4z & = & 18 \end{array}$$

Quindi $z = -9/2$, $y = 4$ e infine dalla prima equazione $x = -4 + 9/2 = 1/2$.

3. Consideriamo il sistema seguente:

$$\begin{array}{rcl} 3x + 2y + z & = & 5 \\ x + 4y + 3z & = & 3 \end{array}$$

La prima è l'equazione di un piano parallelo al piano di equazione $3x + 2y + z = 0$ passante per l'origine e perpendicolare al vettore $A = [3, 2, 1]$. La seconda è l'equazione di un piano parallelo al piano di equazione $x + 4y + 3z = 0$ passante per l'origine e perpendicolare al vettore $B = [1, 4, 3]$.

L'intersezione di due piani, se non paralleli, è una retta.

Cerchiamo le soluzioni comuni alle due equazioni: un vettore $[x, y, z]$ il cui prodotto interno con $[3, 2, 1]$ dà come risultato 5; e con $[1, 4, 3]$ dà come risultato 3.

Calcoliamo se i piani sono paralleli. Sia $A = [3, 2, 1]$ e $B = [1, 4, 3]$. Allora abbiamo:

$$\cos \theta = (A \cdot B) / \|A\| \|B\| = 14 / \sqrt{14} \sqrt{26} = 14 / \sqrt{364} < 1.$$

Quindi i due piani non sono paralleli.

Applichiamo il metodo di eliminazione di Gauss. Scambiamo la prima e la seconda equazione ottenendo

$$\begin{array}{rcl} x + 4y + 3z & = & 3 \\ 3x + 2y + z & = & 5 \end{array}$$

Sottraiamo alla seconda equazione tre volte la prima equazione ottenendo: $x + 4(2-4z)/5 + 3z = 3$

$$\begin{array}{rcl} x & + & 4y & + & 3z & = & 3 \\ & -10y & - & 8z & = & -4 \end{array}$$

da cui si ha $y = \frac{2-4z}{5}$. Sostituiamo nella prima equazione per ottenere le soluzioni in termini del parametro z :

$$x = (3 - (16/5))z + (3 - (8/5)); \quad y = -\frac{2-4z}{5}.$$

Chapter 5

Matrici

Consideriamo nuovamente il sistema lineare:

$$\begin{aligned} 3x + 2y + z &= 5 \\ x + 4y + 3z &= 3 \\ x + y + z &= 0 \end{aligned}$$

Il sistema lineare è completamente determinato dalla seguente tabella bidimensionale con i coefficienti delle variabili

$$\left[\begin{array}{ccc} 3 & 2 & 1 \\ 1 & 4 & 3 \\ 1 & 1 & 1 \end{array} \right]$$

e dal vettore dei termini noti

$$\left[\begin{array}{c} 5 \\ 3 \\ 0 \end{array} \right]$$

Sia i coefficienti delle variabili che il vettore dei termini noti costituiscono delle matrici. La prima ha dimensione 3×3 , la seconda ha dimensione 3×1 . La matrice completa del sistema si ottiene aggiungendo il vettore dei termini noti alla matrice dei coefficienti:

$$\left[\begin{array}{ccc|c} 3 & 2 & 1 & 5 \\ 1 & 4 & 3 & 3 \\ 1 & 1 & 1 & 0 \end{array} \right]$$

In questa sezione introduciamo le matrici in generale e studiamo le loro proprietà algebriche.

5.1 Definizione di matrice

Cominciamo con l'introdurre il concetto di matrice in generale.

Definition 5.1.1. Una *matrice* $A = (a_{ij})$ con m righe ed n colonne (in breve una matrice di tipo $m \times n$) sul campo \mathbb{K} è una famiglia di mn elementi di \mathbb{K} . Ciascun elemento ha un indice di riga ed un indice di colonna. L'elemento a_{ij} ha indice di riga i ed indice di colonna j con $1 \leq i \leq m$ e $1 \leq j \leq n$. La matrice si rappresenta come segue:

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

Se A è una matrice di tipo $m \times n$, denotiamo con A_i il vettore che è la riga i della matrice, e con A^j il vettore che è la colonna j della matrice:

$$A_i = [a_{i1} \ a_{i2} \ \dots \ a_{in}] ; \quad A^j = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \dots \\ a_{mj} \end{bmatrix}$$

Una matrice è *quadrata* se il numero di righe è uguale al numero di colonne. Una matrice quadrata di tipo $n \times n$ viene detta di *ordine n* .

Una matrice quadrata A di ordine n è

1. *simmetrica* se $a_{ij} = a_{ji}$ per ogni $1 \leq i, j \leq n$.
2. *diagonale* se $a_{ij} = 0$ per $i \neq j$.
3. la *matrice identica* se è diagonale e $a_{ii} = 1$ per ogni i (la matrice identica di ordine n si indica con I_n).
4. *triangolare superiore (inferiore)* se $a_{ij} = 0$ per ogni $i > j$ ($i < j$).
5. la *matrice nulla* se $a_{ij} = 0$ per ogni $1 \leq i, j \leq n$.

La *trasposta* di una matrice $A = (a_{ij})$ è una matrice $A^t = (b_{ij})$ le cui righe sono le colonne di A e le cui colonne sono le righe di A . In altre parole, $b_{ij} = a_{ji}$ per ogni i e j . Se A è una matrice di tipo $m \times n$ allora la sua trasposta è una matrice di tipo $n \times m$.

Proposition 5.1.1. *Sia A una matrice. Allora si ha:*

1. $(A^t)^t = A$.
2. Se A è simmetrica e quadrata allora $A^t = A$.

Data una matrice A di tipo $m \times n$, gli elementi a_{ik} con $i = k$ si chiamano *elementi principali* o elementi appartenenti alla diagonale principale. La loro somma si chiama *traccia della matrice*, si indica con $\text{tr}A$ e si ha $\text{tr}A = a_{11} + a_{22} + \dots + a_{mm}$.

Example 9. Le seguenti matrici sono rispettivamente simmetrica, diagonale e la matrice identica:

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

Il vettore riga $A_2 = [2 \ 4 \ 4]$ coincide con la seconda riga della matrice A ,

mentre il vettore colonna $A^3 = \begin{bmatrix} 1 \\ 4 \\ 2 \end{bmatrix}$ corrisponde alla terza colonna della matrice A . La traccia della matrice A è $\text{tr}A = 3 + 4 + 2 = 9$.

La trasposta della matrice simmetrica A coincide con A , mentre la matrice C qui di seguito non è simmetrica e la sua trasposta non coincide con C .

$$C = \begin{bmatrix} 3 & 2 & 7 \\ 1 & 4 & 9 \\ 22 & 1 & 0 \end{bmatrix} \quad C^t = \begin{bmatrix} 3 & 1 & 22 \\ 2 & 4 & 1 \\ 7 & 9 & 0 \end{bmatrix}$$

La seguente matrice D è triangolare superiore:

$$D = \begin{bmatrix} 3 & 2 & 7 \\ 0 & 4 & 9 \\ 0 & 0 & 5 \end{bmatrix}$$

5.2 Operazioni su Matrici

L'insieme delle matrici di tipo $m \times n$ a coefficienti in un campo \mathbb{K} costituisce uno spazio vettoriale rispetto all'operazioni di somma componente per componente e prodotto per uno scalare. Se $A = (a_{ij})$ e $B = (b_{ij})$ sono matrici di

tipo $m \times n$, allora la matrice $C = (c_{ij})$ è la somma di A e B se:

$$c_{ij} = a_{ij} + b_{ij}, \quad \text{per ogni } i, j.$$

Se $r \in \mathbb{K}$ è uno scalare, allora rA è la matrice così definita:

$$(rA)_{ij} = rA_{ij}, \quad \text{per ogni } i, j.$$

Limitandoci alle matrici di tipo 3×3 , in modo più espressivo possiamo scrivere:

$$\begin{aligned} A+B &= \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix} = \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} & a_{13} + b_{13} \\ a_{21} + b_{21} & a_{22} + b_{22} & a_{23} + b_{23} \\ a_{31} + b_{31} & a_{32} + b_{32} & a_{33} + b_{33} \end{bmatrix} \\ cA &= r \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} ra_{11} & ra_{12} & ra_{13} \\ ra_{21} & ra_{22} & ra_{23} \\ ra_{31} & ra_{32} & ra_{33} \end{bmatrix} \end{aligned}$$

L'opposta della matrice A è la matrice $-A$ le cui componenti sono gli elementi $-a_{ij}$.

Example 10.

$$\begin{aligned} \begin{bmatrix} 2 & 3 \\ 1 & -5 \\ 3 & 2 \end{bmatrix} + \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} &= \begin{bmatrix} 3 & 5 \\ 4 & -1 \\ 8 & 8 \end{bmatrix} \quad 4 \begin{bmatrix} 2 & 3 \\ 1 & -5 \\ 3 & 2 \end{bmatrix} = \begin{bmatrix} 8 & 12 \\ 4 & -20 \\ 12 & 8 \end{bmatrix} \\ A &= \begin{bmatrix} 2 & 3 \\ 1 & -5 \\ 3 & 2 \end{bmatrix}; \quad -A = \begin{bmatrix} -2 & -3 \\ -1 & 5 \\ -3 & -2 \end{bmatrix} \end{aligned}$$

5.2.1 Prodotto di una matrice per un vettore

Consideriamo il sistema lineare a coefficienti reali:

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + \cdots + a_{2n}x_n &= b_2 \\ \cdots &\quad \cdots \quad \cdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

Definiamo la matrice A dei coefficienti del sistema ed i vettori colonna delle incognite e dei termini noti:

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

Allora si vede facilmente che la prima equazione lineare $a_{11}x_1 + \dots + a_{1n}x_n = b_1$ si ottiene prendendo il prodotto interno del vettore riga

$$A_1 = [a_{11} \ a_{12} \ \dots \ a_{1n}] \text{ per il vettore colonna } \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{bmatrix} \text{ e ponendolo}$$

uguale a b_1 . Similmente per le altre equazioni lineari. I prodotti interni di questo tipo si rappresentano con il prodotto di matrici:

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + \dots + a_{1n}x_n \\ a_{21}x_1 + \dots + a_{2n}x_n \\ \dots \\ a_{m1}x_1 + \dots + a_{mn}x_n \end{bmatrix}$$

Quindi il sistema lineare si rappresenta globalmente come segue

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

La precedente equazione matriciale si può anche scrivere come combinazione lineare di vettori colonna:

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

Cosa significa questa equazione? Ci chiediamo se il vettore $\begin{bmatrix} b_1 \\ b_2 \\ \dots \\ b_m \end{bmatrix}$ si può

scrivere come combinazione lineare dei vettori colonna $\begin{bmatrix} a_{11} \\ a_{21} \\ \dots \\ a_{m1} \end{bmatrix}, \dots, \begin{bmatrix} a_{1n} \\ a_{2n} \\ \dots \\ a_{mn} \end{bmatrix}$.

Come vedremo nel Capitolo 7, l'equazione vettoriale ha sicuramente soluzione se i vettori colonna sono "linearmente indipendenti".

Example 11. La matrice quadrata dei coefficienti del sistema lineare

$$\begin{aligned} 3x + 2y &= 5 \\ x + 4y &= 3 \end{aligned}$$

ha ordine 2:

$$\begin{bmatrix} 3 & 2 \\ 1 & 4 \end{bmatrix}$$

Il sistema lineare si può scrivere in notazione matriciale come segue:

$$\begin{bmatrix} 3 & 2 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 5 \\ 3 \end{bmatrix}$$

I due vettori $[x, y]$ e $[5, 3]$ sono stati scritti come vettori colonna, cioè come matrici 2×1 .

Il prodotto interno $3x+2y$ del vettore $[3, 2]$ per il vettore $\begin{bmatrix} x \\ y \end{bmatrix}$ è il prodotto interno della prima riga della matrice per l'unica colonna del vettore colonna $\begin{bmatrix} x \\ y \end{bmatrix}$. Allo stesso modo il prodotto interno $x+4y$ del vettore $[1, 4]$ per il vettore $\begin{bmatrix} x \\ y \end{bmatrix}$ è il prodotto interno della seconda riga della matrice per l'unica colonna del vettore colonna $\begin{bmatrix} x \\ y \end{bmatrix}$. Quindi abbiamo:

$$\begin{bmatrix} 3 & 2 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 3x + 2y \\ x + 4y \end{bmatrix}$$

Example 12. Il sistema

$$\begin{aligned} 3x + 2y + z &= 5 \\ x + 4y + 3z &= 3 \\ x + y + z &= 0 \end{aligned}$$

si rappresenta in notazione matriciale come segue:

$$\begin{bmatrix} 3 & 2 & 1 \\ 1 & 4 & 3 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 5 \\ 3 \\ 0 \end{bmatrix}$$

Come prima facciamo il prodotto interno tra la prima (rispettivamente seconda e terza) riga della matrice per l'unica colonna del vettore $\begin{bmatrix} x \\ y \\ z \end{bmatrix}$ per ottenere

$$\begin{bmatrix} 3 & 2 & 1 \\ 1 & 4 & 3 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 3x + 2y + z \\ x + 4y + 3z \\ x + y + z \end{bmatrix} = \begin{bmatrix} 5 \\ 3 \\ 0 \end{bmatrix}$$

La precedente equazione matriciale si può anche scrivere come combinazione lineare di vettori colonna:

$$x \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix} + y \begin{bmatrix} 2 \\ 4 \\ 1 \end{bmatrix} + z \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ 3 \\ 0 \end{bmatrix} = b$$

con b il vettore colonna $[5, 3, 0]^t$. Ci chiediamo se il vettore b si può scrivere

come combinazione lineare dei vettori $\begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}$, $\begin{bmatrix} 2 \\ 4 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix}$.

5.2.2 Prodotto di matrici

Tutte le trasformazioni sulle matrici della precedente sezione si ottengono anche utilizzando il prodotto matriciale che ci accingiamo a definire.

Definition 5.2.1. Siano $A = (a_{ij})$ una matrice di tipo $m \times k$ e $B = (b_{ij})$ una matrice di tipo $k \times n$. Il *prodotto* AB di A e B è una matrice C di tipo $m \times n$ le cui componenti c_{ij} sono ottenute come segue (per ogni i e j):

$$c_{ij} = \sum_{r=1}^k a_{ir} b_{rj}.$$

Osserviamo che sia la riga i di A che la colonna j di B hanno k elementi.

Proposition 5.2.1. Siano $A = (a_{ij})$ una matrice di tipo $m \times k$, $B = (b_{ij})$ una matrice di tipo $k \times n$ e $C = (c_{ij})$ una matrice di tipo $n \times r$. Il prodotto tra matrici è associativo:

$$A(BC) = (AB)C.$$

L'elemento neutro della matrice A è a sinistra la matrice identica di ordine m ed a destra la matrice identica di ordine k :

$$I_m A = A = A I_k.$$

Il prodotto distribuisce rispetto alla somma:

$$A(B + C) = (AB) + (AC); \quad (B + C)A = (BA) + (CA).$$

Il prodotto non è in generale commutativo. Esistono matrici A e B tali che $AB \neq BA$. Abbiamo inoltre per le matrici trasposte:

$$(AB)^t = B^t A^t.$$

Infine se r è uno scalare:

$$(rA)B = A(rB) = r(AB).$$

Se $\mathbf{a} = a_1\mathbf{e}_1 + a_2\mathbf{e}_2 + a_3\mathbf{e}_3$ e $\mathbf{b} = b_1\mathbf{e}_1 + b_2\mathbf{e}_2 + b_3\mathbf{e}_3$ sono vettori dello spazio tridimensionale, le cui coordinate sono espresse rispetto alla base canonica, allora il prodotto interno di \mathbf{a} e \mathbf{b} si può scrivere con il prodotto matriciale: $[a_1, a_2, a_3][b_1, b_2, b_3]^t = a_1b_1 + \dots + a_n b_n$. Questo risultato non vale in generale se consideriamo le coordinate dei vettori \mathbf{a} e \mathbf{b} rispetto ad una base arbitraria.

Example 13. In questo esempio consideriamo tre matrici:

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

e controlliamo che $(AB)C = A(BC)$. Abbiamo:

$$\begin{aligned} AB &= \begin{bmatrix} 7 & -2 & 13 \\ 18 & 2 & 32 \end{bmatrix} & (AB)C &= \begin{bmatrix} 41 & -36 & 44 \\ 74 & -104 & 116 \end{bmatrix} \\ BC &= \begin{bmatrix} 5 & -20 & 20 \\ 18 & -8 & 12 \end{bmatrix} & A(BC) &= \begin{bmatrix} 41 & -36 & 44 \\ 74 & -104 & 116 \end{bmatrix} \end{aligned}$$

Example 14. In questo esempio verifichiamo che il prodotto non è commutativo:

$$\begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix} \begin{bmatrix} 3 & 2 \\ 2 & -2 \end{bmatrix} = \begin{bmatrix} 7 & -2 \\ 18 & 2 \end{bmatrix}$$

mentre

$$\begin{bmatrix} 3 & 2 \\ 2 & -2 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix} = \begin{bmatrix} 11 & 12 \\ -6 & -2 \end{bmatrix}$$

Example 15.

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

Esempi

- $\begin{bmatrix} 2 & 5 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} -2 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \cdot -2 + 5 \cdot 1 \\ 3 \cdot -2 + 1 \cdot 1 \end{bmatrix} = \begin{bmatrix} 1 \\ -5 \end{bmatrix}$
- $\begin{bmatrix} 1 & -2 \\ 4 & 7 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 5 & 6 \end{bmatrix} = \begin{bmatrix} 1 \cdot 2 + (-2) \cdot 5 & 1 \cdot 1 + (-2) \cdot 6 \\ 4 \cdot 2 + 7 \cdot 5 & 4 \cdot 1 + 7 \cdot 6 \end{bmatrix} = \begin{bmatrix} -8 & -11 \\ 43 & 46 \end{bmatrix}$
- $\begin{bmatrix} 2 & 1 \\ 5 & 6 \end{bmatrix} \begin{bmatrix} 1 & -2 \\ 4 & 7 \end{bmatrix} = \begin{bmatrix} 2 \cdot 1 + 1 \cdot 4 & 2 \cdot (-2) + 1 \cdot 7 \\ 5 \cdot 1 + 6 \cdot 4 & 5 \cdot (-2) + 6 \cdot 7 \end{bmatrix} = \begin{bmatrix} 6 & 3 \\ 29 & 32 \end{bmatrix}$
- $\begin{bmatrix} 2 & 3 & 6 \\ 7 & 4 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 & 2 & 1 \\ 3 & 4 & 5 & 6 \\ 1 & 0 & 3 & 8 \end{bmatrix} = \begin{bmatrix} 2 \cdot 2 + 3 \cdot 3 + 6 \cdot 1 & 2 \cdot 0 + 3 \cdot 4 + 6 \cdot 0 & 2 \cdot 2 + 3 \cdot 5 + 6 \cdot 3 & 2 \cdot 1 + 3 \cdot 3 + 6 \cdot 8 \\ 7 \cdot 2 + 4 \cdot 3 + 1 \cdot 1 & 7 \cdot 0 + 4 \cdot 4 + 1 \cdot 0 & 7 \cdot 2 + 4 \cdot 5 + 1 \cdot 3 & 7 \cdot 1 + 4 \cdot 3 + 1 \cdot 8 \end{bmatrix} = \begin{bmatrix} 19 & 12 & 37 & 59 \\ 27 & 16 & 37 & 27 \end{bmatrix}$
- $\begin{bmatrix} 1 & 3 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} 2 & 3 & 1 & 1 \\ 0 & 2 & -1 & -4 \end{bmatrix} = \begin{bmatrix} 2 & 9 & -2 & -11 \\ -2 & -3 & -1 & -1 \end{bmatrix}$

Figure 5.1: Prodotto di Matrici

Example 16. Un grafo $G = (V, E)$ è costituito da un insieme finito $V = \{v_1, v_2, \dots, v_n\}$ di vertici (o nodi) e da un insieme di archi o frecce definite tramite una relazione binaria $E \subseteq V \times V$. Se $(v, u) \in E$ allora esiste un arco orientato che si diparte dal vertice v ed arriva al vertice u :

$$v \longrightarrow u.$$

Un cammino in un grafo è una sequenza di nodi u_0, u_1, \dots, u_k tali che $u_i \rightarrow u_{i+1}$ for every $0 \leq i < k$.

La matrice di adiacenza di un grafo G con n vertici è una matrice quadrata A di ordine n :

$$a_{ij} = \begin{cases} 1 & \text{se } (v_i, v_j) \in E \\ 0 & \text{altrimenti.} \end{cases}$$

La somma degli elementi della riga A_i è pari al numero di archi che escono dal vertice v_i .

Definiamo le potenze della matrice A per induzione come segue: $A^0 = I_n$ e $A^{k+1} = A^k A$ (da non confondersi con i vettori colonna di A). Indichiamo con $(a^k)_{ij}$ le componenti della matrice A^k . Proviamo per induzione su k che $(a^k)_{ij}$ è uguale al numero di cammini di lunghezza k dal nodo v_i al nodo v_j . Il risultato è vero per $A^1 = A$. Un cammino di lunghezza 1 da v_i a v_j è un arco orientato che connette v_i a v_j . L'arco esiste sse $a_{ij} = 1$ sse $(v_i, v_j) \in E$.

Supponiamo che il risultato sia vero per A^k e dimostriamolo per A^{k+1} :

$$(a^{k+1})_{ij} = \sum_{r=1}^n (a^k)_{ir} a_{rj}.$$

Infatti, un cammino di lunghezza $k+1$ da v_i a v_j lo possiamo spezzare come un cammino di lunghezza k da v_i ad un vertice intermedio v_r ed un arco da v_r a v_j . Se calcoliamo quanti sono questi cammini di lunghezza $k+1$ con nodo intermedio v_r , essi sono pari al numero $(a^k)_{ir}$ di cammini di lunghezza k da v_i a v_r se esiste un arco da v_r a v_j , oppure sono 0 se tale arco non esiste. In ogni caso è pari a

$$(a^k)_{ir} a_{rj}.$$

Ne segue la conclusione. Quindi, per ogni k , $(a^k)_{ij}$ è uguale al numero di cammini di lunghezza k da v_i a v_j .

Consideriamo, come esempio, la matrice binaria

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

Essa rappresenta un grafo con tre vertici v_1, v_2, v_3 ed i seguenti archi:

$$v_1 \rightarrow v_2; \quad v_1 \rightarrow v_3; \quad v_2 \rightarrow v_3; \quad v_3 \rightarrow v_1; \quad v_3 \rightarrow v_2$$

Si ha

$$A^2 = AA = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 2 \end{bmatrix}$$

Abbiamo in effetti due cammini di lunghezza due da v_3 in v_3 .

5.3 Moltiplicazione di matrici a blocchi

La moltiplicazione tra matrici si semplifica a volte se utilizziamo la moltiplicazione a blocchi.

Siano A e B matrici rispettivamente di tipo $m \times n$ e di tipo $n \times p$, e sia r un numero minore o uguale ad n . Possiamo decomporre le due matrici in blocchi:

$$A = [C|D]; \quad B = \left[\begin{array}{c} E \\ F \end{array} \right],$$

dove C è di tipo $m \times r$, D è di tipo $m \times (n - r)$, E è di tipo $r \times p$ e F è di tipo $(n - r) \times p$. Allora il prodotto matriciale può essere calcolato come segue:

$$AB = CE + DF.$$

Se dividiamo A e B in quattro blocchi

$$A = \begin{bmatrix} C & D \\ E & F \end{bmatrix}; \quad B = \begin{bmatrix} C' & D' \\ E' & F' \end{bmatrix},$$

allora la moltiplicazione matriciale si esegue come se A e B fossero matrici quadrate di ordine 2:

$$AB = \begin{bmatrix} CC' + DE' & CD' + DF' \\ EC' + FE' & ED' + FF' \end{bmatrix}.$$

Example 17. Siano $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \mid \begin{bmatrix} 5 \\ 3 \end{bmatrix}$ e $B = \begin{bmatrix} 2 & 3 \\ 4 & 8 \\ \hline 1 & 0 \end{bmatrix} \begin{array}{c|cc} 1 & 1 \\ 0 & 0 \end{array}$ due matrici suddivise in blocchi compatibili. Allora si ha:

- $[1, 0] \begin{bmatrix} 2 & 3 \\ 4 & 8 \end{bmatrix} + [5][1, 0] = [2, 3] + [5, 0] = [7, 3].$

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

Quindi

$$AB = \begin{bmatrix} 7 & 3 & 6 & 1 \\ 7 & 8 & 3 & 0 \end{bmatrix}.$$

Example 18. Siano $A = \begin{bmatrix} 1 & 0 & | & 0 \\ 0 & 0 & | & 3 \end{bmatrix}$ e $B = \begin{bmatrix} 0 & 0 & | & 1 & 1 \\ 0 & 0 & | & 0 & 0 \\ 1 & 0 & | & 1 & 0 \end{bmatrix}$ due matrici suddivise in blocchi compatibili. Siccome alcuni blocchi sono costituiti dalla matrice nulla, allora si vede facilmente che $AB = \begin{bmatrix} 0 & 0 & 1 & 1 \\ 3 & 0 & 3 & 0 \end{bmatrix}$.

Chapter 6

Matrici e sistemi lineari

Un sistema lineare di m equazioni in n incognite si scrive come

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

dove la matrice A dei coefficienti è di tipo $m \times n$, il vettore colonna \mathbf{x} delle incognite ha n componenti ed il vettore \mathbf{b} dei termini noti ha m componenti. La matrice completa $[A|\mathbf{b}]$ del sistema di tipo $m \times (n+1)$ rappresenta completamente il sistema.

Le *operazioni elementari* sulle righe della matrice completa $[A|\mathbf{b}]$, definite come:

- Scambio di due righe;
- Moltiplicazione di una riga per uno scalare $r \neq 0$;
- Sostituzione di una riga con la somma della riga stessa con un'altra riga moltiplicata per uno scalare.

determinano un processo di trasformazione tra matrici dello stesso tipo:

$$[A|\mathbf{b}] \Rightarrow [A'|\mathbf{b}'] \Rightarrow [A''|\mathbf{b}''] \Rightarrow \dots$$

Ciascuna di queste matrici complete rappresenta un sistema lineare che ha le stesse soluzioni del sistema di partenza $A\mathbf{x} = \mathbf{b}$. Per esempio, il sistema lineare $A''\mathbf{x} = \mathbf{b}''$ ha le stesse soluzioni di $A\mathbf{x} = \mathbf{b}$. Alla fine si ottiene una matrice a gradini che ci permette di ottenere facilmente le soluzioni del sistema di partenza.

Osserviamo che le operazioni elementari sono reversibili. Se applichiamo una delle operazioni elementari ad una riga di una matrice, ne possiamo sempre applicare un'altra che esegue l'operazione inversa. Per esempio, se moltiplichiamo una riga per lo scalare 3, allora se moltiplichiamo successivamente per $\frac{1}{3}$ riotteniamo la riga di partenza.

Due matrici sono equivalenti (per riga) se possiamo passare dall'una all'altra con operazioni elementari. Questa relazione è una relazione di equivalenza.

Alla fine di questa sezione proveremo che, se le matrici complete di due sistemi lineari sono equivalenti (per riga), allora i sistemi hanno le stesse soluzioni.

6.1 Matrici e metodo di eliminazione

Ritorniamo al sistema lineare

$$\begin{bmatrix} 3 & 2 & 1 \\ 1 & 4 & 3 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 5 \\ 3 \\ 0 \end{bmatrix}$$

e rivediamo i passi effettuati per ottenere la soluzione:

1. Scambia la prima riga con la terza riga;
2. Sottrai la prima equazione dalla seconda;
3. Sottrai 3 volte la prima equazione dalla terza;
4. Somma alla terza equazione un terzo della seconda equazione.

Consideriamo la matrice completa del sistema:

$$\begin{bmatrix} 3 & 2 & 1 & 5 \\ 1 & 4 & 3 & 3 \\ 1 & 1 & 1 & 0 \end{bmatrix}$$

Scriviamo qui di seguito le varie matrici che si ottengono con i vari passaggi

$$\begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 4 & 3 & 3 \\ 3 & 2 & 1 & 5 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 3 & 2 & 3 \\ 3 & 2 & 1 & 5 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 3 & 2 & 3 \\ 0 & -1 & -2 & 5 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 3 & 2 & 3 \\ 0 & 0 & -4/3 & 6 \end{bmatrix}$$

Dalla matrice finale si ottiene facilmente la soluzione.

Definition 6.1.1. Una matrice è *elementare* se ha uno dei seguenti tre formati:

1. Matrice di tipo I che moltiplicando a sinistra una matrice A effettua lo scambio di una riga di A con un'altra riga di A . Per semplicità, consideriamo matrici di tipo 3×3 .

$$E_{1,3} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad E_{2,3} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad E_{1,2} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

2. Matrice di tipo II che moltiplicando a sinistra una matrice A effettua la moltiplicazione di una riga di A per uno scalare r :

$$E_{r1} = \begin{bmatrix} r & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad E_{r2} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & r & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad E_{r3} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & r \end{bmatrix}$$

3. Matrice di tipo III che, moltiplicando a sinistra una matrice A , somma un multiplo di una data riga di A ad un'altra riga data:

$$\begin{aligned} E_{2+r1} &= \begin{bmatrix} 1 & 0 & 0 \\ r & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} & E_{3+r1} &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ r & 0 & 1 \end{bmatrix} & E_{3+r2} &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & r & 1 \end{bmatrix} \\ E_{1+r2} &= \begin{bmatrix} 1 & r & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} & E_{1+r3} &= \begin{bmatrix} 1 & 0 & r \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} & E_{2+r3} &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & r \\ 0 & 0 & 1 \end{bmatrix} \end{aligned}$$

Ritorniamo al sistema lineare prima della definizione. Le operazioni che applichiamo ai coefficienti del sistema lineare, le applichiamo anche ai termini noti. Quindi aggiungiamo una colonna con i termini noti alla matrice A del sistema lineare.

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

Per scambiare la prima riga con l'ultima, moltiplichiamo la matrice $E_{1,3}$, di scambio tra la riga 1 e la riga 3, per la matrice A .

$$E_{1,3} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

Si noti che il vettore $[0, 0, 1]$, riga 1 di $E_{1,3}$, farà diventare la terza riga prima riga, il vettore $[0, 1, 0]$, riga 2 di $E_{1,3}$, manterrà intatta la seconda riga, mentre il vettore $[1, 0, 0]$, riga 3 di $E_{1,3}$, trasferirà la riga 1 al posto della vecchia riga 3.

$$E_{1,3}A = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 3 & 2 & 1 & 5 \\ 1 & 4 & 3 & 3 \\ 1 & 1 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 4 & 3 & 3 \\ 3 & 2 & 1 & 5 \end{bmatrix}$$

Ora vogliamo sottrarre la prima equazione dalla seconda. Consideriamo la matrice $E_{2+(-1)1}$:

$$E_{2+(-1)1} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Allora abbiamo

$$E_{2+(-1)1}E_{1,3}A = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 4 & 3 & 3 \\ 3 & 2 & 1 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 3 & 2 & 3 \\ 3 & 2 & 1 & 5 \end{bmatrix}$$

Sottraiamo 3 volte la prima equazione dalla terza. Consideriamo la seguente matrice:

$$E_{3+(-3)1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix}$$

Allora abbiamo

$$E_{3+(-3)1}E_{2+(-1)1}E_{1,3}A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 3 & 2 & 3 \\ 3 & 2 & 1 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 3 & 2 & 3 \\ 0 & -1 & -2 & 5 \end{bmatrix}$$

Dividiamo la seconda equazione per 3 e poi sommiamo la seconda equazione alla terza. Consideriamo la matrice

$$E_{3+(\frac{1}{3})2} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \frac{1}{3} & 1 \end{bmatrix}$$

Allora abbiamo

$$E_{3+(\frac{1}{3})2} E_{3+(-3)1} E_{2+(-1)1} E_{1,3} A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \frac{1}{3} & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 3 & 2 & 3 \\ 0 & -1 & -2 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 3 & 2 & 3 \\ 0 & 0 & -4/3 & 6 \end{bmatrix}$$

Il fatto che il prodotto tra matrici è associativo ci permette anche di moltiplicare prima tutte le matrici $E_{3+(\frac{1}{3})2} E_{3+(-3)1} E_{2+(-1)1} E_{1,3}$ e poi applicare il risultato ad A per ottenere il risultato finale.

6.2 Matrice inversa

Le operazioni che abbiamo applicato alla matrice A sono tutte reversibili nel senso che ciascuna delle matrici elementari è invertibile.

Definition 6.2.1. Una matrice quadrata A di ordine n è *invertibile* se esiste una matrice B dello stesso tipo tale che

$$AB = I_n = BA,$$

dove I_n è la matrice identica.

Come dimostreremo nel seguente lemma, esiste al più una matrice inversa di A ; nel caso in cui esiste l'inversa della matrice A essa si indica con A^{-1} .

Lemma 6.2.1. *Siano A e B matrici quadrate di ordine n .*

1. *La matrice inversa A^{-1} di A è unica.*
2. $(AB)^{-1} = B^{-1}A^{-1}$.
3. *Se A è invertibile, il sistema lineare $A\mathbf{x} = \mathbf{b}$ di n equazioni in n incognite ha un'unica soluzione data da $\mathbf{x} = A^{-1}\mathbf{b}$.*

Proof. (1) Supponiamo che C sia un'altra inversa, cioè $AC = CA = I_n$. Allora abbiamo:

$$C = CI_n = C(AA^{-1}) = (CA)A^{-1} = I_nA^{-1} = A^{-1}.$$

$$(2) (AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AI_nA^{-1} = AA^{-1} = I_n.$$

(3) Sia $\mathbf{x} = A^{-1}\mathbf{b}$. Allora si ha: $A(A^{-1}\mathbf{b}) = (AA^{-1})\mathbf{b} = I_n\mathbf{b} = \mathbf{b}$. \square

Lemma 6.2.2. *Le matrici elementari sono invertibili.*

Example 19. La matrice $E_{1,3}$ (che scambia la riga 1 e la riga 3) è invertibile con inversa la matrice stessa:

$$E_{1,3}E_{1,3} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I_3$$

È infatti chiaro che scambiare due volte di seguito la riga uno e la riga tre riporta alla situazione iniziale.

Example 20. La matrice E_{3+r2} è invertibile con inversa $E_{3+(-r)2}$:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & r & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -r & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

6.3 Matrici a gradini

Definition 6.3.1. Una *matrice a gradini* è una matrice che verifica le seguenti condizioni:

1. Tutte i vettori riga nulli sono nella parte bassa della matrice;
2. Il primo elemento in una riga non nulla è un 1;
3. Date due righe successive i e $i + 1$ non nulle, il primo elemento non nullo della riga $i + 1$ si trova a destra del primo elemento non nullo della riga i .

Una *matrice ridotta (per riga)* è una matrice a gradini che verifica anche la seguente condizione:

4. Se una colonna contiene il primo elemento non nullo di una riga, allora tutti gli altri elementi della colonna sono nulli.

Example 21. La prima matrice è in forma a gradini, mentre la seconda è in forma ridotta:

$$\begin{bmatrix} 1 & 3 & 0 & 2 & 5 & 4 \\ 0 & 1 & 0 & 5 & 6 & 7 \\ 0 & 0 & 0 & 1 & 7 & 7 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 4 \\ 0 & 1 & 0 & 0 & 0 & 7 \\ 0 & 0 & 0 & 1 & 0 & 7 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Definition 6.3.2. Due matrici A e B di dimensione $m \times n$ sono *equivalenti per riga*, e scriviamo $A \equiv_r B$, se la matrice B può essere ottenuta dalla matrice A applicando operazioni elementari (di tipo I, II, III).

Si noti che la relazione \equiv_r è una relazione di equivalenza perché le matrici elementari sono invertibili e quindi anche il prodotto di matrici elementari è invertibile: se $B = EA$ per una matrice elementare E , allora $A = E^{-1}B$.

Proposition 6.3.1. *Ogni matrice non nulla è equivalente ad una matrice in forma ridotta. Più precisamente, per ogni matrice A non nulla esiste una matrice F , che è prodotto di matrici elementari, tale che FA è una matrice in forma ridotta.*

Proof. Sia A la matrice di partenza.

Come ottenere una matrice equivalente a gradini:

Consideriamo la prima colonna (da sinistra) con almeno un elemento diverso da zero. Sia j l'indice di colonna e supponiamo che il primo elemento non nullo dall'alto si trovi nella riga i . Scambiamo la riga i con la riga 1, ottenendo la matrice B . Così $b_{1j} \neq 0$. Dividiamo la riga 1 di B per b_{1j} , ottenendo la matrice C , dove $c_{1j} = 1$. Successivamente, per ogni elemento c_{sj} ($2 \leq s \leq m$) diverso da zero, sommiamo $-c_{sj}$ volte la prima riga alla riga s . Otteniamo una matrice D in cui tutti gli elementi della colonna j sono nulli tranne il primo che è 1. Consideriamo la sottomatrice di D ottenuta eliminando la prima riga di D . Applichiamo la stessa procedura alla sottomatrice. Iterando il ragionamento alla fine arriviamo ad una matrice a gradini.

Come ottenere una matrice in forma ridotta da una matrice a gradini:

Sia $H = (h_{ij})$ una matrice a gradini. Applica la procedura seguente dall'alto verso il basso. Considera una riga non nulla, diciamo la riga i . Allora il primo elemento da sinistra della riga i è un 1. Supponiamo che si trovi nella colonna j , cioè, $h_{ij} = 1$. Allora, per ogni riga k ($1 \leq k < i$) con elemento non nullo in posizione kj , cioè $h_{kj} \neq 0$, somma $-h_{kj}$ volte la riga i alla riga k .

Alla fine nella colonna j avremo tutti elementi nulli tranne un 1 in posizione ij . \square

Proposition 6.3.2. *Siano $A\mathbf{x} = \mathbf{a}$ e $B\mathbf{x} = \mathbf{b}$ due sistemi lineari ciascuno con m equazioni in n incognite. Se le matrici complete $A|\mathbf{a}$ e $B|\mathbf{b}$ sono equivalenti allora*

- (i) *I due sistemi hanno le stesse soluzioni.*
- (ii) *Le soluzioni del sistema $A\mathbf{x} = \mathbf{a}$ si ottengono trasformando la matrice completa $[A|\mathbf{a}]$ in forma ridotta.*

Proof. (i) Per ipotesi esiste una matrice G , che è prodotto di matrici elementari, tale che $G(B|\mathbf{b}) = A|\mathbf{a}$. Moltiplicando a blocchi, segue che $GB = A$ e $G\mathbf{b} = \mathbf{a}$. Supponiamo che il vettore colonna \mathbf{x} sia una soluzione del sistema lineare $B\mathbf{x} = \mathbf{b}$. Allora $A\mathbf{x} = GB\mathbf{x} = G\mathbf{b} = \mathbf{a}$. Per simmetria otteniamo la tesi. \square

Un sistema lineare è *omogeneo* se il vettore dei termini noti è il vettore nullo.

Proposition 6.3.3. *Un sistema omogeneo $A\mathbf{x} = \mathbf{0}$ di m equazioni in n incognite ammette sempre una soluzione non nulla se $n > m$.*

Proof. Riduciamo A in forma ridotta ottenendo la matrice B . Siano r il numero di righe non nulle di B . Allora la matrice C di dimensione $r \times n$, formata dalle prime r righe di B , non ha righe nulle. Siccome $r \leq m < n$ possiamo risolvere il sistema con le prime r incognite che dipendono dalle altre $n - r$. Queste ultime possono prendere valori arbitrari. \square

Chapter 7

Spazi vettoriali

Prima di proseguire nel capitolo invitiamo il lettore a rileggersi la definizione di spazio vettoriale in Sezione 2.2.

Riportiamo nel seguito alcuni esempi “non-standard” di spazi vettoriali.

Example 22. (*Spazio vettoriale dei polinomi reali*) Un polinomio reale è una funzione $p : \mathbb{R} \rightarrow \mathbb{R}$ che è esprimibile come

$$p(x) = a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n$$

con coefficienti reali a_i . Per esempio, i seguenti sono polinomi: $3x + 2$, $x^2 + 5x + 1$, etc. I polinomi costituiscono uno spazio vettoriale reale.

Example 23. (*Spazio vettoriale delle sequenze infinite di reali*) L’insieme di tutte le successioni $(a_n)_{n \geq 0}$ di numeri reali è uno spazio vettoriale sul campo reale.

Example 24. (*Spazio vettoriale delle funzioni a valori reali*) Sia X un insieme. Allora l’insieme di tutte le funzioni da X ad \mathbb{R} è uno spazio vettoriale. Se $f, g : X \rightarrow \mathbb{R}$ sono funzioni e r è uno scalare, definiamo:

$$(f + g)(x) = f(x) + g(x); \quad (rf)(x) = r \cdot f(x), \text{ for all } x \in X.$$

L’Esempio 23 è un caso particolare di questo esempio: ogni successione $(a_n)_{n \geq 0}$ è una funzione $a : \mathbb{N} \rightarrow \mathbb{R}$.

7.1 Sottospazi

Definition 7.1.1. Sia V uno spazio vettoriale sul campo \mathbb{K} . Un sottoinsieme non vuoto U di V è un *sottospazio vettoriale* di V se la somma vettoriale di vettori di U è ancora in U e lo stesso accade per il prodotto di un vettore di U per un arbitrario scalare:

1. $\mathbf{v}, \mathbf{w} \in U \Rightarrow \mathbf{v} + \mathbf{w} \in U;$
2. $\mathbf{v} \in U \wedge r \in \mathbb{K} \Rightarrow r\mathbf{v} \in U.$

Dalla seconda condizione della definizione precedente il vettore nullo $\mathbf{0}$ appartiene ad ogni sottospazio vettoriale.

Ogni spazio vettoriale V ammette due sottospazi banali: il primo $\{\mathbf{0}\}$ costituito dal solo vettore nullo e il secondo da V stesso.

Lemma 7.1.1. *L'intersezione di due o più sottospazi vettoriali di V è ancora un sottospazio vettoriale.*

Proof. Siano W e U due sottospazi. Si vede facilmente che, se $\mathbf{v}, \mathbf{w} \in W \cap U$, allora anche che $\mathbf{v} + \mathbf{w} \in W \cap U$ e $r\mathbf{v} \in W \cap U$ per ogni scalare r . \square

Sia V uno spazio vettoriale e $\mathbf{x}_1, \dots, \mathbf{x}_n \in V$ vettori. Un vettore $\mathbf{v} \in V$ è una *combinazione lineare* dei vettori $\mathbf{x}_1, \dots, \mathbf{x}_n$ se esistono scalari r_1, \dots, r_n tali che

$$\mathbf{v} = r_1\mathbf{x}_1 + \cdots + r_n\mathbf{x}_n.$$

Definition 7.1.2. Sia $X \subseteq V$ un sottoinsieme di uno spazio vettoriale V sul campo \mathbb{K} . Il *sottospazio Span(X) di V generato da X* è definito come l'intersezione di tutti i sottospazi di V che contengono X .

$\text{Span}(X)$ è costituito da tutte le combinazioni lineari finite di elementi di X a coefficienti nel campo \mathbb{K} :

$$\text{Span}(X) = \{r_1\mathbf{v}_1 + \cdots + r_k\mathbf{v}_k : r_1, \dots, r_k \in \mathbb{K} \text{ e } \mathbf{v}_1, \dots, \mathbf{v}_k \in X\}$$

Nella proposizione seguente proviamo questo risultato nell'ipotesi che X sia finito.

Proposition 7.1.2. *Sia V uno spazio vettoriale e $\mathbf{x}_1, \dots, \mathbf{x}_n \in V$ vettori. L'insieme delle combinazioni lineari di $\mathbf{x}_1, \dots, \mathbf{x}_n$ è un sottospazio vettoriale di V che coincide con $\text{Span}(\mathbf{x}_1, \dots, \mathbf{x}_n)$.*

Proof. Se $\mathbf{v} = c_1\mathbf{x}_1 + \dots + c_n\mathbf{x}_n$ e $\mathbf{w} = d_1\mathbf{x}_1 + \dots + d_n\mathbf{x}_n$, allora $\mathbf{v} + \mathbf{w} = (c_1 + d_1)\mathbf{x}_1 + \dots + (c_n + d_n)\mathbf{x}_n$ e $m\mathbf{v} = (mc_1)\mathbf{x}_1 + \dots + (mc_n)\mathbf{x}_n$ per ogni scalare m . Infine, ogni combinazione lineare di $\mathbf{x}_1, \dots, \mathbf{x}_n$ appartiene ad ogni sottospazio vettoriale che contiene i vettori $\mathbf{x}_1, \dots, \mathbf{x}_n$. \square

Proposition 7.1.3. 1. I sottospazi vettoriali non banali del piano sono le rette passanti per l'origine.

2. I sottospazi vettoriali non banali dello spazio sono i piani e le rette passanti per l'origine.

Le rette (i piani) che non passano per l'origine NON sono sottospazi vettoriali, perché il vettore nullo non appartiene alla retta (al piano).

Ricordiamo che un sistema lineare è omogeneo se il vettore dei termini noti è il vettore nullo.

Proposition 7.1.4. L'insieme delle soluzioni di un sistema lineare omogeneo di m equazioni in n incognite è un sottospazio vettoriale dello spazio \mathbb{R}^n .

Proof. Sia A una matrice di tipo $m \times n$, \mathbf{x}, \mathbf{y} vettori colonna di lunghezza n tali che $A\mathbf{x} = \mathbf{0}$ e $A\mathbf{y} = \mathbf{0}$. Allora $A(\mathbf{x} + \mathbf{y}) = A\mathbf{x} + A\mathbf{y} = \mathbf{0} + \mathbf{0} = \mathbf{0}$ ed inoltre $A(r\mathbf{x}) = r(A\mathbf{x}) = r\mathbf{0} = \mathbf{0}$. \square

Proposition 7.1.5. Sia $A\mathbf{x} = \mathbf{b}$ un sistema lineare con A matrice di tipo $m \times n$. Il sistema ammette soluzione se il vettore \mathbf{b} appartiene al sottospazio $\text{Span}(A^1, \dots, A^n)$ generato dai vettori colonna di A . In tal caso $\mathbf{b} = c_1A^1 + \dots + c_nA^n$ per opportuni scalari c_1, \dots, c_n e la soluzione del sistema è $\mathbf{x} = [c_1 \dots c_n]^t$.

Concludiamo la sezione con esempi che chiarificano le proposizioni precedenti.

Example 25. La retta di equazione $2x + 3y = 0$ è un sottospazio vettoriale di \mathbb{R}^2 (si consulti Proposizione 7.1.3):

- Se $[x_1, x_2]$ e $[y_1, y_2]$ appartengono alla retta (i.e., $2x_1 + 3x_2 = 0$ e $2y_1 + 3y_2 = 0$), allora anche $[x_1 + y_1, x_2 + y_2]$ appartiene alla retta ($2(x_1 + y_1) + 3(x_2 + y_2) = 0$).
- Se $[x_1, x_2]$ appartiene alla retta (i.e., $2x_1 + 3x_2 = 0$), allora anche $r[x_1, x_2] = [rx_1, rx_2]$ appartiene alla retta (i.e., $r(2x_1 + 3x_2) = 0$).

Example 26. Il piano di equazione $2x + 3y + 4z = 0$ è un sottospazio vettoriale di \mathbb{R}^3 (si consulti Proposizione 7.1.3). Le rette passanti per l'origine sono anch'esse sottospazi vettoriali di \mathbb{R}^3 in quanto intersezione di due piani passanti per l'origine (Si consulti Lemma 7.1.1).

Example 27. Si consideri lo spazio vettoriale delle matrici quadrate di ordine n . I seguenti sono sottospazi vettoriali:

- L'insieme delle matrici triangolari superiori;
- L'insieme delle matrici diagonali;
- L'insieme delle matrici simmetriche.

Example 28. Si consideri la matrice

$$A = \begin{bmatrix} 2 & 1 & 1 \\ 4 & -6 & 0 \\ -4 & -2 & -2 \end{bmatrix}$$

L'insieme delle combinazioni lineari dei vettori colonna della matrice

$$c_1 \begin{bmatrix} 2 \\ 4 \\ -4 \end{bmatrix} + c_2 \begin{bmatrix} 1 \\ -6 \\ -2 \end{bmatrix} + c_3 \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix}$$

costituisce un sottospazio vettoriale W di \mathbb{R}^3 (si consulti Proposizione 7.1.2).

Se il vettore $\begin{bmatrix} a \\ b \\ c \end{bmatrix}$ appartiene al sottospazio W allora il sistema lineare

$$\begin{bmatrix} 2 & 1 & 1 \\ 4 & -6 & 0 \\ -4 & -2 & -2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

ammette soluzione.

Example 29. L'insieme dei polinomi di grado minore o uguale ad n è un sottospazio vettoriale dello spazio dei polinomi.

Example 30. Siano $\mathbf{x}_1 = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}$ e $\mathbf{x}_2 = \begin{bmatrix} 5 \\ 4 \\ 3 \end{bmatrix}$ due vettori nello spazio tridimensionale. Allora le combinazioni lineari di \mathbf{x}_1 e \mathbf{x}_2

$$c_1 \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} + c_2 \begin{bmatrix} 5 \\ 4 \\ 3 \end{bmatrix}$$

descrivono l'insieme dei punti del piano di equazione $x - 2y + z = 0$, che è un sottospazio vettoriale di \mathbb{R}^3 . Infatti, le equazioni $x = 5c_2$, $y = c_1 + 4c_2$ e $z = 2c_1 + 3c_2$ costituiscono le equazioni parametriche di un piano passante per l'origine. Da $c_2 = \frac{x}{5}$ si ricava $5y = 5c_1 + 4x$ e $5z = 10c_1 + 3x$. Infine, sostituendo $5c_1 = 5y - 4x$ nell'equazione $5z = 10c_1 + 3x$ si ottiene l'equazione lineare del piano $x - 2y + z = 0$.

Example 31. Siano $p(x) = x + 3$ e $q(x) = x^2 + 2$ due polinomi. Allora le combinazioni lineari $rp(x) + sq(x)$ con $r, s \in \mathbb{R}$ costituiscono il sottospazio dei polinomi del tipo

$$sx^2 + rx + (2s + 3r)$$

al variare di r, s tra i reali.

Example 32. Siano $A = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ e $C = \begin{bmatrix} 0 & 0 \\ 0 & 4 \end{bmatrix}$ tre matrici.

L'insieme delle combinazioni lineari di A , B e C determina il sottospazio delle matrici che hanno la seguente forma (c_1, c_2, c_3 arbitrari numeri reali):

$$\begin{bmatrix} c_1 & c_2 \\ 0 & c_3 \end{bmatrix}$$

7.2 Vettori linearmente indipendenti

Definition 7.2.1. I vettori $\mathbf{x}_1, \dots, \mathbf{x}_n$ di uno spazio vettoriale V si dicono *linearmente dipendenti* se il vettore nullo $\mathbf{0}$ è una combinazione lineare di $\mathbf{x}_1, \dots, \mathbf{x}_n$ con coefficienti scalari non tutti nulli; altrimenti si dicono *linearmente indipendenti*.

Example 33. Riconsideriamo i tre vettori colonna dell'Esempio 28. Essi sono linearmente dipendenti perché

$$-3 \begin{bmatrix} 2 \\ 4 \\ -4 \end{bmatrix} - 2 \begin{bmatrix} 1 \\ -6 \\ -2 \end{bmatrix} + 8 \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix} = \mathbf{0}.$$

Quindi il sistema omogeneo

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

ammette soluzioni diverse dal vettore nullo $\mathbf{0}$. Infatti si vede facilmente che la terza riga della matrice è un multiplo della prima, così il sistema lineare ammette una retta di soluzioni che sono l'intersezione del piano $2x+y+z=0$ (che è lo stesso piano di equazione $-4x-2y-2z=0$) e del piano $4x-6y=0$. Si noti che se i vettori colonna sono lineramente dipendenti, anche i vettori riga $[2, 1, 1]$, $[4, -6, 0]$ e $[-4, -2, -2]$ lo sono:

$$2[2, 1, 1] + 0[4, -6, 0] + [-4, -2, -2] = \mathbf{0}.$$

Quest'ultima uguaglianza la possiamo scrivere anche così:

$$\begin{bmatrix} 2 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 & 1 \\ 4 & -6 & 0 \\ -4 & -2 & -2 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}$$

Oppure prendendo le matrici trasposte:

$$\begin{bmatrix} 2 & 4 & -4 \\ 1 & -6 & -2 \\ 1 & 0 & -2 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Proposition 7.2.1. *Se $A = (a_{ij})$ è una matrice di tipo $m \times n$ a coefficienti reali e $\mathbf{x} = \begin{bmatrix} x_1 \\ \dots \\ x_n \end{bmatrix}$ un vettore colonna non nullo, allora le seguenti condizioni sono equivalenti:*

1. $A\mathbf{x} = \mathbf{0}$;

2. \mathbf{x} è un vettore perpendicolare agli m vettori riga della matrice:

$$\text{Prodotto interno: } A_i \cdot \mathbf{x} = \sum_{j=1}^n a_{ij}x_j = 0, \text{ per ogni } 1 \leq i \leq m.$$

3. Gli n vettori colonna di A sono linearmente dipendenti:

$$x_1 A^1 + \cdots + x_n A^n = x_1 \begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ \vdots \\ a_{m2} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix} = \mathbf{0}.$$

Corollary 7.2.2. Sia A una matrice di tipo $m \times n$. Le colonne di A sono linearmente indipendenti se e solo se il sistema $A\mathbf{x} = \mathbf{0}$ ammette il vettore nullo come unica soluzione.

Corollary 7.2.3. Sia A una matrice di tipo $m \times n$ con $n > m$. Allora le colonne di A sono linearmente dipendenti.

Proof. Segue dalla Proposizione 7.2.1 e dalla Proposizione 6.3.3. \square

Example 34. Consideriamo la matrice

$$A = \begin{bmatrix} 3 & 4 & 2 \\ 4 & -6 & 0 \\ 1 & 2 & -2 \end{bmatrix}$$

I tre vettori colonna $\begin{bmatrix} 3 \\ 4 \\ 1 \end{bmatrix}$, $\begin{bmatrix} 4 \\ -6 \\ 2 \end{bmatrix}$ e $\begin{bmatrix} 2 \\ 0 \\ -2 \end{bmatrix}$ sono linearmente indipendenti.

Per dimostrarlo è sufficiente verificare che il vettore nullo è l'unica soluzione del sistema lineare omogeneo $A\mathbf{x} = \mathbf{0}$.

7.3 Basi

Definition 7.3.1. Una *base* di uno spazio vettoriale V è un insieme di vettori linearmente indipendenti che generano tutto lo spazio.

Nel seguito consideriamo soltanto spazi vettoriali con basi costituite da un numero finito di elementi.

Sia $\mathbf{x}_1, \dots, \mathbf{x}_n$ una base di V . Allora ogni vettore $\mathbf{v} \in V$ si scrive in maniera unica come combinazione lineare della base. Infatti se $\mathbf{v} = a_1\mathbf{x}_1 + \dots + a_n\mathbf{x}_n = b_1\mathbf{x}_1 + \dots + b_n\mathbf{x}_n$ allora $0 = (a_1 - b_1)\mathbf{x}_1 + \dots + (a_n - b_n)\mathbf{x}_n$. Siccome $\mathbf{x}_1, \dots, \mathbf{x}_n$ sono linearmente indipendenti, si ricava $a_i - b_i = 0$ per ogni i , da cui $a_i = b_i$.

Se $\mathbf{v} = a_1\mathbf{x}_1 + \dots + a_n\mathbf{x}_n$ allora $[a_1, \dots, a_n]$ è il *vettore delle coordinate* di \mathbf{v} rispetto alla base data.

Example 35. I vettori $\mathbf{e}_1 = [1, 0, 0]$, $\mathbf{e}_2 = [0, 1, 0]$ e $\mathbf{e}_3 = [0, 0, 1]$ sono la *base canonica* di \mathbb{R}^3 . Tre vettori qualsiasi linearmente indipendenti costituiscono sempre una base di \mathbb{R}^3 .

Example 36. Lo spazio dei polinomi reali non ammette una base finita, ma soltanto una base infinita

$$1, x, x^2, \dots, x^n, \dots$$

Il sottospazio dei polinomi di grado ≤ 5 ammette una base finita:

$$1, x, x^2, x^3, x^4, x^5.$$

Lo stesso risultato vale per il sottospazio dei polinomi di grado $\leq n$.

Example 37. Sia \mathbb{C} lo spazio vettoriale dei numeri complessi sul campo reale. Allora, i vettori 1 e $i = \sqrt{-1}$ costituiscono una base.

Example 38. Sia \mathbb{C} lo spazio vettoriale dei numeri complessi sul campo complesso. Allora, il vettore 1 è una base.

Proposition 7.3.1. *Sia V uno spazio vettoriale di base $\mathbf{v}_1, \dots, \mathbf{v}_m$. Allora ogni insieme di elementi $\mathbf{w}_1, \dots, \mathbf{w}_n$ con $n > m$ è linearmente dipendente. In particolare, due basi qualsiasi di uno spazio vettoriale hanno lo stesso numero di vettori. Questo numero si dice dimensione dello spazio vettoriale e si indica con $\dim V$.*

Proof. Rappresentiamo \mathbf{w}_i come combinazione lineare della base $\mathbf{v}_1, \dots, \mathbf{v}_m$:

$$\mathbf{w}_i = \sum_{j=1}^m a_{ji} \mathbf{v}_j = a_{1i} \mathbf{v}_1 + a_{2i} \mathbf{v}_2 + \dots + a_{mi} \mathbf{v}_m. \quad (7.1)$$

I coefficienti a_{ji} costituiscono una matrice A di dimensione $m \times n$, per cui si ha:

$$[\mathbf{w}_1, \dots, \mathbf{w}_n] = [\mathbf{v}_1, \dots, \mathbf{v}_m] \begin{bmatrix} a_{11} & \dots & a_{1m} & \dots & a_{1n} \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & \dots & a_{mm} & \dots & a_{mn} \end{bmatrix}$$

Si noti che (i) ogni colonna della matrice A ha qualche coefficiente diverso da zero; (ii) i coefficienti della matrice sono scalari, mentre le sequenze $[\mathbf{v}_1, \dots, \mathbf{v}_m]$ e $[\mathbf{w}_1, \dots, \mathbf{w}_n]$ hanno elementi in V .

Risolviamo il sistema lineare omogeneo:

$$\begin{bmatrix} a_{11} & \dots & a_{1m} & \dots & a_{1n} \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & \dots & a_{mm} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \dots \\ x_n \end{bmatrix} = \begin{bmatrix} 0 \\ \dots \\ 0 \end{bmatrix}.$$

La soluzione non banale $\begin{bmatrix} b_1 \\ \dots \\ b_n \end{bmatrix}$ esiste perché $n > m$ (Proposizione 6.3.3).

Allora abbiamo:

$$\begin{aligned} [\mathbf{w}_1, \dots, \mathbf{w}_n] \begin{bmatrix} b_1 \\ \dots \\ b_n \end{bmatrix} &= [\mathbf{v}_1, \dots, \mathbf{v}_m] \begin{bmatrix} a_{11} & \dots & a_{1m} & \dots & a_{1n} \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & \dots & a_{mm} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} b_1 \\ \dots \\ b_n \end{bmatrix} \\ &= [\mathbf{v}_1, \dots, \mathbf{v}_m] \begin{bmatrix} 0 \\ \dots \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ \dots \\ 0 \end{bmatrix} \end{aligned}$$

E quindi i vettori $\mathbf{w}_1, \dots, \mathbf{w}_n$ sono linearmente dipendenti. \square

Proposition 7.3.2. *Ogni insieme di vettori linearmente indipendenti può essere esteso ad una base. Qualsiasi insieme di generatori dello spazio può essere ridotto ad una base.*

Example 39. • \mathbb{R}^3 ha dimensione 3. La base canonica di \mathbb{R}^3 è costituita

dai vettori $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ e $\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$. Ogni altra base è costituita da tre vettori

linearmente indipendenti. Per esempio, i tre vettori $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$ e $\begin{bmatrix} 2 \\ 0 \\ 3 \end{bmatrix}$

sono linearmente indipendenti e quindi costituiscono una base, perché il sottospazio vettoriale delle soluzioni del sistema lineare omogeneo

$$\begin{bmatrix} 1 & 1 & 2 \\ 1 & 1 & 0 \\ 1 & 0 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

è costituito solo dal vettore nullo. Le coordinate di un punto dello spazio dipendono dalla base scelta. Per esempio, se il punto P ha

coordinate $P = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ rispetto alla base canonica, allora le coordinate dello stesso punto rispetto alla base non canonica definita prima sono:

$P = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ Quindi il concetto di coordinata è dipendente dalla base.

- Lo spazio vettoriale reale \mathbb{C} dei numeri complessi ha dimensione 2. La base canonica sono i vettori 1 e i . Altre basi sono, per esempio, 5 e $3i$; oppure $2 + 3i$ e $1 + i$.
- Lo spazio vettoriale dei numeri complessi sul campo \mathbb{C} ha dimensione 1. La base canonica è data dal vettore 1 .
- Lo spazio delle matrici $m \times n$ ha dimensione mn . La base canonica è costituita dalle matrici A $m \times n$ per cui esistono indici ij tali che $a_{ij} = 1$ mentre tutte le altre componenti $a_{hk} = 0$ per $h \neq i$ e $k \neq j$.
- Lo spazio delle matrici 3×3 triangolari superiori ha dimensione 6.
- Lo spazio dei polinomi di grado ≤ 3 ha dimensione 4 ed ha come base canonica i polinomi $1, x, x^2, x^3$.

Chapter 8

Trasformazioni lineari e matrici

8.1 Trasformazioni lineari

Definition 8.1.1. Siano V e W spazi vettoriali sullo stesso campo. Una funzione $f : V \rightarrow W$ è una *trasformazione lineare* se verifica la seguente proprietà, per ogni $\mathbf{x}, \mathbf{y} \in V$ e scalari r, s :

$$f(r\mathbf{x} + s\mathbf{y}) = rf(\mathbf{x}) + sf(\mathbf{y}).$$

Si noti che ponendo $r = s = 0$ si ricava $f(\mathbf{0}) = \mathbf{0}$.

Example 40. Consideriamo il vettore $\mathbf{x} = [3, 2, -2]$. La funzione $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ definita da

$$f(\mathbf{y}) = 3y_1 + 2y_2 - 2y_3$$

è una trasformazione lineare:

•

$$\begin{aligned} f(\mathbf{y} + \mathbf{z}) &= 3(y_1 + z_1) + 2(y_2 + z_2) - 2(y_3 + z_3) \\ &= (3y_1 + 2y_2 - 2y_3) + (3z_1 + 2z_2 - 2z_3) \\ &= f(\mathbf{y}) + f(\mathbf{z}). \end{aligned}$$

•

$$\begin{aligned} f(r\mathbf{y}) &= 3(ry_1) + 2(ry_2) - 2(ry_3) \\ &= r(3y_1 + 2y_2 - 2y_3) \\ &= rf(\mathbf{y}). \end{aligned}$$

In generale, se fissiamo un vettore $\mathbf{x} = [x_1, x_2, x_3] \in \mathbb{R}^3$, allora la funzione $f : \mathbb{R}^3 \rightarrow \mathbb{R}$, definita tramite il prodotto interno:

$$f(\mathbf{y}) = \mathbf{x} \cdot \mathbf{y}^t = x_1y_1 + x_2y_2 + x_3y_3, \quad \text{per ogni } \mathbf{y} = [y_1, y_2, y_3] \in \mathbb{R}^3, \quad (8.1)$$

è una trasformazione lineare. Richiamiamo dalla Sezione 2.3 le proprietà del prodotto interno che dimostrano la linearità della funzione descritta in (8.1): $\mathbf{x} \cdot (\mathbf{y} + \mathbf{z}) = (\mathbf{x} \cdot \mathbf{y}) + (\mathbf{x} \cdot \mathbf{z})$ e $\mathbf{x} \cdot (r\mathbf{y}) = r(\mathbf{x} \cdot \mathbf{y})$.

Example 41. Sia Pol lo spazio vettoriale dei polinomi in una variabile x . La funzione $f : \text{Pol} \rightarrow \mathbb{R}^2$ definita da:

$$f(a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n) = [a_{n-1}, a_n],$$

è un'applicazione lineare. La funzione $g : \text{Pol} \rightarrow \mathbb{R}$ definita da

$$g(a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n) = a_0$$

non è una applicazione lineare. Infatti, $g(2x^2 + 3x + 1) = 2$, mentre $g(2x^2) + g(3x + 1) = 2 + 3 = 5$.

Sia $f : V \rightarrow W$ una trasformazione lineare. L'*immagine* di f è definita come

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

mentre il *nucleo* di f è

$$\ker(f) = \{\mathbf{v} \in V : f(\mathbf{v}) = \mathbf{0}\}.$$

Example 42. Una trasformazione lineare iniettiva dello spazio trasforma rette in rette, e piani in piani. Una trasformazione lineare iniettiva del piano trasforma rette in rette. Per esempio, la trasformazione lineare iniettiva f del piano, definita da $f[x, y] = [2x + 3y, 2x]$, trasforma la retta $2x + y = 5$ (di equazione parametrica $x = 2 + t$; $y = 1 - 2t$) nella retta di equazione parametrica $x' = 2(2 + t) + 3(1 - 2t)$, $y' = 2(2 + t)$.

Theorem 8.1.1. Il nucleo di una trasformazione lineare $f : V \rightarrow W$ è un sottospazio di V , mentre l'immagine di f è un sottospazio di W . Si ha la seguente relazione:

$$\dim V = \dim \ker(f) + \dim \text{Im}(f).$$

Proof. Dimostriamo che il nucleo è un sottospazio. Siano $\mathbf{v}, \mathbf{t} \in V$ vettori e r uno scalare. Se $f(\mathbf{v}) = \mathbf{0}$ e $f(\mathbf{t}) = \mathbf{0}$ allora $f(\mathbf{v} + \mathbf{t}) = f(\mathbf{v}) + f(\mathbf{t}) = \mathbf{0} + \mathbf{0} = \mathbf{0}$ e $f(r\mathbf{v}) = rf(\mathbf{v}) = r\mathbf{0} = \mathbf{0}$.

Dimostriamo che $\text{Im}(f)$ è un sottospazio. Siano $\mathbf{w}_1, \mathbf{w}_2 \in \text{Im}(f)$. Allora esistono $\mathbf{v}_1, \mathbf{v}_2 \in V$ tali che $\mathbf{w}_i = f(\mathbf{v}_i)$ ($i = 1, 2$). Allora $\mathbf{w}_1 + \mathbf{w}_2 = f(\mathbf{v}_1 + \mathbf{v}_2)$ e $r\mathbf{w}_1 = f(r\mathbf{v}_1)$.

Proviamo ora la relazione tra dimensione del nucleo e dimensione dell'immagine. Sia $\mathbf{v}_1, \dots, \mathbf{v}_k \in V$ una base del nucleo e sia $\mathbf{w}_1, \dots, \mathbf{w}_r \in W$ una base dell'immagine di f . Consideriamo r elementi $\mathbf{v}_{k+1}, \dots, \mathbf{v}_{k+r} \in V$ tali che $f(\mathbf{v}_{k+i}) = \mathbf{w}_i$.

Dimostriamo che i vettori $\mathbf{v}_1, \dots, \mathbf{v}_k, \mathbf{v}_{k+1}, \dots, \mathbf{v}_{k+r}$ sono linearmente indipendenti. Infatti se

$$c_1\mathbf{v}_1 + \dots + c_k\mathbf{v}_k + c_{k+1}\mathbf{v}_{k+1} + \dots + c_{k+r}\mathbf{v}_{k+r} = \mathbf{0},$$

allora

$$\begin{aligned} \mathbf{0} &= f(c_1\mathbf{v}_1 + \dots + c_k\mathbf{v}_k + c_{k+1}\mathbf{v}_{k+1} + \dots + c_{k+r}\mathbf{v}_{k+r}) \\ &= c_1f(\mathbf{v}_1) + \dots + c_kf(\mathbf{v}_k) + c_{k+1}f(\mathbf{v}_{k+1}) + \dots + c_{k+r}f(\mathbf{v}_{k+r}) \\ &= c_1\mathbf{0} + \dots + c_k\mathbf{0} + c_{k+1}\mathbf{w}_1 + \dots + c_{k+r}\mathbf{w}_r \\ &= c_{k+1}\mathbf{w}_1 + \dots + c_{k+r}\mathbf{w}_r \end{aligned}$$

E quindi i vettori $\mathbf{w}_1, \dots, \mathbf{w}_r \in W$ oppure i vettori $\mathbf{v}_1, \dots, \mathbf{v}_k \in V$ sarebbero linearmente dipendenti. Assurdo.

Verifichiamo che i vettori $\mathbf{v}_1, \dots, \mathbf{v}_k, \mathbf{v}_{k+1}, \dots, \mathbf{v}_{k+r}$ generano lo spazio V . Sia $\mathbf{x} \in V$. Se $f(\mathbf{x}) = \mathbf{0}$ allora \mathbf{x} è combinazione lineare di $\mathbf{v}_1, \dots, \mathbf{v}_k$, altrimenti $f(\mathbf{x}) = d_1\mathbf{w}_1 + \dots + d_r\mathbf{w}_r$. E quindi $f(\mathbf{x} - (d_1\mathbf{v}_{k+1} + \dots + d_r\mathbf{v}_{k+r})) = \mathbf{0}$. Scriviamo quindi $\mathbf{x} - (d_1\mathbf{v}_{k+1} + \dots + d_r\mathbf{v}_{k+r})$ come combinazione lineare di $\mathbf{v}_1, \dots, \mathbf{v}_k$ ed otteniamo il risultato. \square

Lemma 8.1.2. *Sia $f : V \rightarrow W$ una trasformazione lineare.*

1. *f è iniettiva sse $\ker(f) = \{\mathbf{0}\}$.*
2. *Se f è iniettiva, allora*
 - $\dim V = \dim \text{Im}(f)$.
 - *Se $\mathbf{v}_1, \dots, \mathbf{v}_n$ sono linearmente indipendenti, allora $f(\mathbf{v}_1), \dots, f(\mathbf{v}_n)$ sono linearmente indipendenti.*

Proof. (1) Abbiamo $f(\mathbf{x}) = f(\mathbf{y})$ sse $f(\mathbf{x} - \mathbf{y}) = \mathbf{0}$. \square

Lemma 8.1.3. *Sia V uno spazio vettoriale di base $\mathbf{v}_1, \dots, \mathbf{v}_n$. Ogni trasformazione lineare $f : V \rightarrow W$ è univocamente determinata dai valori $f(\mathbf{v}_1), \dots, f(\mathbf{v}_n)$ assunti dai vettori della base. Se $\mathbf{v} \in V$ ha coordinate c_1, \dots, c_n rispetto alla data base allora*

$$f(\mathbf{v}) = c_1 f(\mathbf{v}_1) + \cdots + c_n f(\mathbf{v}_n). \quad (8.2)$$

Viceversa, ogni funzione $g : \{\mathbf{v}_1, \dots, \mathbf{v}_n\} \rightarrow W$ può univocamente essere estesa tramite (8.2) ad una trasformazione lineare da V a W .

Example 43. Si consideri la trasformazione lineare $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ dell'Esempio 40 definita da $f(\mathbf{x}) = 3x_1 + 2x_2 - 2x_3$. La dimensione del nucleo di f è 2 perché il nucleo è il piano ortogonale al vettore $[3, 2, -2]$ di equazione $3x_1 + 2x_2 - 2x_3 = 0$, per cui dal Teorema 8.1.1 la dimensione dell'immagine deve essere 1.

Example 44. Consideriamo lo spazio vettoriale infinito dimensionale Pol dei polinomi reali e fissiamo un numero reale, per esempio 3. Allora la funzione $f : \text{Pol} \rightarrow \mathbb{R}$, definita come segue (per ogni polinomio $p(x) = a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n$):

$$f(a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n) = a_03^n + a_13^{n-1} + \cdots + a_{n-1}3^1 + a_n,$$

è una trasformazione lineare. Per esempio, se $p(x) = x^2 + 5x - 1$ allora

$$f(x^2 + 5x - 1) = 3^2 + 5 \times 3 - 1 = 23.$$

Il nucleo di f è il sottospazio vettoriale determinato dall'insieme dei polinomi $p(x)$ che ammettono 3 come radice: $p(3) = 0$. L'immagine di f è tutto \mathbb{R} .

Example 45. Sia V uno spazio di dimensione 2 con base v_1, v_2 e W uno spazio di dimensione 3 con base w_1, w_2, w_3 . Dal Lemma 8.1.3 la funzione $f : V \rightarrow W$ definita da

$$f(v_1) = 3w_1 + 5w_2 - 2w_3; \quad f(v_2) = w_1 + w_3$$

è estendibile ad una trasformazione lineare.

Lemma 8.1.4. *Le seguenti funzioni sono trasformazioni lineari:*

1. *La funzione identica $I : V \rightarrow V$, definita da*

$$I(\mathbf{v}) = \mathbf{v}, \text{ per ogni vettore } \mathbf{v} \in V.$$

2. *La composizione $g \circ f : V \rightarrow U$ di due trasformazioni lineari $f : V \rightarrow W$ e $g : W \rightarrow U$, definita da*

$$(g \circ f)(\mathbf{v}) = g(f(\mathbf{v})), \text{ per ogni vettore } \mathbf{v} \in V.$$

8.2 La matrice di una trasformazione lineare

Sia $f : V \rightarrow W$ una trasformazione lineare. Fissiamo una base $\mathbf{v}_1, \dots, \mathbf{v}_n$ di V ed una base $\mathbf{w}_1, \dots, \mathbf{w}_m$ di W . Allora, l'immagine $f(\mathbf{v}_j)$ di ogni vettore della base di V deve essere combinazione lineare dei vettori della base di W :

$$f(\mathbf{v}_j) = a_{1j}\mathbf{w}_1 + \cdots + a_{mj}\mathbf{w}_m, \text{ per ogni } 1 \leq j \leq n.$$

Consideriamo la matrice A di dimensione $m \times n$ la cui colonna j è determinata

dai coefficienti $\begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix}$ dalle coordinate di $f(\mathbf{v}_j)$.

Per ogni vettore $\mathbf{v} = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n \in V$, si ha

$$\begin{aligned} f(\mathbf{v}) &= f(c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n) \\ &= c_1f(\mathbf{v}_1) + \cdots + c_nf(\mathbf{v}_n) \\ &= c_1(a_{11}\mathbf{w}_1 + \cdots + a_{m1}\mathbf{w}_m) + \cdots + c_n(a_{1n}\mathbf{w}_1 + \cdots + a_{mn}\mathbf{w}_m) \\ &= (c_1a_{11} + c_2a_{12} + \cdots + c_na_{1n})\mathbf{w}_1 + \cdots + (c_1a_{m1} + c_2a_{m2} + \cdots + c_na_{mn})\mathbf{w}_m \end{aligned}$$

Sia $\mathbf{c} = [c_1 \dots c_n]^t$ il vettore colonna delle coordinate di \mathbf{v} . Le coordinate di $f(\mathbf{v})$ si calcolano utilizzando il prodotto matriciale:

$$A\mathbf{c} = \begin{bmatrix} A_1 \cdot \mathbf{c} \\ A_2 \cdot \mathbf{c} \\ \vdots \\ A_m \cdot \mathbf{c} \end{bmatrix}$$

Proposition 8.2.1. *Siano V, W, U spazi vettoriali sullo stesso campo di scalari e siano $f : V \rightarrow W$, $g : W \rightarrow U$ trasformazioni lineari. Siano A la matrice di f e B la matrice di g rispetto a basi scelte in ciascuno dei tre spazi.*

- La matrice BA (prodotto di matrici) è la matrice di $g \circ f : V \rightarrow U$.

Example 46. Sia $f : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ la trasformazione lineare definita da

$$f(\mathbf{v}_1) = 4\mathbf{w}_1 + \mathbf{w}_2 + \mathbf{w}_3; \quad f(\mathbf{v}_2) = \mathbf{w}_1 + 3\mathbf{w}_2,$$

con $\mathbf{v}_1, \mathbf{v}_2$ base di \mathbb{R}^2 e $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3$ base di \mathbb{R}^3 . Allora la matrice della trasformazione lineare è:

$$\begin{bmatrix} 4 & 1 \\ 1 & 3 \\ 1 & 0 \end{bmatrix}$$

Se $\mathbf{v} = 5\mathbf{v}_1 + 3\mathbf{v}_2$ è un vettore di \mathbb{R}^2 , allora

$$\begin{bmatrix} 4 & 1 \\ 1 & 3 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 5 \\ 3 \end{bmatrix} = \begin{bmatrix} 23 \\ 14 \\ 5 \end{bmatrix}$$

Quindi $f(\mathbf{v}) = 23\mathbf{w}_1 + 14\mathbf{w}_2 + 5\mathbf{w}_3$.

Ogni trasformazione lineare $f : V \rightarrow W$ assume una forma matriciale veramente semplice se si scelgono le basi di V e W opportunamente.

Denotiamo con I_n la matrice identica di ordine n e con $0_{m,n}$ la matrice nulla di tipo $m \times n$.

Proposition 8.2.2. *Sia $f : V \rightarrow W$ una trasformazione lineare dallo spazio vettoriale V di dimensione n allo spazio vettoriale W di dimensione m . Supponiamo che $r = \dim \text{Im}(f)$ e $k = \dim \ker(f)$ con $n = k + r$. Allora esistono basi di V e W tali che la matrice A (di dimensione $m \times n$) di f rispetto a queste basi assume la forma*

$$A = \begin{bmatrix} I_r & 0_{r,k} \\ 0_{m-r,r} & 0_{m-r,k} \end{bmatrix}$$

Proof. Sia v_1, \dots, v_k una base del nucleo di f . Completiamo v_1, \dots, v_k ad una base di V : $u_1, \dots, u_r, v_1, \dots, v_k$ con $r + k = n$. Le immagini $f(u_1), \dots, f(u_r)$ tramite f dei vettori u_1, \dots, u_r sono non nulle e costituiscono una base di $\text{Im}(f)$. Completiamo $f(u_1), \dots, f(u_r)$ ad una base di W :

$$f(u_1), \dots, f(u_r), w_1, \dots, w_{m-r}.$$

La matrice A di f rispetto a queste due basi verifica le condizioni della proposizione. \square

Remark 1. Supponiamo che gli spazi di partenza e di arrivo della trasformazione lineare f della Proposizione 8.2.2 coincidono: $f : V \rightarrow V$. Allora la matrice A di f assume la forma della proposizione soltanto per opportune basi distinte di V . La forma descritta non sarà in generale assunta se la base di partenza coincide con la base di arrivo.

8.3 La trasformazione lineare di una matrice

Sia V uno spazio di base $\mathbf{v}_1, \dots, \mathbf{v}_n$ e W uno spazio di base $\mathbf{w}_1, \dots, \mathbf{w}_m$ sullo stesso campo numerico e sia $A = (a_{ij})$ una matrice di tipo $m \times n$ ad elementi scalari.

La matrice A determina

1. Il sottospazio di V generato dai vettori di V le cui coordinate sono le righe della matrice. Tale sottospazio è denotato con $\text{Span}(A_1, \dots, A_m)$.
2. Il sottospazio di W generato dai vettori di W le cui coordinate sono le colonne della matrice. Tale sottospazio è denotato con $\text{Span}(A^1, \dots, A^n)$.
3. Una trasformazione lineare $f_A : V \rightarrow W$ definita come segue. Per ogni vettore $\mathbf{v} = c_1\mathbf{v}_1 + \dots + c_n\mathbf{v}_n \in V$, consideriamo le sue coordinate come

vettore colonna $\mathbf{c} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}$. Definiamo le coordinate di $f_A(\mathbf{v})$ rispetto alla base $\mathbf{w}_1, \dots, \mathbf{w}_m$ con il prodotto matriciale:

$$A\mathbf{c} = A \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} A_1 \cdot \mathbf{c} \\ A_2 \cdot \mathbf{c} \\ \vdots \\ A_m \cdot \mathbf{c} \end{bmatrix} = c_1 A^1 + \dots + c_n A^n = \begin{bmatrix} \sum_{i=1}^n c_i a_{1i} \\ \sum_{i=1}^n c_i a_{2i} \\ \vdots \\ \sum_{i=1}^n c_i a_{mi} \end{bmatrix}$$

Così si ha $f_A(\mathbf{v}) = (\sum_{i=1}^n c_i a_{1i})\mathbf{w}_1 + \dots + (\sum_{i=1}^n c_i a_{mi})\mathbf{w}_m$. f_A è lineare per le proprietà del prodotto matriciale: $A(\mathbf{x} + \mathbf{y}) = A\mathbf{x} + A\mathbf{y}$; $A(c\mathbf{x}) = c(A\mathbf{x})$. Dal Teorema 8.1.1 si ha:

$$n = \dim \ker(f_A) + \dim \text{Im}(f_A) = \dim \ker(f_a) + \dim \text{Span}(A^1, \dots, A^n)$$

dove $\text{Span}(A^1, \dots, A^n) = \text{Im}(f_A)$ è il sottospazio di W generato dai vettori colonna della matrice A .

Il sottospazio generato dai vettori riga $\text{Span}(A_1, \dots, A_m)$ è lo spazio ortogonale al sottospazio $\ker(f_A)$, cioè $\text{Span}(A_1, \dots, A_m) = \ker(f_A)^\perp$.

Abbiamo $A\mathbf{c} = \begin{bmatrix} A_1 \cdot \mathbf{c} \\ A_2 \cdot \mathbf{c} \\ \vdots \\ A_m \cdot \mathbf{c} \end{bmatrix} = \mathbf{0}$ sse \mathbf{c} è ortogonale ad ogni vettore nello

spazio $\text{Span}(A_1, \dots, A_m)$:

$$(d_1 A_1 + \dots + d_m A_m) \cdot \mathbf{c} = d_1 (A_1 \cdot \mathbf{c}) + \dots + d_m (A_m \cdot \mathbf{c}) = d_1 \mathbf{0} + \dots + d_m \mathbf{0} = \mathbf{0}.$$

Quindi, dalla Proposizione 11.2.1 si ricava:

$$\begin{aligned} n &= \text{numero delle colonne} \\ &= \dim \ker(f_A) + \dim(\text{spazio delle righe}) \\ &= \dim \ker(f_A) + \dim(\text{spazio delle colonne}). \end{aligned}$$

E quindi lo spazio delle colonne e quello delle righe hanno la stessa dimensione.

4. Una trasformazione lineare ${}_A f : W \rightarrow V$ definita come segue. Per ogni vettore $\mathbf{w} \in W$, consideriamo le sue coordinate come vettore riga $\mathbf{d} = [d_1, d_2, \dots, d_m]$. Definiamo le coordinate di ${}_A f(\mathbf{w})$ rispetto alla base $\mathbf{v}_1, \dots, \mathbf{v}_n$ con il prodotto matriciale:

$$\begin{aligned} \mathbf{d}A &= [d_1, d_2, \dots, d_m]A = [\mathbf{d}A^1, \mathbf{d}A^2, \dots, \mathbf{d}A^n] = d_1 A_1 + \dots + d_m A_m = \\ &= \left[\sum_{i=1}^m d_i a_{i1}, \sum_{i=1}^m d_i a_{i2}, \dots, \sum_{i=1}^m d_i a_{in} \right] \end{aligned}$$

Si vede facilmente che ${}_A f$ è lineare per le proprietà del prodotto matriciale. Valgono proprietà analoghe al punto (3):

$$\begin{aligned} m &= \text{numero delle righe} \\ &= \dim \ker({}_A f) + \dim(\text{spazio delle colonne}) \\ &= \dim \ker({}_A f) + \dim(\text{spazio delle righe}). \end{aligned}$$

Example 47. Consideriamo la matrice 3×4 :

$$A = \begin{bmatrix} 3 & 4 & 2 & 1 \\ 4 & -6 & 0 & 0 \\ 1 & 2 & -2 & 2 \end{bmatrix}$$

La matrice A determina:

- (i) Una trasformazione lineare $f_A : \mathbb{R}^4 \rightarrow \mathbb{R}^3$ rispetto alle basi canoniche di \mathbb{R}^4 e \mathbb{R}^3 , che è definita tramite il prodotto matriciale a destra. Per

esempio, $f_A \begin{bmatrix} 1 \\ 0 \\ 3 \\ 5 \end{bmatrix}$ si calcola con il prodotto matriciale:

$$\begin{bmatrix} 3 & 4 & 2 & 1 \\ 4 & -6 & 0 & 0 \\ 1 & 2 & -2 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 3 \\ 5 \end{bmatrix} = \begin{bmatrix} 14 \\ -12 \\ 5 \end{bmatrix}$$

Attenzione: Se consideriamo una base diversa di \mathbb{R}^4 , la matrice A definisce un'altra trasformazione lineare!

- (ii) Una trasformazione lineare ${}_Af : \mathbb{R}^3 \rightarrow \mathbb{R}^4$ rispetto alle basi canoniche di \mathbb{R}^4 e \mathbb{R}^3 , che è definita tramite il prodotto matriciale a sinistra. Per esempio, ${}_Af[0, 1, 2]$ si calcola con il prodotto matriciale:

$$[0, 1, 2] \begin{bmatrix} 3 & 4 & 2 & 1 \\ 4 & -6 & 0 & 0 \\ 1 & 2 & -2 & 2 \end{bmatrix} = [6, -2, -4, 4]$$

- (iii) Il sottospazio di \mathbb{R}^3 generato dai vettori colonna $A^1 = \begin{bmatrix} 3 \\ 4 \\ 1 \end{bmatrix}$, $A^2 = \begin{bmatrix} 4 \\ -6 \\ 2 \end{bmatrix}$, $A^3 = \begin{bmatrix} 2 \\ 0 \\ -2 \end{bmatrix}$ e $A^4 = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}$ le cui coordinate sono le colonne della matrice. Tale sottospazio è denotato con $\text{Span}(A^1, A^2, A^3, A^4)$.

Il sottospazio $\text{Span}(A^1, A^2, A^3, A^4)$ di \mathbb{R}^3 coincide con il sottospazio $\text{Im}(f_A)$ perché

$$\text{Im}(f_A) = \{A\mathbf{c} : \mathbf{c} \in \mathbb{R}^4\} = \{c_1A^1 + c_2A^2 + c_3A^3 + c_4A^4 : \mathbf{c} \in \mathbb{R}^4\} = \text{Span}(A^1, A^2, A^3, A^4).$$

- (iv) Il sottospazio di \mathbb{R}^4 generato dai vettori riga $A_1 = [3, 4, 2, 1]$, $A_2 = [4, -6, 0, 0]$ e $A_3 = [1, 2, -2, 2]$ le cui coordinate sono le righe della matrice. Tale sottospazio è denotato con $\text{Span}(A_1, A_2, A_3)$.

Il sottospazio $\text{Span}(A_1, A_2, A_3)$ di \mathbb{R}^4 è il sottospazio ortogonale allo spazio $\ker(f_A)$ perché

$$\ker(f_A) = \{\mathbf{c} : A\mathbf{c} = \mathbf{0}\} = \{\mathbf{c} : A_i \cdot \mathbf{c} = 0 \text{ per ogni } 1 \leq i \leq 3\} = \text{Span}(A_1, A_2, A_3)^\perp.$$

Dal Teorema 8.1.1 e dalla Proposizione 11.2.1 si ha

$$\begin{aligned} 4 &= \dim \ker(f_A) + \dim \text{Im}(f_A) \\ &= \dim \ker(f_A) + \dim \text{spazio colonne} \\ &= \dim \ker(f_A) + \dim \text{spazio righe} \end{aligned}$$

Quindi la dimensione dello spazio delle colonne coincide con la dimensione dello spazio delle righe.

8.4 Isomorfismi e cambi di base

Una trasformazione lineare è un *isomorfismo* se è bigettiva. Due spazi vettoriali sono *isomorfi* se esiste un isomorfismo tra di loro.

La funzione inversa $f^{-1} : W \rightarrow V$ di un isomorfismo $f : V \rightarrow W$ è anch'essa un isomorfismo lineare.

Proposition 8.4.1. *Due spazi vettoriali sullo stesso campo sono isomorfi se hanno la stessa dimensione.*

Proof. Supponiamo che V e W abbiano la stessa dimensione. Siano $\mathbf{v}_1, \dots, \mathbf{v}_n$ una base di V e $\mathbf{w}_1, \dots, \mathbf{w}_n$ una base di W . Se $\mathbf{v} \in V$, allora possiamo rappresentare \mathbf{v} in maniera unica tramite le sue coordinate: $\mathbf{v} = c_1\mathbf{v}_1 + \dots + c_n\mathbf{v}_n$ per opportuni scalari c_1, \dots, c_n . Allora definiamo

$$f(\mathbf{v}) = c_1\mathbf{w}_1 + \dots + c_n\mathbf{w}_n.$$

Proviamo che f è un isomorfismo. Siano $\mathbf{v} = c_1\mathbf{v}_1 + \dots + c_n\mathbf{v}_n$ e $\mathbf{t} = d_1\mathbf{v}_1 + \dots + d_n\mathbf{v}_n$ due vettori di V .

- f è iniettiva: Se $f(\mathbf{v}) = f(\mathbf{t})$, allora $c_1\mathbf{w}_1 + \dots + c_n\mathbf{w}_n = d_1\mathbf{w}_1 + \dots + d_n\mathbf{w}_n$. Per la indipendenza lineare dei \mathbf{w}_i si ha $c_i = d_i$ per ogni i e quindi $\mathbf{v} = \mathbf{t}$.
- f è surgettiva: Se $\mathbf{w} = s_1\mathbf{w}_1 + \dots + s_n\mathbf{w}_n \in W$ allora $f(s_1\mathbf{v}_1 + \dots + s_n\mathbf{v}_n) = \mathbf{w}$.

□

Proposition 8.4.2. *Sia V uno spazio di dimensione n e W uno spazio di dimensione m . Allora lo spazio vettoriale delle trasformazioni lineari da V a W è isomorfo allo spazio vettoriale delle matrici $m \times n$.*

Proof. Fissiamo una base $\mathbf{v}_1, \dots, \mathbf{v}_n$ di V ed una base $\mathbf{w}_1, \dots, \mathbf{w}_m$ di W . Alla trasformazione lineare

$$f : V \rightarrow W$$

associamo la matrice $A_f = [A^1, \dots, A^n]$ con m righe ed n colonne (A^i è la colonna i della matrice A_f) tale che

$$f(\mathbf{v}_i) = A^i \begin{bmatrix} \mathbf{w}_1 \\ \vdots \\ \mathbf{w}_m \end{bmatrix}$$

Viceversa, ad una matrice A con m righe ed n colonne associamo la trasformazione lineare $f_A : V \rightarrow W$ definita da

$$f_A(\mathbf{v}_i) = \sum_{k=1}^m a_{ki} \mathbf{w}_k.$$

□

Proposition 8.4.3. *Sia V un arbitrario spazio vettoriale di dimensione n di base $\mathbf{v}_1, \dots, \mathbf{v}_n$, ed A una matrice quadrata di ordine n . La matrice A è invertibile sse la trasformazione lineare $f_A : V \rightarrow V$, definita da $f_A(\mathbf{v}_i) = \sum_{k=1}^n a_{ki} \mathbf{v}_k$, è un isomorfismo. In tal caso, la matrice inversa A^{-1} di A è la matrice della trasformazione lineare f^{-1} .*

Proof. (\Leftarrow) Sia f_A un isomorfismo e $I_V : V \rightarrow V$ la funzione identica. Allora f_A^{-1} è una trasformazione lineare tale che $f_A \circ f_A^{-1} = I_V = f_A^{-1} \circ f_A$. Sia B la matrice della trasformazione lineare f_A^{-1} . Allora si ha $AB = I = BA$ con I_n matrice identica. Quindi A è invertibile e $B = A^{-1}$.

(\Rightarrow) Sia A è invertibile. Il lettore è invitato a completare la prova. □

8.4.1 Cambio di base

Sia V uno spazio vettoriale di dimensione n , sia $\mathbf{v}_1, \dots, \mathbf{v}_n$ una prima base e $\mathbf{t}_1, \dots, \mathbf{t}_n$ una seconda base. Consideriamo la trasformazione lineare identica

$I : V \rightarrow V$. Rappresentiamo la trasformazione lineare identica con la matrice quadrata A di ordine n del cambio di base:

$$\mathbf{v}_i = a_{1i}\mathbf{t}_1 + \cdots + a_{ni}\mathbf{t}_n$$

Se le coordinate del vettore $\mathbf{v} = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n$ sono il vettore colonna $\mathbf{c} = \begin{bmatrix} c_1 \\ c_2 \\ \dots \\ c_n \end{bmatrix}$ nella prima base, allora le coordinate di \mathbf{v} saranno $A\mathbf{c} = \begin{bmatrix} \sum_{i=1}^n a_{1i}c_i \\ \sum_{i=1}^n a_{2i}c_i \\ \dots \\ \sum_{i=1}^n a_{ni}c_i \end{bmatrix}$ nella seconda base.

La matrice di un cambio di base è invertibile, perché il cambio di base inverso è rappresentato dalla matrice inversa.

Definition 8.4.1. Due matrici quadrate A e B di ordine n sullo stesso campo, si dicono *simili* sse esiste una matrice invertibile P di ordine n tale che

$$B = PAP^{-1}.$$

Proposition 8.4.4. Se una trasformazione lineare $f : V \rightarrow V$ è rappresentata dalla matrice A rispetto alla base $\mathbf{v}_1, \dots, \mathbf{v}_n$ e dalla matrice B rispetto alla base $\mathbf{w}_1, \dots, \mathbf{w}_n$, allora A e B sono matrici simili.

Example 48. Consideriamo la base canonica $\mathbf{e}_1 = (1, 0)$ e $\mathbf{e}_2 = (0, 1)$ su \mathbf{R}^2 e la base $\mathbf{t}_1 = (1, 1)$, $\mathbf{t}_2 = (3, 2)$. Allora

$$\mathbf{e}_1 = -2\mathbf{t}_1 + \mathbf{t}_2; \quad \mathbf{e}_2 = 3\mathbf{t}_1 - \mathbf{t}_2$$

Quindi la matrice sarà

$$\begin{bmatrix} -2 & 3 \\ 1 & -1 \end{bmatrix}$$

Allora il vettore di coordinate $\begin{bmatrix} 4 \\ 4 \end{bmatrix}$ rispetto alla base canonica avrà coordinate

$\begin{bmatrix} 4 \\ 0 \end{bmatrix}$ rispetto alla seconda base non canonica:

$$\begin{bmatrix} -2 & 3 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 4 \\ 4 \end{bmatrix} = \begin{bmatrix} 4 \\ 0 \end{bmatrix} = 4\mathbf{t}_1.$$

Il vettore è lo stesso, cambiano solo le coordinate.

Example 49. (*Matrici elementari*) Le matrici elementari rappresentano o dei cambiamenti di base oppure degli isomorfismi. Cominciamo con i cambiamenti di base.

Le matrici elementari di tipo I scambiano le componenti di una base. Per esempio, la matrice

$$E_{1,3} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

fa passare dalla base $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ alla base $\mathbf{v}_3, \mathbf{v}_2, \mathbf{v}_1$.

Le matrici elementari di tipo II moltiplicano per uno scalare un componente della base. Per esempio,

$$E_1^c = \begin{bmatrix} c & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

fa passare dalla base $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ alla base $c\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$.

Infine le matrici elementari di tipo III, sostituiscono un componente della base con una sua combinazione lineare. Per esempio,

$$E_{2+c1} = \begin{bmatrix} 1 & 0 & 0 \\ c & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

fa passare dalla base $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ alla base $\mathbf{v}_1, \mathbf{v}_2 + c\mathbf{v}_1, \mathbf{v}_3$.

In tutti i casi precedenti le matrici elementari rappresentano la trasformazione lineare identica.

Dalla Proposizione 8.4.3 segue che le matrici elementari in quanto invertibili rappresentano degli isomorfismi. Fissata la base $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$, la matrice $E_{1,3}$ di tipo I rappresenta l'isomorfismo

$$f_{E_{1,3}}(\mathbf{v}_1) = \mathbf{v}_3; \quad f_{E_{1,3}}(\mathbf{v}_2) = \mathbf{v}_2; \quad f_{E_{1,3}}(\mathbf{v}_3) = \mathbf{v}_1.$$

La matrice elementare di tipo II E_1^c rappresenta l'isomorfismo

$$f_{E_1^c}(\mathbf{v}_1) = c\mathbf{v}_1; \quad f_{E_1^c}(\mathbf{v}_2) = \mathbf{v}_2; \quad f_{E_1^c}(\mathbf{v}_3) = \mathbf{v}_3.$$

Infine la matrice elementare E_{2+c1} di tipo III rappresenta l'isomorfismo

$$f_{E_{2+c1}}(\mathbf{v}_1) = \mathbf{v}_1; \quad f_{E_{2+c1}}(\mathbf{v}_2) = \mathbf{v}_2 + c\mathbf{v}_1; \quad f_{E_{2+c1}}(\mathbf{v}_3) = \mathbf{v}_3.$$

8.5 Sistemi lineari e trasformazioni lineari

Consideriamo un sistema lineare omogeneo di m equazioni in n incognite: $A\mathbf{x} = \mathbf{0}$ con A matrice di tipo $m \times n$. Fissiamo le basi canoniche in \mathbb{R}^n e \mathbb{R}^m . Allora la matrice A rappresenta una trasformazione lineare $f_A : \mathbb{R}^n \rightarrow \mathbb{R}^m$. Abbiamo una soluzione non banale del sistema omogeneo sse $\ker(f_A) \neq \{\mathbf{0}\}$. Ricordiamo dal Teorema 8.1.1 che

$$n = \dim \mathbb{R}^n = \dim \ker(f_A) + \dim \text{Im}(f_A).$$

Se $n > m$, sicuramente il sistema ha soluzione (altrimenti, $n = \dim \text{Im}(f_A) \leq m$).

Dato il vettore colonna non nullo $\mathbf{b} \in \mathbb{R}^m$, il sistema lineare $A\mathbf{x} = \mathbf{b}$ ammette soluzione se $\mathbf{b} \in \text{Im}(f_A)$.

Spieghiamo ora il metodo di eliminazione di Gauss in termini di composizione di trasformazioni lineari. Sia A una matrice $m \times n$ e sia $f_A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ la trasformazione lineare associata alla matrice (rispetto alle basi canoniche). Se $g : \mathbb{R}^m \rightarrow \mathbb{R}^m$ è un arbitrario isomorfismo rappresentato dalla matrice quadrata B di ordine m , allora

$$A\mathbf{x} = \mathbf{0} \text{ sse } \mathbf{x} \in \ker(f_A) \text{ sse } \mathbf{x} \in \ker(g \circ f_A) \text{ sse } BA\mathbf{x} = \mathbf{0}.$$

Se il sistema lineare ha un vettore \mathbf{b} di termini noti non tutti nulli, si ha:

$$A\mathbf{x} = \mathbf{b} \text{ sse } \mathbf{b} \in \text{Im}(f_A) \text{ sse } g(\mathbf{b}) \in \text{Im}(g \circ f_A) \text{ sse } BA\mathbf{x} = B\mathbf{b}.$$

L'isomorfismo g che utilizziamo nel metodo di eliminazione di Gauss è composizione di isomorfismi determinati da matrici elementari.

Chapter 9

Determinante

Il determinante $\det(A)$ di una matrice quadrata A ha un significato geometrico.

9.1 Determinante di una matrice di tipo 2×2

Consideriamo due vettori colonna nel piano: $P = [p_1, p_2]^t$ e $Q = [q_1, q_2]^t$. Con i vettori P e Q possiamo costruire la matrice

$$A = \begin{bmatrix} p_1 & q_1 \\ p_2 & q_2 \end{bmatrix}$$

le cui colonne coincidono con i vettori P e Q rispettivamente. Il luogo dei punti $rP + sQ$ ($0 \leq r, s \leq 1$) costituisce il parallelogramma delimitato dall'origine O degli assi cartesiani e dai punti $P, Q, P + Q$. Sappiamo che i vettori P e Q formano un angolo $\theta \leq \pi$, il cui coseno è calcolato con il prodotto interno: $\cos \theta = \frac{\|P\|\|Q\|}{P \cdot Q}$. L'area $\|P\|\|Q\| \sin(\theta)$ del parallelogramma $O, P, Q, P+Q$ è il modulo del determinante della matrice A (si veda la figura a pagina 94), mentre il segno del determinante dipende dall'orientazione dei vettori P e Q rispetto ai vettori della base canonica del piano cartesiano. Calcoliamo ora l'area in termini delle coordinate dei vettori P e Q :

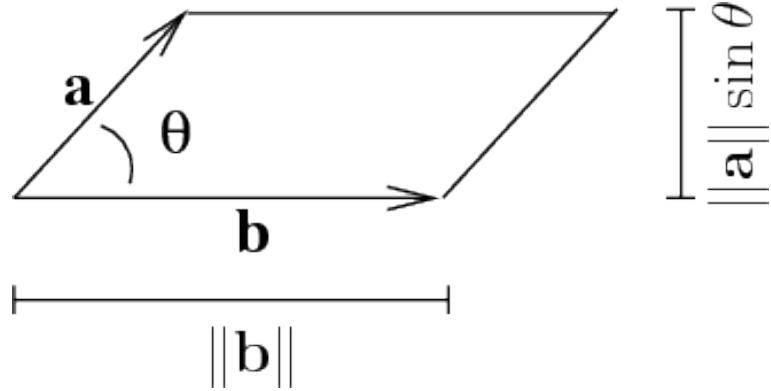


Figure 9.1: Area del parallelogramma determinato dai vettori \mathbf{a} e \mathbf{b} nel piano

$$\begin{aligned}
 \|P\|\|Q\| \sin(\theta) &= \|P\|\|Q\| \sqrt{1 - \cos(\theta)^2} \\
 &= \|P\|\|Q\| \sqrt{1 - \frac{(P \cdot Q)^2}{(P \cdot P)(Q \cdot Q)}} \\
 &= \sqrt{(P \cdot P)(Q \cdot Q) - (P \cdot Q)^2} \\
 &= \sqrt{(p_1^2 + p_2^2)(q_1^2 + q_2^2) - (p_1 q_1 + p_2 q_2)^2} \\
 &= \sqrt{p_1^2 q_2^2 + p_2^2 q_1^2 - 2p_1 p_2 q_1 q_2} \\
 &= \sqrt{(p_1 q_2 - p_2 q_1)^2} \\
 &= |p_1 q_2 - p_2 q_1|.
 \end{aligned}$$

Il segno del determinante è positivo se la rotazione che porta il vettore P in Q attraversando l'angolo θ è antioraria, mentre il segno è negativo se la rotazione è oraria. In conclusione,

$$\det(A) = p_1 q_2 - p_2 q_1.$$

Il determinante è 0 se i vettori P e Q sono allineati.

9.2 Determinante di una matrice di tipo 3×3

Analizziamo ora la situazione nello spazio. Consideriamo tre vettori colonna $\mathbf{p} = [p_1, p_2, p_3]^t$, $\mathbf{q} = [q_1, q_2, q_3]^t$, $\mathbf{r} = [r_1, r_2, r_3]^t$ e la matrice

$$B = \begin{bmatrix} p_1 & q_1 & r_1 \\ p_2 & q_2 & r_2 \\ p_3 & q_3 & r_3 \end{bmatrix}$$

Il luogo dei punti $up + sq + tr$ ($0 \leq u, s, t \leq 1$) costituisce il parallelepipedo delimitato dall'origine O degli assi e dai vettori $\mathbf{p}, \mathbf{q}, \mathbf{r}$ (si veda la figura a pagina 95). Il volume del parallelepipedo è il modulo del determinante della matrice B , mentre il segno del determinante dipende dall'orientazione dei tre vettori rispetto alla orientazione dei tre assi cartesiani.

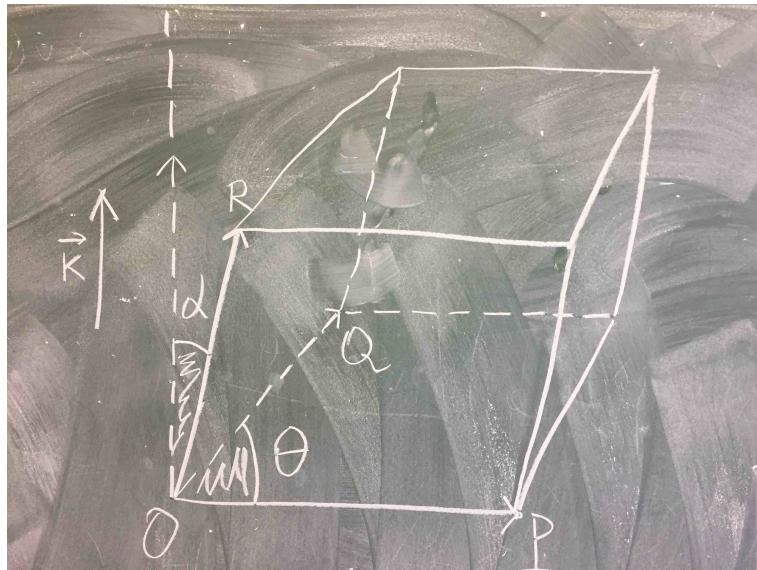


Figure 9.2: Area del parallelepipedo delimitato dai vettori $\mathbf{p}, \mathbf{q}, \mathbf{r}$ nello spazio

Calcoliamo l'area del parallelepipedo eseguendo il ragionamento con l'ausilio della figura di pagina 95. I vettori \mathbf{p} e \mathbf{q} determinano un piano di equazione parametrica $u\mathbf{p} + s\mathbf{q}$ ($u, s \in \mathbb{R}$). Indichiamo con \mathbf{k} il vettore unitario perpendicolare al piano di \mathbf{p} e \mathbf{q} . Il verso di \mathbf{k} è determinato dalla figura applicando la regola della mano destra (si veda la figura di pagina 96). L'area del parallelogramma determinato dai vettori \mathbf{p} e \mathbf{q} è $\|\mathbf{p}\| \|\mathbf{q}\| \sin(\theta)$. L'area del parallelepipedo è pari all'area di questo parallelogramma moltiplicata per l'altezza del parallelepipedo rispetto al piano determinato dai vettori \mathbf{p} e \mathbf{q} .

L'altezza è $\|\mathbf{r}\| \cos \alpha$. In conclusione, l'area è $\|\mathbf{p}\| \|\mathbf{q}\| \sin \theta \|\mathbf{r}\| \cos \alpha$. Se indichiamo con $\mathbf{p} \times \mathbf{q}$ il vettore $\|\mathbf{p}\| \|\mathbf{q}\| \sin(\theta) \mathbf{k}$ allora l'area orientata (cioé, con segno) si può ottenere con il prodotto interno:

$$(\mathbf{p} \times \mathbf{q}) \cdot \mathbf{r}.$$

Calcoliamo l'area del parallelepipedo determinato da tre vettori in termini delle coordinate dei tre vettori.

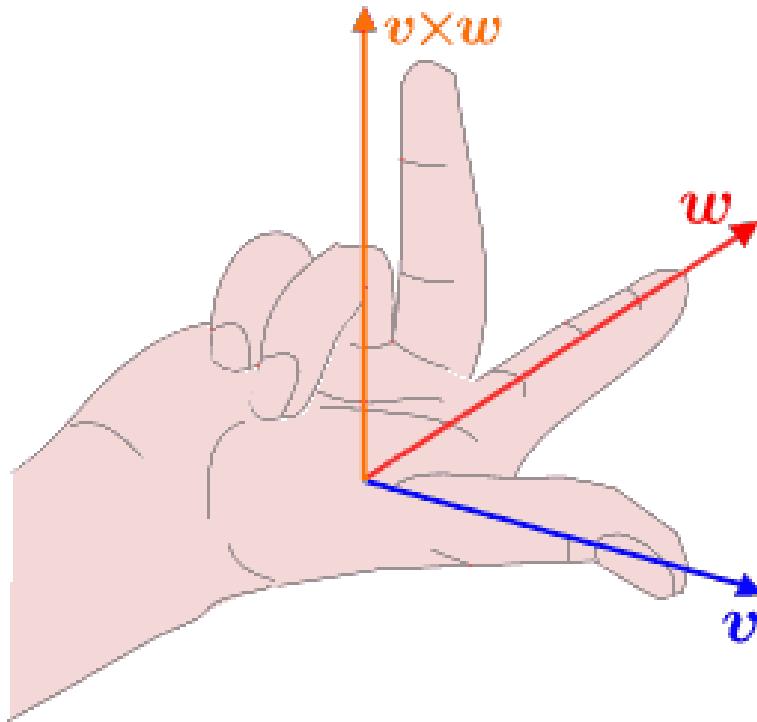


Figure 9.3: Regola della mano destra

Consideriamo prima i due vettori \mathbf{p} e \mathbf{q} , e sia $\theta_{\mathbf{pq}}$ l'angolo da essi formato nel piano $u\mathbf{p} + s\mathbf{q}$ ($u, s \in \mathbb{R}$). Abbiamo già visto che l'area del parallelogramma determinato da \mathbf{p} e \mathbf{q} è $\|\mathbf{p}\| \|\mathbf{q}\| \sin(\theta_{\mathbf{pq}})$. Quest'area è anche uguale a

$\|\mathbf{p}\| \|\mathbf{p} + \mathbf{q}\| \sin(\theta_{\mathbf{p}(\mathbf{p}+\mathbf{q})})$, dove $\theta_{\mathbf{p}(\mathbf{p}+\mathbf{q})}$ è l'angolo formato dai vettori \mathbf{p} e $\mathbf{p} + \mathbf{q}$. Sia \mathbf{k} il vettore unitario perpendicolare al piano $u\mathbf{p} + s\mathbf{q}$ ($u, s \in \mathbb{R}$). Il verso di \mathbf{k} è scelto in maniera tale che i tre vettori $\mathbf{p}, \mathbf{q}, \mathbf{k}$ abbiano la stessa orientazione dei tre vettori unitari $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$. Allora il *prodotto vettoriale* $\mathbf{p} \times \mathbf{q}$ di \mathbf{p} e \mathbf{q} è definito come il vettore

$$\mathbf{p} \times \mathbf{q} = \|\mathbf{p}\| \|\mathbf{q}\| \sin(\theta_{\mathbf{pq}}) \mathbf{k}$$

la cui lunghezza è pari all'area del parallelogramma formato dai due vettori. Il prodotto vettoriale verifica le seguenti condizioni:

1. $\mathbf{p} \times \mathbf{p} = \mathbf{0}$; (il parallelogramma ha area 0)

2. $\mathbf{p} \times \mathbf{q} = -\mathbf{q} \times \mathbf{p}$; (lo scambio dei due vettori fa passare dalla base $\mathbf{p}, \mathbf{q}, \mathbf{k}$ alla base $\mathbf{q}, \mathbf{p}, \mathbf{k}$ e quindi cambia l'orientazione rispetto alla base canonica)
3. $s(\mathbf{p} \times \mathbf{q}) = (s\mathbf{p}) \times \mathbf{q} = \mathbf{p} \times (s\mathbf{q})$; (allungare o accorciare un lato del parallelogramma di un fattore scalare s ha l'effetto di modificare la corrispondente area dello stesso fattore)
4. $\mathbf{p} \times (\mathbf{q} + \mathbf{r}) = \mathbf{p} \times \mathbf{q} + \mathbf{p} \times \mathbf{r}$; (La distributività $\mathbf{p} \times (\mathbf{q} + \mathbf{r}) = \mathbf{p} \times \mathbf{q} + \mathbf{p} \times \mathbf{r}$ vale perché si ha (si consulti la figura a pagina 97 dove $\alpha = \theta_{\mathbf{p}(\mathbf{q}+\mathbf{r})}$, $\beta = \theta_{\mathbf{pq}}$ e $\gamma = \theta_{\mathbf{pr}}$): $\|\mathbf{q} + \mathbf{r}\| \sin(\theta_{\mathbf{p}(\mathbf{q}+\mathbf{r})}) = \|\mathbf{q}\| \sin(\theta_{\mathbf{pq}}) + \|\mathbf{r}\| \sin(\theta_{\mathbf{pr}})$.)
5. Il prodotto vettoriale non è associativo ma soddisfa l'identità di Jacobi: $\mathbf{p} \times (\mathbf{q} \times \mathbf{r}) + \mathbf{q} \times (\mathbf{r} \times \mathbf{p}) + \mathbf{r} \times (\mathbf{p} \times \mathbf{q}) = 0$.

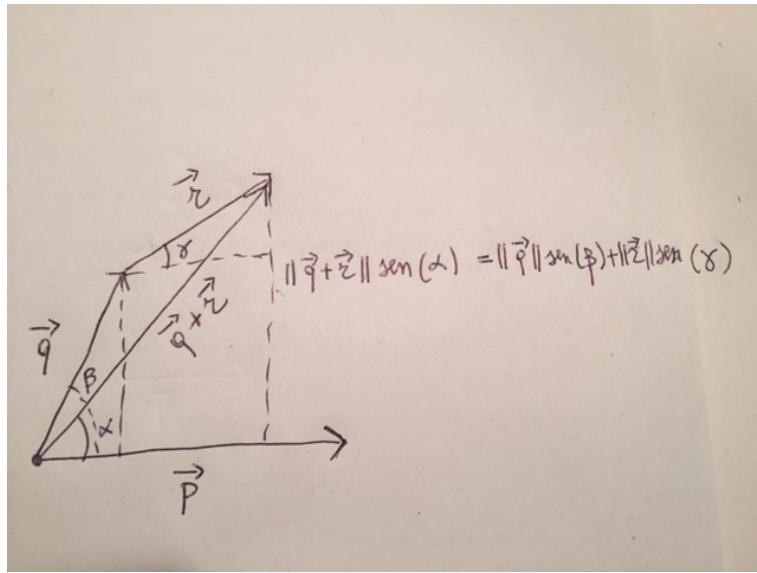


Figure 9.4: Distributività del prodotto vettoriale

Definition 9.2.1. Il determinante della matrice A di tipo 3×3 formata dai vettori colonna $A^1 = \mathbf{p}$, $A^2 = \mathbf{q}$ e $A^3 = \mathbf{r}$ è uguale a $(\mathbf{p} \times \mathbf{q}) \cdot \mathbf{r}$:

$$\det \begin{bmatrix} p_1 & q_1 & r_1 \\ p_2 & q_2 & r_2 \\ p_3 & q_3 & r_3 \end{bmatrix} = (\mathbf{p} \times \mathbf{q}) \cdot \mathbf{r}$$

dove $\mathbf{p} = [p_1, p_2, p_3]^t$, $\mathbf{q} = [q_1, q_2, q_3]^t$ e $\mathbf{r} = [r_1, r_2, r_3]^t$.

Calcoliamo le coordinate del vettore $\mathbf{p} \times \mathbf{q}$ utilizzando la proprietà distributiva del prodotto vettoriale rispetto alla somma vettoriale:

$$\begin{aligned}\mathbf{p} \times \mathbf{q} &= (p_1\mathbf{e}_1 + p_2\mathbf{e}_2 + p_3\mathbf{e}_3) \times (q_1\mathbf{e}_1 + q_2\mathbf{e}_2 + q_3\mathbf{e}_3) \\ &= p_1\mathbf{e}_1 \times (q_2\mathbf{e}_2 + q_3\mathbf{e}_3) + p_2\mathbf{e}_2 \times (q_1\mathbf{e}_1 + q_3\mathbf{e}_3) + p_3\mathbf{e}_3 \times (q_1\mathbf{e}_1 + q_2\mathbf{e}_2) \\ &= (p_1q_2 - p_2q_1)(\mathbf{e}_1 \times \mathbf{e}_2) + (p_1q_3 - p_3q_1)(\mathbf{e}_1 \times \mathbf{e}_3) + (p_2q_3 - p_3q_2)(\mathbf{e}_2 \times \mathbf{e}_3) \\ &= (p_1q_2 - p_2q_1)\mathbf{e}_3 + (p_1q_3 - p_3q_1)(-\mathbf{e}_2) + (p_2q_3 - p_3q_2)\mathbf{e}_1 \\ &= (p_1q_2 - p_2q_1)\mathbf{e}_3 - (p_1q_3 - p_3q_1)\mathbf{e}_2 + (p_2q_3 - p_3q_2)\mathbf{e}_1\end{aligned}$$

La seguente tabella spiega il prodotto vettoriale dei vettori della base canonica di \mathbb{R}^3 :

\times	\mathbf{e}_1	\mathbf{e}_2	\mathbf{e}_3
\mathbf{e}_1	$\mathbf{0}$	\mathbf{e}_3	$-\mathbf{e}_2$
\mathbf{e}_2	$-\mathbf{e}_3$	$\mathbf{0}$	\mathbf{e}_1
\mathbf{e}_3	\mathbf{e}_2	$-\mathbf{e}_1$	$\mathbf{0}$

Remark 2. Il piano parametrico $r\mathbf{p} + s\mathbf{q}$ ($r, s \in \mathbb{R}$) passante per l'origine degli assi è definito dalla seguente equazione lineare:

$$(p_2q_3 - p_3q_2)x - (p_1q_3 - p_3q_1)y + (p_1q_2 - p_2q_1)z = 0.$$

Quindi il vettore di coordinate $[p_2q_3 - p_3q_2, p_1q_3 - p_3q_1, p_1q_2 - p_2q_1]$ è perpendicolare al piano generato dai vettori \mathbf{p} e \mathbf{q} .

Infine il determinante della matrice calcolato rispetto alla terza colonna è:

$$\mathbf{r} \cdot (\mathbf{p} \times \mathbf{q}) = r_1(p_2q_3 - p_3q_2) - r_2(p_1q_3 - p_3q_1) + r_3(p_1q_2 - p_2q_1) = \mathbf{p} \cdot (\mathbf{q} \times \mathbf{r}) = \mathbf{q} \cdot (\mathbf{p} \times \mathbf{r})$$

Il determinante è lineare in \mathbf{p} , \mathbf{q} ed \mathbf{r} . Inoltre se due dei tre vettori sono uguali il determinante è nullo.

Nella sezione seguente studieremo il determinante in generale.

9.3 Il determinante di una matrice quadrata di ordine n

Sia A una matrice quadrata di ordine n sul campo numerico \mathbb{K} . Ad A possiamo associare uno scalare, il suo *determinante* $\det(A) \in \mathbb{K}$. Per comodità

9.3. IL DETERMINANTE DI UNA MATRICE QUADRATA DI ORDINE N99

a volte denotiamo il determinante tramite le colonne (oppure le righe) della matrice A : $\det(A^1, \dots, A^n)$. Il determinante è univocamente determinato dalle seguenti tre proprietà:

(i) Il determinante come funzione di una colonna è lineare:

$$\det(A^1, \dots, B+C, \dots, A^n) = \det(A^1, \dots, B, \dots, A^n) + \det(A^1, \dots, C, \dots, A^n);$$

$$\det(A^1, \dots, cA^j, \dots, A^n) = c \cdot \det(A^1, \dots, A^j, \dots, A^n).$$

(ii) Se due colonne sono uguali, cioè $A^i = A^j$ ($i \neq j$), allora $\det(A) = 0$.

(iii) $\det(I_n) = 1$, dove I_n è la matrice identica.

Per semplicità utilizzeremo talvolta la notazione $|A|$ al posto di $\det(A)$.

Example 50. Il determinante di una matrice diagonale è il prodotto degli elementi della diagonale. Nel caso di dimensione 2 rappresenta l'area di un rettangolo con segno:

$$\begin{vmatrix} a & 0 \\ 0 & b \end{vmatrix} =_{(i)} a \begin{vmatrix} 1 & 0 \\ 0 & b \end{vmatrix} =_{(i)} ab \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} =_{(iii)} ab.$$

Analizziamo alcune conseguenze delle tre proprietà (i)-(iii).

Proposition 9.3.1. *Valgono le seguenti proprietà:*

(iv) *Se due colonne sono scambiate, il determinante cambia di segno.*

(v) *Se si somma ad una colonna un multiplo scalare di un'altra colonna il valore del determinante non cambia.*

Proof. (iv) Per semplicità consideriamo le prime due colonne. Si ha $\det(A^1 + A^2, A^1 + A^2, \dots) = 0$ perché due colonne sono uguali. Per linearità si ottiene:

$$\begin{aligned} 0 &= \det(A^1 + A^2, A^1 + A^2, \dots) \\ &= \det(A^1, A^1 + A^2, \dots) + \det(A^2, A^1 + A^2, \dots) \\ &= \det(A^1, A^1, \dots) + \det(A^1, A^2, \dots) + \det(A^2, A^1 + A^2, \dots) \\ &= 0 + \det(A^1, A^2, \dots) + \det(A^2, A^1 + A^2, \dots) \\ &= \det(A^1, A^2, \dots) + \det(A^2, A^1, \dots) + \det(A^2, A^2, \dots) \\ &= \det(A^1, A^2, \dots) + \det(A^2, A^1, \dots) + 0 \\ &= \det(A^1, A^2, \dots) + \det(A^2, A^1, \dots). \end{aligned}$$

da cui si ha la conclusione.

(v) Sia $\det(A) = \det(\dots, A^i, \dots, A^j, \dots)$, mettendo in evidenza le colonne i e j . Sommiamo alla colonna i c volte la colonna j :

$$\begin{aligned}\det(\dots, A^i + cA^j, \dots, A^j, \dots) &= \det(\dots, A^i, \dots, A^j, \dots) + c \cdot \det(\dots, A^j, \dots, A^j, \dots) \\ &= \det(A) + 0 \\ &= \det(A).\end{aligned}$$

□

Example 51. Calcoliamo il determinante di

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

applicando le regole (i)-(vi).

$$|A| = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = \begin{vmatrix} a & b \\ 0 & d \end{vmatrix} + \begin{vmatrix} 0 & b \\ c & d \end{vmatrix}$$

Inoltre,

$$\begin{vmatrix} a & b \\ 0 & d \end{vmatrix} = \begin{vmatrix} a & 0 \\ 0 & d \end{vmatrix} + \begin{vmatrix} a & b \\ 0 & 0 \end{vmatrix} = ad + \frac{b}{a} \begin{vmatrix} a & a \\ 0 & 0 \end{vmatrix} = ad.$$

In maniera simile,

$$\begin{vmatrix} 0 & b \\ c & d \end{vmatrix} = \begin{vmatrix} 0 & 0 \\ c & d \end{vmatrix} + \begin{vmatrix} 0 & b \\ c & 0 \end{vmatrix} = 0 - \begin{vmatrix} b & 0 \\ 0 & c \end{vmatrix} = -bc$$

applicando Proposizione 9.3.1(iv).

Proposition 9.3.2. I vettori colonna A^1, \dots, A^n sono linearmente dipendenti sse $\det(A^1, \dots, A^n) = 0$.

Proof. Se i vettori colonna sono linearmente dipendenti, allora possiamo scrivere un opportuno vettore colonna, per esempio il primo, come combinazione lineare degli altri. Se fosse il primo avremmo: $\det(A^1, \dots, A^n) = \det(c_2 A^2 + \dots + c_n A^n, A^2, \dots, A^n) = 0$ per (i)-(v).

Supponiamo ora che $\det(A^1, \dots, A^n) = 0$ e ipotizziamo per assurdo che i vettori colonna siano linearmente indipendenti. Allora A^1, \dots, A^n generano lo spazio vettoriale \mathbb{K}^n . Sia B una matrice qualsiasi il cui determinante è $\neq 0$. Allora ogni vettore colonna B^i di B è combinazione lineare $B^i = c_1^i A^1 + \dots + c_n^i A^n$. Ne segue che $0 \neq \det(B) = \det(B^1, \dots, B^n) = \det(c_1^1 A^1 + \dots + c_n^1 A^n, \dots, c_1^n A^n + \dots + c_n^n A^n) = 0$ per le regole (i)-(v). Assurdo. \square

Proposizione 9.3.3. *Una matrice quadrata A a coefficienti nel campo \mathbb{K} è invertibile se e solo se il suo determinante è diverso da zero.*

Proof. (\Leftarrow) Dalla Proposizione 9.3.2 le colonne A^1, \dots, A^n sono linearmente indipendenti. Quindi ogni vettore dello spazio vettoriale \mathbb{K}^n si scrive come combinazione lineare di A^1, \dots, A^n . In particolare i vettori $\mathbf{e}_1, \dots, \mathbf{e}_n$ della base canonica di \mathbb{K}^n . Quindi,

$$\mathbf{e}_1 = b_{11}A^1 + \dots + b_{n1}A^n, \dots, \mathbf{e}_n = b_{1n}A^1 + \dots + b_{nn}A^n.$$

e sia B la matrice tale che $B^i = [b_{1i}, \dots, b_{ni}]^t$. Allora la matrice A è la matrice del cambiamento di base dalla base A^1, \dots, A^n alla base $\mathbf{e}_1, \dots, \mathbf{e}_n$, mentre la matrice B è la matrice del cambiamento di base inverso. In conclusione, $I_n = BA$.

(\Rightarrow) Se A è invertibile, allora il sistema lineare $A\mathbf{x} = x_1A^1 + \dots + x_nA^n = \mathbf{b}$ ha un'unica soluzione $\mathbf{x} = A^{-1}\mathbf{b}$. Siccome il vettore \mathbf{b} è arbitrario allora i vettori A^1, \dots, A^n costituiscono una base di V . Quindi sono linearmente indipendenti e dalla Proposizione 9.3.2 il determinante di A è diverso da zero. \square

Proposizione 9.3.4. 1. $|A^t| = |A|$;

$$2. |AB| = |A| |B|;$$

$$3. |A^{-1}| = |A|^{-1}.$$

9.4 Calcolo del determinante

Con la regola di Cramer possiamo risolvere il sistema lineare $A\mathbf{x} = \mathbf{b}$ utilizzando i determinanti.

Theorem 9.4.1. (Regola di Cramer) *Sia A una matrice quadrata di ordine n con determinante diverso da 0. Se \mathbf{b} è un vettore colonna e c_1, \dots, c_n scalari tali che*

$$c_1 A^1 + \cdots + c_n A^n = \mathbf{b}$$

allora per ogni i abbiamo

$$c_i = \frac{\det(A^1, \dots, A^{i-1}, \mathbf{b}, A^{i+1}, \dots, A^n)}{\det(A^1, \dots, A^n)}.$$

Proof. Si sostituisca in $\det(A^1, \dots, A^{i-1}, \mathbf{b}, A^{i+1}, \dots, A^n)$ il vettore \mathbf{b} con la combinazione lineare $c_1 A^1 + \cdots + c_n A^n$ e si applichino le regole di calcolo (i)-(v). \square

Sia $A = (a_{ij})$ una matrice quadrata di ordine n . Fissati i e j indichiamo con A_{ij} la matrice quadrata $(n-1) \times (n-1)$ ottenuta da A cancellando la riga i e la colonna j .

Calcolo del determinante di una matrice quadrata A rispetto alla riga i :

$$|A| = (-1)^{i+1} a_{i1} |A_{i1}| + \cdots + (-1)^{i+n} a_{in} |A_{in}| = \sum_{k=1}^n (-1)^{i+k} |A_{ik}|$$

Calcolo del determinante di una matrice quadrata A rispetto alla colonna j :

$$|A| = (-1)^{j+1} a_{1j} |A_{1j}| + \cdots + (-1)^{j+n} a_{nj} |A_{nj}| = \sum_{k=1}^n (-1)^{j+k} |A_{kj}|$$

Example 52. Il determinante di una matrice di tipo 3×3 sviluppato rispetto alla prima riga:

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = a(ei-fh)-b(di-fg)+c(dh-eg) = aei+bfg+cdh-afh-bdi-ceg$$

Lemma 9.4.2. *Il determinante di una matrice triangolare superiore o inferiore è dato dal prodotto degli elementi nella diagonale.*

Sia A una matrice quadrata. Possiamo semplificare il calcolo del determinante applicando ad A la seguente operazione elementare che non modifica il determinante: sommare ad una data colonna il multiplo scalare di un'altra colonna.

Example 53. Sia $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$ una matrice. Sostituiamo la seconda colonna A^2 con $A^2 - 2A^1$ e la terza colonna A^3 con $A^3 - 3A^1$:

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 0 & 0 \\ 4 & -3 & -6 \\ 7 & -6 & -12 \end{bmatrix} = B$$

Sostituiamo poi la terza colonna $B^3 = \begin{bmatrix} 0 \\ -6 \\ -12 \end{bmatrix}$ di B con $B^3 - 2B^2$:

$$\begin{bmatrix} 1 & 0 & 0 \\ 4 & -3 & -6 \\ 7 & -6 & -12 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 0 & 0 \\ 4 & -3 & 0 \\ 7 & -6 & 0 \end{bmatrix}$$

La matrice finale è triangolare superiore. Il suo determinante è dato dal prodotto degli elementi della diagonale ed è quindi nullo. In conclusione, la matrice A di partenza ha determinante nullo.

Determinante di una trasformazione lineare

Sia $f : V \rightarrow V$ una trasformazione lineare. Sia v_1, \dots, v_n una base e sia A la matrice quadrata di ordine n che rappresenta f rispetto alla base v_1, \dots, v_n .

Se consideriamo un'altra base di V , w_1, \dots, w_n , allora consideriamo la matrice B che rappresenta f rispetto a questa nuova base. Dalla Sezione 8.4.1 sappiamo che esiste una matrice invertibile C che corrisponde al cambiamento di base. Così si ha:

$$B = C^{-1}AC$$

e quindi

$$|B| = |C|^{-1} |A| |C| = |A|.$$

Theorem 9.4.3. *Il determinante di una trasformazione lineare f è indipendente dalla scelta della base (e quindi della matrice che rappresenta f).*

Data una trasformazione lineare f , possiamo scrivere quindi $\det(f)$, intendendo con questo il determinante di una qualsiasi matrice che rappresenta f .

Corollary 9.4.4. Si ha per trasformazioni lineari componibili g, f e per isomorfismi lineari h :

$$\begin{aligned} \det(g \circ f) &= \det(g)\det(f) \\ e \\ 1 &= \det(h^{-1} \circ h) = \det(h^{-1})\det(h) = \det(h)^{-1}\det(h). \end{aligned}$$

9.4.1 Calcolo matrice inversa con il determinante

Metodo di Cramer

E^j è il vettore colonna con 1 nel posto j e 0 in tutti gli altri posti.

Proposition 9.4.5. Sia $A = (a_{ij})$ una matrice quadrata con determinante non nullo. Se $B = (b_{ij})$ è la matrice inversa di A , allora applicando la regola di Cramer al sistema lineare:

$$b_{1j}A^1 + \cdots + b_{nj}A^n = A \begin{bmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{nj} \end{bmatrix} = E^j$$

abbiamo

$$b_{ij} = \frac{\det(A^1, \dots, A^{i-1}, E^j, A^{i+1}, \dots, A^n)}{|A|}.$$

Metodo dei cofattori

La matrice dei cofattori di una matrice quadrata A di ordine n , detta anche matrice dei complementi algebrici, è un'altra matrice quadrata di ordine n il cui elemento nella posizione generica i, j è il cofattore (o complemento algebrico) di A relativo alla posizione i, j :

$$\text{cof}_{ij}(A) = (-1)^{i+j}\det(A_{ij})$$

dove A_{ij} è il minore di A ottenuto cancellando la riga i e la colonna j .

$$\text{cof } A = \begin{pmatrix} \text{cof}_{1,1}(A) & \dots & \text{cof}_{1,n}(A) \\ \vdots & \ddots & \vdots \\ \text{cof}_{n,1}(A) & \dots & \text{cof}_{n,n}(A) \end{pmatrix}$$

La matrice inversa di A è:

$$A^{-1} = \frac{1}{\det(A)} \cdot (\text{cof } A)^T$$

Metodo di Gauss-Jordan

Spieghiamo il metodo con un esempio. Consideriamo la matrice A da invertire:

$$A = \begin{bmatrix} 2 & 1 & 1 \\ 4 & -6 & 0 \\ -2 & 7 & 2 \end{bmatrix}$$

Aggiungiamo la matrice identica in fondo:

$$\begin{bmatrix} 2 & 1 & 1 & 1 & 0 & 0 \\ 4 & -6 & 0 & 0 & 1 & 0 \\ -2 & 7 & 2 & 0 & 0 & 1 \end{bmatrix}$$

L'idea è di eseguire le solite operazioni di triangolarizzazione sulla matrice 3×6 in maniera tale da arrivare ad una matrice

$$\begin{bmatrix} 1 & 0 & 0 & b_{11} & b_{12} & b_{13} \\ 0 & 1 & 0 & b_{21} & b_{22} & b_{23} \\ 0 & 0 & 1 & b_{31} & b_{32} & b_{33} \end{bmatrix}$$

Allora la matrice 3×3 in fondo sarà l'inversa di A .

Cominciamo sottraendo il doppio della prima riga alla seconda riga:

$$\begin{bmatrix} 2 & 1 & 1 & 1 & 0 & 0 \\ 0 & -8 & -2 & -2 & 1 & 0 \\ -2 & 7 & 2 & 0 & 0 & 1 \end{bmatrix}$$

Sommiamo la prima riga all'ultima riga:

$$\begin{bmatrix} 2 & 1 & 1 & 1 & 0 & 0 \\ 0 & -8 & -2 & -2 & 1 & 0 \\ 0 & 8 & 3 & 1 & 0 & 1 \end{bmatrix}$$

Sommiamo la seconda riga all'ultima:

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

Ora cerchiamo di azzerare le componenti 12 e 22, sottraendo la terza riga alla prima:

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

e sommando due volte la terza alla seconda:

$$\begin{bmatrix} 2 & 1 & 0 & 2 & -1 & -1 \\ 0 & -8 & 0 & -4 & 3 & 2 \\ 0 & 0 & 1 & -1 & 1 & 1 \end{bmatrix}$$

Per azzerare la componente 12 si procede sommando un ottavo della seconda alla prima:

$$\begin{bmatrix} 2 & 0 & 0 & \frac{12}{8} & -\frac{5}{8} & -\frac{6}{8} \\ 0 & -8 & 0 & -4 & 3 & 2 \\ 0 & 0 & 1 & -1 & 1 & 1 \end{bmatrix}$$

Infine si divide la prima riga per 2, la seconda per -8:

$$\begin{bmatrix} 1 & 0 & 0 & \frac{12}{16} & -\frac{5}{16} & -\frac{6}{16} \\ 0 & 1 & 0 & \frac{4}{8} & -\frac{3}{8} & -\frac{2}{8} \\ 0 & 0 & 1 & -1 & 1 & 1 \end{bmatrix}$$

Quindi la matrice inversa è:

$$A^{-1} = \begin{bmatrix} \frac{12}{16} & -\frac{5}{16} & -\frac{6}{16} \\ \frac{4}{8} & -\frac{3}{8} & -\frac{2}{8} \\ -1 & 1 & 1 \end{bmatrix}$$

Chapter 10

Autovettori e Autovalori

Una trasformazione lineare di uno spazio vettoriale V in se stesso si chiama endomorfismo lineare.

10.1 Definizione di autovalore e autovettore

Definition 10.1.1. Sia V uno spazio vettoriale ed $f : V \rightarrow V$ un endomorfismo lineare. Un *autovettore* è un vettore non nullo $\mathbf{v} \in V$ per cui esiste uno scalare λ tale che

$$f(\mathbf{v}) = \lambda\mathbf{v}.$$

Lo scalare λ viene detto *autovalore*.

Lo scalare 0 è un autovalore di f sse il nucleo $\ker(f)$ ha dimensione ≥ 1 .

Remark 3. Se V è uno spazio vettoriale sul campo reale, \mathbf{v} è un autovettore di autovalore $\lambda > 0$ sse la trasformazione lineare trasforma la retta $r\mathbf{v}$ ($r \in \mathbb{R}$) in se stessa, dilatandola se $\lambda \geq 1$ oppure contraendola se $0 < \lambda < 1$. Se l'autovalore è negativo abbiamo anche un ribaltamento.

Proposition 10.1.1. *Sia f un endomorfismo lineare dello spazio vettoriale V e sia λ un autovalore di f .*

- (i) *L'insieme E_λ degli autovettori di autovalore λ è un sottospazio vettoriale di V .*
- (ii) *Se $\lambda_1, \dots, \lambda_k$ sono autovalori distinti di f e \mathbf{v}_i è un autovettore dell'autovalore λ_i ($i = 1, \dots, k$), allora i vettori $\mathbf{v}_1, \dots, \mathbf{v}_k$ sono linearmente indipendenti.*

Proof. (i) Siano \mathbf{v} e \mathbf{w} due vettori tali che $f(\mathbf{v}) = \lambda\mathbf{v}$ e $f(\mathbf{w}) = \lambda\mathbf{w}$. Allora

$$f(r\mathbf{v} + s\mathbf{w}) = rf(\mathbf{v}) + sf(\mathbf{w}) = r(\lambda\mathbf{v}) + s(\lambda\mathbf{w}) = \lambda(r\mathbf{v} + s\mathbf{w}).$$

(ii) Per induzione su k . Se $c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k = \mathbf{0}$, allora

$$f(c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k) = c_1\lambda_1\mathbf{v}_1 + \cdots + c_k\lambda_k\mathbf{v}_k = \mathbf{0}.$$

e sostituendo $c_k\mathbf{v}_k = -c_1\mathbf{v}_1 - \cdots - c_{k-1}\mathbf{v}_{k-1}$ si ottiene:

$$c_1\lambda_1\mathbf{v}_1 + \cdots + \lambda_k(-c_1\mathbf{v}_1 - \cdots - c_{k-1}\mathbf{v}_{k-1}) = c_1\mathbf{v}_1(\lambda_1 - \lambda_k) + \cdots + c_{k-1}\mathbf{v}_{k-1}(\lambda_{k-1} - \lambda_k) = \mathbf{0}.$$

che contraddice l'ipotesi induttiva che $\mathbf{v}_1, \dots, \mathbf{v}_{k-1}$ sono linearmente indipendenti. \square

Definition 10.1.2. Il sottospazio vettoriale E_λ degli autovettori di λ si chiama *autospazio*. La dimensione dell'autospazio E_λ si dice *molteplicità geometrica* dell'autovalore λ .

Sia $I : V \rightarrow V$ la funzione identica: $I(\mathbf{v}) = \mathbf{v}$ for every $\mathbf{v} \in V$.

Proposition 10.1.2. *Sia f un endomorfismo lineare dello spazio vettoriale V di dimensione n . Allora le seguenti condizioni sono equivalenti:*

- (i) λ è un autovalore di f ;
- (ii) $\dim \ker(f - \lambda I) \geq 1$;
- (iii) L'endomorfismo lineare $f - \lambda I$ non è invertibile;
- (iv) $\det(f - \lambda I) = 0$.

Proof. λ è un autovalore di f sse esiste un vettore non nullo \mathbf{v} tale che $(f - \lambda I)(\mathbf{v}) = \mathbf{0}$. \square

Se A è la matrice di $f : V \rightarrow V$ rispetto ad una fissata base di V , allora $\det(f - \lambda I) = \det(A - \lambda I_n)$ è un polinomio di grado n nell'incognita λ .

Definition 10.1.3. Il polinomio $p_f(\lambda)$ (di grado n) definito come il determinante $\det(f - \lambda I)$ si chiama *polinomio caratteristico* di f .

Gli autovalori di f sono gli zeri del suo polinomio caratteristico $p_f(\lambda) = 0$. Se V ha dimensione n , allora f ha al più n autovalori.

Il polinomio caratteristico è indipendente dalla scelta della matrice A rappresentante f . Ricordiamo che le matrici A e B sono simili se esiste una matrice invertibile P di cambiamento di base tale che $B = PAP^{-1}$ (si consulti la Definizione 8.4.1). Matrici simili determinano lo stesso polinomio caratteristico e quindi gli stessi autovalori. Infatti si ha:

$$\begin{aligned}\det(B - \lambda I_n) &= \det(PAP^{-1} - \lambda I_n) \\ &= \det(PAP^{-1} - \lambda PI_nP^{-1}) \\ &= \det(P(A - \lambda I_n)P^{-1}) \\ &= \det(P)\det(A - \lambda I_n)\det(P)^{-1} \\ &= \det(A - \lambda I_n).\end{aligned}$$

Il polinomio caratteristico di una matrice quadrata A è il polinomio caratteristico dell'endomorfismo lineare determinato da A .

Si noti che la matrice A e la sua trasposta A^t hanno lo stesso polinomio caratteristico, perché $\det(A - \lambda I_n) = \det(A - \lambda I_n)^t = \det(A^t - \lambda I_n)$.

Proposition 10.1.3. *Il polinomio caratteristico di una matrice quadrata $A = (a_{ij})$ triangolare superiore (o inferiore) è dato da*

$$p(\lambda) = (a_{11} - \lambda)(a_{22} - \lambda) \cdots (a_{nn} - \lambda).$$

Gli autovalori di una matrice triangolare superiore (o inferiore) sono gli elementi della diagonale.

Definition 10.1.4. Sia λ_0 un autovalore di un endomorfismo lineare f . Si dice *molteplicità algebrica* di λ_0 il più grande k tale che il polinomio $(\lambda - \lambda_0)^k$ divide il polinomio caratteristico $p_f(\lambda)$:

$$p_f(\lambda) = (\lambda - \lambda_0)^k q(\lambda),$$

dove $q(\lambda)$ è un opportuno polinomio in λ tale che $q(\lambda_0) \neq 0$.

Proposition 10.1.4. *Sia λ_0 un autovalore di un endomorfismo lineare $f : V \rightarrow V$. Allora la molteplicità geometrica di λ_0 è sempre minore o uguale alla sua molteplicità algebrica.*

Proof. Sia r la dimensione dell'autospazio E_{λ_0} e sia $\mathbf{v}_1, \dots, \mathbf{v}_r$ una base di E_{λ_0} . Si noti che $\mathbf{v}_1, \dots, \mathbf{v}_r$ sono autovettori di autovalore λ_0 . Completiamo

$\mathbf{v}_1, \dots, \mathbf{v}_r$ ad una base $\mathbf{v}_1, \dots, \mathbf{v}_r, \mathbf{v}_{r+1}, \dots, \mathbf{v}_n$ di V . Rispetto a questa base la matrice A di f ha la forma

$$A = \left[\begin{array}{c|c} \lambda_0 I_r & B \\ \hline O_{n-r} & C \end{array} \right]$$

Quindi

$$A - \lambda I_n = \left[\begin{array}{c|c} (\lambda_0 - \lambda) I_r & B \\ \hline O_{n-r} & C - \lambda I_{n-r} \end{array} \right]$$

da cui si ricava $\det(A - \lambda I_n) = (\lambda_0 - \lambda)^r p_C(\lambda)$. Ne segue che la molteplicità geometrica r è minore o uguale a quella algebrica. \square

Example 54. Sia $A = (a_{ij})$ una matrice di tipo 2×2 a coefficienti in un campo \mathbb{K} . Calcoliamo il determinante $\det(A - \lambda I_2)$:

$$\det(A - \lambda I_2) = \begin{vmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{vmatrix} = \lambda^2 - c_1 \lambda + c_0,$$

dove $c_0 = \det(A)$ e $c_1 = \text{tr}(A) = a_{11} + a_{22}$. Allora $\det(A - \lambda I_2) = 0$ sse $\lambda^2 + c_1 \lambda + c_0 = 0$. Quindi per trovare gli autovalori di A bisogna risolvere una equazione di secondo grado nell'incognita λ . Le due soluzioni dell'equazione $\lambda^2 + c_1 \lambda + c_0 = 0$ esistono sempre se $\mathbb{K} = \mathbb{C}$ è il campo dei numeri complessi, mentre possono o non essere esistenti nel caso in cui $\mathbb{K} = \mathbb{R}$ è il campo dei numeri reali. Per esempio, l'equazione $\lambda^2 + 1 = 0$ non ammette soluzioni reali.

Example 55. Nel caso di una matrice $A = (a_{ij})$ di tipo 3×3 a coefficienti in un campo \mathbb{K} si ha:

$$\det(A - \lambda I_3) = \begin{vmatrix} a_{11} - \lambda & a_{12} & a_{13} \\ a_{21} & a_{22} - \lambda & a_{23} \\ a_{31} & a_{32} & a_{33} - \lambda \end{vmatrix} = -\lambda^3 + c_2 \lambda^2 - c_1 \lambda + c_0,$$

dove

$$c_0 = \det(A); \quad c_2 = \text{tr}(A) = a_{11} + a_{22} + a_{33}$$

e

$$c_1 = \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}.$$

Allora $\det(A - \lambda I_3) = 0$ sse $-\lambda^3 + c_2 \lambda^2 + c_1 \lambda + c_0 = 0$. L'equazione di terzo grado $-\lambda^3 + c_2 \lambda^2 + c_1 \lambda + c_0 = 0$ ha sempre tre soluzioni z_1, z_2, z_3 (contate con

la dovuta molteplicità) nel campo dei numeri complessi: se le tre soluzioni non sono tutte reali, allora una soluzione è reale, mentre le altre due sono complesse coniugate.

Remark 4. Se consideriamo una rotazione $g : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ del piano reale (si veda il prossimo capitolo) di un angolo $0 < \theta < \frac{\pi}{2}$ in senso antiorario, nessun vettore mantiene la stessa direzione e quindi non devono esistere autovalori. La matrice della rotazione rispetto alla base canonica è (si consulti il Capitolo 11):

$$A = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}$$

Calcoliamo il polinomio caratteristico che è il determinante della matrice

$$\begin{vmatrix} \cos \theta - \lambda & -\sin \theta \\ \sin \theta & \cos \theta - \lambda \end{vmatrix} = (\cos \theta - \lambda)^2 + \sin^2 \theta = \lambda^2 - 2 \cos \theta \lambda + 1.$$

Se calcoliamo le soluzioni troviamo

$$\lambda = \cos \theta \pm \sqrt{\cos^2 \theta - 1} = \cos \theta \pm \sqrt{-\sin^2 \theta} = \cos \theta \pm i \sin \theta,$$

dove i è l'unità immaginaria. Quindi gli autovalori sono numeri complessi non reali. Siccome ogni polinomio ha radici nel campo dei numeri complessi, mentre non è detto che ne abbia nel campo dei numeri reali, conviene spesso lavorare direttamente con trasformazioni lineari complesse. Consideriamo quindi la matrice A come una matrice che definisce una endomorfismo lineare complesso $h : \mathbb{C}^2 \rightarrow \mathbb{C}^2$ rispetto alla base canonica $\mathbf{e}_1 = (1, 0)$ e $\mathbf{e}_2 = (0, 1)$. Calcoliamo l'autospazio associato all'autovalore $\lambda_1 = \cos \theta + i \sin \theta$. Dobbiamo risolvere il sistema lineare omogeneo

$$(A - \lambda_1 I_2) \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -i \sin \theta & -\sin \theta \\ \sin \theta & -i \sin \theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -ix \sin \theta - y \sin \theta \\ x \sin \theta - iy \sin \theta \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

dove x e y sono numeri complessi. Otteniamo l'equazione della retta complessa che descrive i vettori dell'autospazio di λ_1 : $-ix \sin \theta - y \sin \theta = 0$, che dividendo per $\sin \theta$ diventa $y = -ix$. Un autovettore è il vettore $\begin{bmatrix} -1 \\ i \end{bmatrix}$. Analogamente, possiamo calcolare la retta complessa $y = ix$ degli autovettori di autovalore $\lambda_2 = \cos \theta - i \sin \theta$. Un autovettore di autovalore λ_2 è il vettore $\begin{bmatrix} 1 \\ i \end{bmatrix}$.

Example 56. Consideriamo una trasformazione lineare $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ che è una rotazione attorno all'asse y di un angolo π . Allora $f(\mathbf{e}_1) = -\mathbf{e}_1$, $f(\mathbf{e}_2) = \mathbf{e}_2$ e $f(\mathbf{e}_3) = -\mathbf{e}_3$. La matrice della rotazione è:

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

ed abbiamo sicuramente: (i) l'autovalore $\lambda = 1$ con autovettore \mathbf{e}_2 ; (ii) l'autovalore $\lambda = -1$ con autovettori \mathbf{e}_1 ed \mathbf{e}_3 . Il lettore provi che non vi sono altri autovalori.

Example 57. Sia

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

Calcoliamo il polinomio caratteristico di A :

$$\begin{vmatrix} 1 - \lambda & 2 \\ 3 & 4 - \lambda \end{vmatrix} = (1 - \lambda)(4 - \lambda) - 6 = \lambda^2 - 5\lambda - 2.$$

Le soluzioni dell'equazione $\lambda^2 - 5\lambda - 2 = 0$ sono: $\lambda = \frac{5 \pm \sqrt{33}}{2}$. Quindi la matrice ha due autovalori.

Calcoliamo ora l'autospazio associato all'autovalore $\lambda = \frac{5 + \sqrt{33}}{2}$. Dobbiamo risolvere il sistema omogeneo associato alla matrice

$$\begin{bmatrix} 1 - (\frac{5 + \sqrt{33}}{2}) & 2 \\ 3 & 4 - (\frac{5 + \sqrt{33}}{2}) \end{bmatrix} = \begin{bmatrix} \frac{-3 - \sqrt{33}}{2} & 2 \\ 3 & \frac{3 - \sqrt{33}}{2} \end{bmatrix}$$

Otteniamo quindi due rette $-(3 + \sqrt{33})x + 4y = 0$ e $6x + (3 - \sqrt{33})y = 0$ passanti per l'origine. Moltiplicando la prima equazione per $3 - \sqrt{33}$ si ottiene: $24x + 4(3 - \sqrt{33})y = 0$ e dividendo per 4 si ottiene infine la seconda equazione. Quindi le due equazioni definiscono la stessa retta per l'origine. Se poniamo $x = 1$ allora $y = \frac{9}{\sqrt{33}}$. Quindi l'autospazio associato all'autovalore $\frac{5 + \sqrt{33}}{2}$ è il sottospazio vettoriale del piano generato dal vettore $[1, \frac{9}{\sqrt{33}}]$.

Example 58. Sia

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

Il determinante della matrice A è nullo: $\det(A) = -2 \begin{vmatrix} 1 & 1 \\ 2 & 2 \end{vmatrix} = 0$. Ne segue che il sistema omogeneo $A\mathbf{x} = \mathbf{0}$ ammette soluzioni $\mathbf{x} \neq \mathbf{0}$ non nulle. Calcoliamo il polinomio caratteristico di A sviluppando il determinante rispetto alla prima riga e scopriremo che 0 sarà un autovalore della matrice A :

$$\begin{aligned} A &= \begin{vmatrix} 1-\lambda & 0 & 1 \\ 2 & -\lambda & 2 \\ 2 & 2 & 3-\lambda \end{vmatrix} = (1-\lambda) \begin{vmatrix} -\lambda & 2 \\ 2 & 3-\lambda \end{vmatrix} + \begin{vmatrix} 2 & -\lambda \\ 2 & 2 \end{vmatrix} = \\ &= (1-\lambda)(-\lambda(3-\lambda)-4) + 4 + 2\lambda = -\lambda^3 + 4\lambda^2 + 3\lambda. \end{aligned}$$

L'equazione $\lambda^3 - 4\lambda^2 - 3\lambda = 0$ ammette come soluzioni gli autovalori: $\lambda = 0$ e $\lambda = 2 \pm \sqrt{7}$.

10.2 Matrici diagonalizzabili

Se una matrice è diagonale, allora il vettore della base canonica $\mathbf{e}_i = [0, \dots, 1, \dots, 0]^t$ è un autovettore di autovalore a_{ii} :

$$\begin{bmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_{nn} \end{bmatrix} \begin{bmatrix} 0 \\ \dots \\ 1 \\ \dots \\ 0 \end{bmatrix} = a_{ii} \begin{bmatrix} 0 \\ \dots \\ 1 \\ \dots \\ 0 \end{bmatrix}$$

Ogni matrice diagonale dilata (o contrae) il vettore \mathbf{e}_i della base canonica di $a_{ii} \geq 1$ ($0 < a_{ii} < 1$). Se il segno di a_{ii} è negativo, si ha anche un ribaltamento.

Per esempio, abbiamo:

$$\begin{bmatrix} 2 & 0 \\ 0 & 5 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ 0 \end{bmatrix}; \quad \begin{bmatrix} 2 & 0 \\ 0 & 5 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 5 \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Proposition 10.2.1. *Sia $f : V \rightarrow V$ un endomorfismo lineare e $\mathbf{v}_1, \dots, \mathbf{v}_n$ una base di V . Allora gli elementi della base sono autovettori di f sse la matrice di f rispetto alla base $\mathbf{v}_1, \dots, \mathbf{v}_n$ è diagonale.*

Definition 10.2.1. Un endomorfismo lineare $f : V \rightarrow V$ è *diagonalizzabile* se esiste una base di V costituita da autovettori di f .

Corollary 10.2.2. Una matrice quadrata A di ordine n (ad elementi nel campo \mathbb{K}) è diagonalizzabile sse lo spazio vettoriale \mathbb{K}^n ha una base $\mathbf{v}_1, \dots, \mathbf{v}_n$ di autovettori di A . In tal caso, se P è la matrice del cambiamento di base dalla base $\mathbf{v}_1, \dots, \mathbf{v}_n$ alla base canonica abbiamo che $P^{-1}AP$ è una matrice diagonale.

Corollary 10.2.3. Se una matrice quadrata di ordine n ha n autovalori distinti, allora è diagonalizzabile.

Proof. Siano $\lambda_1, \dots, \lambda_n$ gli n autovalori distinti e sia \mathbf{v}_i un autovettore di λ_i per $i = 1, \dots, n$. Dalla Proposizione 10.1.1(ii) i vettori $\mathbf{v}_1, \dots, \mathbf{v}_n$ costituiscono una base di autovettori dello spazio vettoriale. La conclusione segue dal Corollario 10.2.2. \square

Sia $D = P^{-1}AP$ la matrice diagonale del Corollario 10.2.2. Allora si ha:

$$A = PDP^{-1}$$

Se A è diagonalizzabile e si conoscono la matrice invertibile P e la matrice diagonale D , allora è facile calcolare le potenze della matrice A : $A^2 = AA = P^{-1}DDP = P^{-1}D^2P$, $A^3 = AAA = P^{-1}D^3P$, etc. Se

$$D = \begin{bmatrix} d_1 & 0 & \dots & 0 \\ 0 & d_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & d_n \end{bmatrix}$$

allora

$$D^k = \begin{bmatrix} d_1^k & 0 & \dots & 0 \\ 0 & d_2^k & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & d_n^k \end{bmatrix}$$

Il teorema seguente precisa meglio il risultato della Proposizione 10.2.1.

Theorem 10.2.4. Sia V uno spazio vettoriale di dimensione n sul campo \mathbb{K} , sia $f : V \rightarrow V$ un endomorfismo lineare e siano $\lambda_1, \dots, \lambda_k$ tutti gli autovalori distinti di f . Sia n_{alg}^i (resp. n_{gm}^i) la molteplicità algebrica (geometrica) dell'autovalore λ_i per $i = 1, \dots, k$.

Allora le seguenti condizioni sono equivalenti:

- (i) f è diagonalizzabile;
- (ii) Per ogni vettore $\mathbf{v} \in V$, esistono (unici) autovettori \mathbf{w}_i di autovalore λ_i ($i = 1, \dots, k$) tale che $\mathbf{v} = \mathbf{w}_1 + \dots + \mathbf{w}_k$;
- (iii) $n_{gm}^1 + \dots + n_{gm}^k = n$;
- (iv) f ha tutti gli autovalori nel campo \mathbb{K} e $n_{alg}^i = n_{gm}^i$ per ogni $i = 1, \dots, k$.

Proof. (i) \Rightarrow (ii) Per ipotesi V ammette una base di autovettori.

(ii) \Rightarrow (iii) Per ipotesi la somma delle dimensioni degli autospazi è n :

$$\dim E_{\lambda_1} + \dots + \dim E_{\lambda_k} = n.$$

Infine $n_{gm}^i = \dim E_{\lambda_i}$.

(iii) \Rightarrow (iv) Da $n_{gm}^i \leq n_{alg}^i$.

(iv) \Rightarrow (i) Ovvio. \square

Sia $f : V \rightarrow V$ diagonalizzabile ed A la matrice di f rispetto ad una base $\mathbf{v}_1, \dots, \mathbf{v}_n$ di V . Sia $\mathbf{w}_1, \dots, \mathbf{w}_n$ una base di autovettori di f , e P la matrice tale che P^i è il vettore colonna delle coordinate di \mathbf{w}_i rispetto alla base $\mathbf{v}_1, \dots, \mathbf{v}_n$. Allora abbiamo $P^{-1}AP = D$ è una matrice diagonale. Le n colonne di P possono essere suddivise in k insiemi di cardinalità $\dim E_{\lambda_1}, \dots, \dim E_{\lambda_k}$. I vettori colonna del primo insieme sono una base del sottospazio E_{λ_1} , e così via per gli altri insiemi. Gli elementi della diagonale di D sono gli autovalori. Un autovalore occorrerà tante volte in D quanto la dimensione del suo autospazio.

Example 59. Calcoliamo gli autovettori ed autovalori della matrice $A = \begin{bmatrix} 2 & 1 \\ 3 & 2 \end{bmatrix}$. Se $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ è un autovettore di autovalore λ abbiamo

$$(A - \lambda I)\mathbf{x} = \mathbf{0}.$$

Il precedente sistema lineare omogeneo ha soluzione non nulla sse

$$\det(A - \lambda I_2) = \begin{vmatrix} 2 - \lambda & 1 \\ 3 & 2 - \lambda \end{vmatrix} = (2 - \lambda)^2 - 3 = 0.$$

L'equazione di secondo grado

$$\lambda^2 - 4\lambda + 1 = 0$$

ha soluzioni

$$\lambda_1 = 2 + \sqrt{3}; \quad \lambda_2 = 2 - \sqrt{3}.$$

Le due matrici che si ottengono sono:

$$B = A - (2 + \sqrt{3})I = \begin{bmatrix} -\sqrt{3} & 1 \\ 3 & -\sqrt{3} \end{bmatrix}; \quad C = A - (2 - \sqrt{3})I = \begin{bmatrix} \sqrt{3} & 1 \\ 3 & \sqrt{3} \end{bmatrix}.$$

Ora calcoliamo gli autovettori. Per $\lambda_1 = 2 + \sqrt{3}$ lo spazio degli autovettori, che è lo spazio delle soluzioni del sistema omogeneo $B\mathbf{x} = \mathbf{0}$, ha dimensione 1 ed è descritto dalla retta $y = \sqrt{3}x$. Per $\lambda = 2 - \sqrt{3}$ lo spazio degli autovettori, che è lo spazio delle soluzioni del sistema omogeneo $C\mathbf{x} = \mathbf{0}$, ha dimensione 1 ed è descritto dalla retta $y = -\sqrt{3}x$.

La matrice A è quindi diagonalizzabile perché esiste una base di autovettori $\mathbf{v}_1 = [1, \sqrt{3}]$ e $\mathbf{v}_2 = [1, -\sqrt{3}]$. La matrice del cambiamento di base è

$$P = \begin{bmatrix} 1 & 1 \\ \sqrt{3} & -\sqrt{3} \end{bmatrix}$$

e $P^{-1}AP$ è la matrice diagonale con gli autovalori nella diagonale.

Calcoliamo l'inversa di P :

$$P^{-1} = \frac{1}{\det(P)} \begin{bmatrix} -\sqrt{3} & -\sqrt{3} \\ -1 & 1 \end{bmatrix}^t = \begin{bmatrix} 1/2 & 1/2 \\ \frac{1}{2\sqrt{3}} & \frac{-1}{2\sqrt{3}} \end{bmatrix}^t = \begin{bmatrix} 1/2 & \frac{1}{2\sqrt{3}} \\ 1/2 & \frac{-1}{2\sqrt{3}} \end{bmatrix}$$

Allora

$$\begin{aligned} P^{-1}AP &= \begin{bmatrix} 1/2 & \frac{1}{2\sqrt{3}} \\ 1/2 & \frac{-1}{2\sqrt{3}} \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ \sqrt{3} & -\sqrt{3} \end{bmatrix} = \\ &= \begin{bmatrix} 1/2 & \frac{1}{2\sqrt{3}} \\ 1/2 & \frac{-1}{2\sqrt{3}} \end{bmatrix} \begin{bmatrix} 2 + \sqrt{3} & 2 - \sqrt{3} \\ 3 + 2\sqrt{3} & 3 - 2\sqrt{3} \end{bmatrix} = \begin{bmatrix} 2 + \sqrt{3} & 0 \\ 0 & 2 - \sqrt{3} \end{bmatrix} \end{aligned}$$

Example 60. Sia $A = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 2 & 0 \\ 1 & 0 & 2 \end{bmatrix}$ una matrice. Vogliamo

1. Determinare gli autovalori di A e le relative molteplicità.
2. Determinare gli autospazi di A e trovare, se esiste, una base di \mathbb{R}^3 formata da autovettori di A .

3. Calcolare una matrice P invertibile tale che $P^{-1}AP$ sia diagonale.

La matrice A è triangolare inferiore. Quindi gli autovalori sono gli elementi sulla diagonale: l'autovalore 1 con molteplicità 1, e l'autovalore 2 con molteplicità 2. Un altro modo per calcolare gli autovalori è tramite il polinomio caratteristico:

$$\begin{vmatrix} 1 - \lambda & 0 & 0 \\ -1 & 2 - \lambda & 0 \\ 1 & 0 & 2 - \lambda \end{vmatrix} = (1 - \lambda)(2 - \lambda)^2.$$

Le radici del polinomio caratteristico $(1 - \lambda)(2 - \lambda)^2 = 0$ sono 1, 2, 2.

Per trovare gli autospazi bisogna risolvere i sistemi lineari omogenei

$$\begin{bmatrix} 1 - 1 & 0 & 0 \\ -1 & 2 - 1 & 0 \\ 1 & 0 & 2 - 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

e

$$\begin{bmatrix} 1 - 2 & 0 & 0 \\ -1 & 2 - 2 & 0 \\ 1 & 0 & 2 - 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 \\ -1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Il primo sistema equivale a $x_2 = x_1$ e $x_3 = -x_1$. L'autospazio delle soluzioni ha come base il vettore $(1, 1, -1)$. Il secondo sistema ha per soluzione il piano $x_1 = 0$, che ha per base i vettori $[0, 1, 0]$ e $[0, 0, 1]$. Quindi una base di \mathbb{R}^3 fatta di autovettori è composta dai vettori $[1, 1, -1]$, $[0, 1, 0]$ e $[0, 0, 1]$. Rispetto alla base composta dagli autovettori $[1, 1, -1]$, $[0, 1, 0]$ e $[0, 0, 1]$ la matrice diagonale D simile ad A è:

$$D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

La matrice P tale che $D = P^{-1}AP$ è diagonale, è la matrice le cui colonne sono gli autovettori:

$$P = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}$$

Example 61. Siano dati in \mathbb{R}^3 i vettori

$$v_1 = [0, 1, -1]; \quad v_2 = [2, 0, 1]; \quad v_3 = [1, 2, 0].$$

1. Verificare che esiste una unica trasformazione lineare $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ avente $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ come autovettori associati, rispettivamente, agli autovalori $0, 3, 6$.
2. Determinare (a) la matrice A associata ad f ; (b) $\ker(f)$ e $\text{Im}(f)$.

Dobbiamo avere:

$$f(\mathbf{v}_1) = \mathbf{0}; \quad f(\mathbf{v}_2) = 3\mathbf{v}_2; \quad f(\mathbf{v}_3) = 6\mathbf{v}_3.$$

I tre vettori $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ sono linearmente indipendenti e quindi costituiscono una base di \mathbb{R}^3 . Ne segue che la funzione f è unica perché è definita su ogni vettore $\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3$ di \mathbb{R}^3 : $f(c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3) = c_1f(\mathbf{v}_1) + c_2f(\mathbf{v}_2) + c_3f(\mathbf{v}_3) = 3c_2\mathbf{v}_2 + 6c_3\mathbf{v}_3$.

La matrice di f rispetto alla base $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ di autovettori è la matrice diagonale

$$D = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 6 \end{bmatrix}$$

Per calcolare la matrice A rispetto alla base canonica di \mathbb{R}^3 dobbiamo considerare la matrice P le cui colonne sono le coordinate degli autovettori rispetto alla base canonica:

$$P = \begin{bmatrix} 0 & 2 & 1 \\ 1 & 0 & 2 \\ -1 & 1 & 0 \end{bmatrix}$$

Siccome $D = P^{-1}AP$, ricaviamo che $A = PDP^{-1}$. L'inversa della matrice P è:

$$P^{-1} = \begin{bmatrix} 2/3 & -1/3 & -4/3 \\ 2/3 & -1/3 & -1/3 \\ -1/3 & 2/3 & 2/3 \end{bmatrix}$$

E quindi

$$\begin{aligned} A = PDP^{-1} &= \begin{bmatrix} 0 & 2 & 1 \\ 1 & 0 & 2 \\ -1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 6 \end{bmatrix} \begin{bmatrix} 2/3 & -1/3 & -4/3 \\ 2/3 & -1/3 & -1/3 \\ -1/3 & 2/3 & 2/3 \end{bmatrix} = \\ &= \begin{bmatrix} 2 & 2 & 2 \\ -4 & 8 & 8 \\ 2 & -1 & -1 \end{bmatrix} \end{aligned}$$

10.3 Algoritmo di Google

Questo esempio è la terza lezione nel sito Winter2009/RalucaRemus (che troverete con un motore di ricerca).

Supponiamo di avere quattro siti web (www.page1.com, www.page2.com, www.page3.com, www.page4.com) con links tra i siti descritti dal seguente grafo: Il nodo 1 ha tre archi uscenti. Ciò significa che vi sono nella pagina p_1

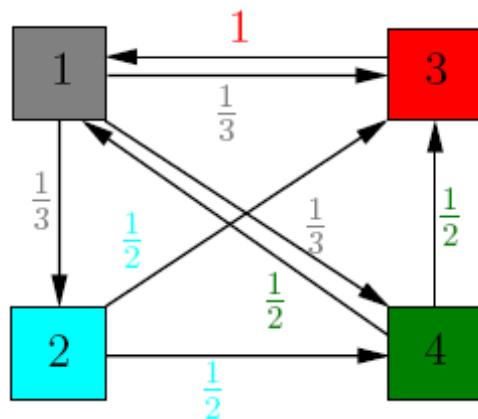


Figure 10.1: Il grafo con archi pesati

tre link, il primo diretto alla pagina p_2 , il secondo alla pagina p_3 e l'ultimo alla pagina p_4 . Similmente per gli altri archi uscenti dagli altri nodi.

Supponiamo che un utente si trovi nella pagina p_1 . Immaginiamo che gli eventi “passa alla pagina p_2 ”, “passa alla pagina p_3 ”, “passa alla pagina p_4 ” siano equiprobabili. Quindi se un utente si troverà nella pagina p_1 vi è una probabilità $\frac{1}{3}$ che passi alla pagina p_2 , e così via. Lo stesso discorso si applica per gli altri nodi con probabilità possibilmente diverse, che dipendono dal numero di links.

Siccome il nodo 1 ha tre archi uscenti, trasferisce un terzo della sua importanza a ciascuno dei tre nodi riceventi. In generale, se un nodo ha k archi uscenti, trasferisce $\frac{1}{k}$ della sua importanza a ciascuno dei nodi riceventi.

I principi sui quali si basa PageRank sono quindi i seguenti:

- Una pagina importante riceve links da pagine importanti.
- Una pagina importante ha pochi links verso altre pagine.

Questi principi vengono formalizzati nella seguente formula: indicando con $r(p)$ il rango della pagina web p (cioé la sua importanza relativa) e con $|p|$ il numero di links dalla pagina p verso altre pagine, abbiamo

$$r(p) = \sum_{q \rightarrow p} \frac{r(q)}{|q|}.$$

In questa formula, la somma è effettuata su tutte le pagine q che hanno un link verso p . Il contributo di una pagina q è quindi direttamente proporzionale all'importanza (rango) di q ed inversamente proporzionale al numero di links da q verso altre pagine.

La matrice A di transizione del grafo mette in ciascuna colonna A^i il trasferimento di importanza dal nodo i agli altri nodi, mentre ciascuna riga A_i della matrice rappresenta l'importanza che il nodo i riceve dagli altri nodi. In altre parole, si ha:

$$a_{ij} = \begin{cases} \frac{1}{|p_j|} & \text{se esiste un link da } p_j \text{ a } p_i \\ 0 & \text{altrimenti.} \end{cases}$$

e quindi nel nostro esempio abbiamo

$$A = \begin{bmatrix} 0 & 0 & 1 & \frac{1}{2} \\ \frac{1}{3} & 0 & 0 & 0 \\ \frac{1}{3} & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{2} & 0 & 0 \end{bmatrix}$$

La matrice A è invertibile. Sviluppiamo il determinante rispetto alla terza colonna.

$$\det(A) = \det\left(\begin{bmatrix} \frac{1}{3} & 0 & 0 \\ \frac{1}{3} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{2} & 0 \end{bmatrix}\right) = \frac{1}{3} \cdot \left(-\frac{1}{4}\right) = -\frac{1}{12}.$$

Denotiamo con $\mathbf{r} = \begin{bmatrix} r(p_1) \\ r(p_2) \\ r(p_3) \\ r(p_4) \end{bmatrix}$ il rango delle quattro pagine. Allora dobbiamo avere:

$$A\mathbf{r} = \begin{bmatrix} 0 & 0 & 1 & \frac{1}{2} \\ \frac{1}{3} & 0 & 0 & 0 \\ \frac{1}{3} & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{2} & 0 & 0 \end{bmatrix} \begin{bmatrix} r(p_1) \\ r(p_2) \\ r(p_3) \\ r(p_4) \end{bmatrix} = \begin{bmatrix} \sum_{q \rightarrow p_1} \frac{r(q)}{|q|} \\ \sum_{q \rightarrow p_2} \frac{r(q)}{|q|} \\ \sum_{q \rightarrow p_3} \frac{r(q)}{|q|} \\ \sum_{q \rightarrow p_4} \frac{r(q)}{|q|} \end{bmatrix} = \begin{bmatrix} r(p_1) \\ r(p_2) \\ r(p_3) \\ r(p_4) \end{bmatrix}$$

In altre parole, il vettore \mathbf{r} è un autovettore di autovalore 1. Risolvendo il sistema lineare corrispondente

$$\begin{aligned} r(p_1) &= r(p_3) + \frac{r(p_4)}{2} \\ r(p_2) &= \frac{r(p_1)}{3} \\ r(p_3) &= \frac{r(p_1)}{3} + \frac{r(p_2)}{2} + \frac{r(p_4)}{2} \\ r(p_4) &= \frac{r(p_1)}{3} + \frac{r(p_2)}{2} \end{aligned}$$

si ottiene facilmente che abbiamo una retta di soluzioni $c(12, 4, 9, 6)$ al variare di c . Prendendo il vettore la cui somma delle coordinate è 1 otteniamo

$$\text{PageRank vector} \equiv \begin{bmatrix} 0.38 \\ 0.12 \\ 0.29 \\ 0.19 \end{bmatrix}$$

In altre parole, la pagina p_1 ha importanza 0.38 e così via per le altre pagine. Su 100 utenti, 38 visiteranno la pagina p_1 .

Si suggerisce di cercare con un motore di ricerca il file jkhoury/Google.pdf dove viene spiegato in dettaglio l'algoritmo di Google.

10.4 Numeri di Fibonacci

Questo esempio è preso da jkhoury/fibonacci.htm

All'inizio dell'anno abbiamo una coppia di conigli maschio e femmina. Le regole sono le seguenti: Dopo due mesi ogni coppia produce una coppia mista (maschio, femmina) e da quel momento una nuova coppia mista ogni mese successivo. Nessun coniglio muore.

Indichiamo con F_n il numero di conigli dopo n mesi. Abbiamo:

$$F_0 = 1; \quad F_1 = 1; \quad F_n = F_{n-1} + F_{n-2}.$$

La successione cresce rapidamente. Dopo 55 mesi abbiamo $F_{55} = 139.583.862.445$.

Esiste un'espressione che ci permette di trovare F_n facilmente? La risposta è positiva se si conosce il processo di diagonalizzazione di una matrice.

Consideriamo la matrice quadrata

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

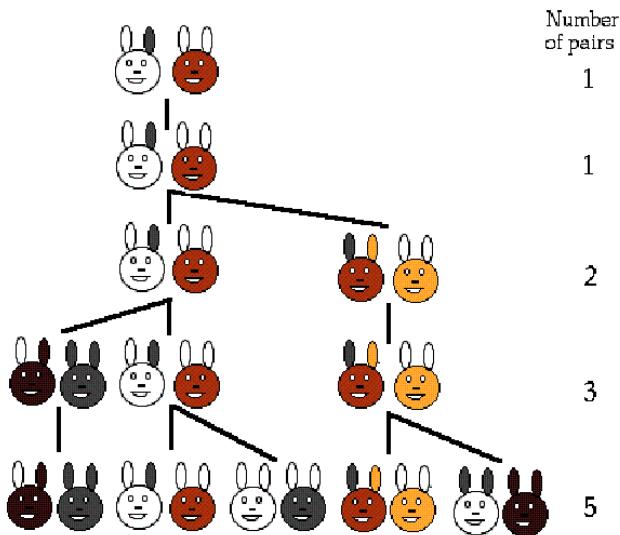


Figure 10.2: Coppie di conigli e successione di Fibonacci

Allora partendo dal vettore

$$\begin{bmatrix} F_1 \\ F_0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

si ricava che

$$\begin{bmatrix} F_{n+1} \\ F_n \end{bmatrix} = A \begin{bmatrix} F_n \\ F_{n-1} \end{bmatrix}$$

o in altri termini

$$\begin{bmatrix} F_{n+1} \\ F_n \end{bmatrix} = AA \cdots A \begin{bmatrix} 1 \\ 1 \end{bmatrix} = A^n \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad (n\text{-volte})$$

La matrice A è diagonalizzabile, ossia esiste una matrice invertibile P ed una matrice diagonale D tale che $A = P^{-1}DP$ da cui si ricava $A^n = P^{-1}D^nP$.

Siccome è facile calcolare una potenza D^n di una matrice diagonale, è anche facile calcolare A^n .

Calcoliamo gli autovalori di A , per cui deve essere $\det(A - \lambda I) = 0$:

$$\det(A - \lambda I) = \det\begin{pmatrix} 1 - \lambda & 1 \\ 1 & -\lambda \end{pmatrix} = -(1 - \lambda)\lambda - 1 = 0.$$

In altri termini,

$$\lambda^2 - \lambda - 1 = 0.$$

Le due soluzioni sono reali

$$\lambda_1 = \frac{1 + \sqrt{5}}{2}, \quad \lambda_2 = \frac{1 - \sqrt{5}}{2}.$$

Se il vettore \mathbf{x} è un autovettore corrispondente all'autovalore λ_i ($i = 1, 2$) allora abbiamo che $A\mathbf{x} = \lambda_i \mathbf{x}$, che si può scrivere come $(A - \lambda_i I)\mathbf{x} = \mathbf{0}$, che è un sistema omogeneo. Risolvendo tale sistema omogeneo si scopre che il vettore $w_1 = (\frac{1+\sqrt{5}}{2}, 1)$ è una base per lo spazio degli autovettori dell'autovalore $\lambda_1 = \frac{1+\sqrt{5}}{2}$, mentre il vettore $w_2 = (\frac{1-\sqrt{5}}{2}, 1)$ è una base per lo spazio degli autovettori dell'autovalore $\lambda_2 = \frac{1-\sqrt{5}}{2}$. Dal Teorema 10.2.4 si ricava che i vettori w_1, w_2 costituiscono una base di \mathbb{R}^2 .

Sempre dal Teorema 10.2.4 si ricava che la matrice i cui vettori colonna sono w_1 e w_2 è la matrice invertibile che diagonalizza A :

$$P = \begin{bmatrix} \frac{1+\sqrt{5}}{2} & \frac{1-\sqrt{5}}{2} \\ 1 & 1 \end{bmatrix}$$

La matrice inversa è

$$P^{-1} = \begin{bmatrix} \frac{1}{\sqrt{5}} & \frac{-1+\sqrt{5}}{2\sqrt{5}} \\ \frac{-1}{2\sqrt{5}} & \frac{1+\sqrt{5}}{2\sqrt{5}} \end{bmatrix}$$

mentre la matrice diagonale D è

$$D = \begin{bmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{bmatrix}$$

Allora si ha:

$$\begin{bmatrix} F_{n+1} \\ F_n \end{bmatrix} = A^n \begin{bmatrix} 1 \\ 1 \end{bmatrix} = PD^n P^{-1} \begin{bmatrix} 1 \\ 1 \end{bmatrix} =$$

$$\begin{bmatrix} \frac{1+\sqrt{5}}{2} & \frac{1-\sqrt{5}}{2} \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \left(\frac{1+\sqrt{5}}{2}\right)^n & 0 \\ 0 & \left(\frac{1-\sqrt{5}}{2}\right)^n \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{5}} & \frac{-1+\sqrt{5}}{2\sqrt{5}} \\ \frac{-1}{2\sqrt{5}} & \frac{1+\sqrt{5}}{2\sqrt{5}} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Infine moltiplicando le matrici si ricava:

$$F_n = \frac{1}{\sqrt{5}} \left[\left(\frac{1 + \sqrt{5}}{2} \right)^{n+1} - \left(\frac{1 - \sqrt{5}}{2} \right)^{n+1} \right]$$

Un calcolatore potente in poco tempo calcola $F_{100} = 573147844013817084101$.

Il numero

$$\phi = \frac{1 + \sqrt{5}}{2}$$

è la famosa *sezione aurea* o *divina proporzione*. È utilizzato in arte e architettura per dare simmetria alla rappresentazioni figurative geometriche (si consulti il libro di T. Livio: La sezione aurea). Il rettangolo aureo è un rettangolo le cui proporzioni sono basate sulla sezione aurea. Ciò significa che il rapporto $\frac{a}{b}$ fra il lato maggiore a e quello minore b è identico a quello fra il lato minore b e il segmento $a - b$ ottenuto sottraendo b dal lato maggiore a (il che implica che entrambi i rapporti siano ϕ). Quindi se abbiamo un rettangolo aureo di lati a e b con $a > b$ si ha: Se $a > b$ sono lunghezze non nulle, allora

$$\phi = \frac{a+b}{a} = \frac{a}{b} = \frac{b}{a-b}.$$

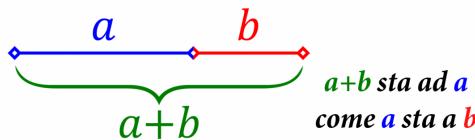


Figure 10.3: Sezione aurea

Siccome $\phi^2 - \phi - 1 = 0$, si ha anche che $\phi = \frac{1+\phi}{\phi}$. Così il rettangolo di lati ϕ e 1 è un rettangolo aureo.

Si ha anche che il rapporto di due numeri di Fibonacci consecutivi tende alla sezione aurea quando l'indice n tende all'infinito.

$$\phi = \lim_{n \rightarrow \infty} \frac{F_n}{F_{n-1}}.$$

10.5 Crittografia

I caratteri dell’alfabeto italiano sono 26. Aggiungiamo un ulteriore carattere che rappresenta il “blank”, lo spazio vuoto. Codifichiamo questi 27 caratteri con dei numeri arbitrari. Per semplificare i conti riduciamo il numero di caratteri a 11 compreso il blank:

$$A = 345; \quad B = 12438; \quad C = 79; \quad D = 987; \quad E = 30078; \quad F = 675;$$

$$G = 5499; \quad I = 9090; \quad O = 555; \quad R = 777; \quad \text{blank} = 647.$$

Allora la frase “GIOCO BARO” viene codificata in maniera elementare dalla seguente successione di numeri:

$$5499, 9090, 555, 79, 555, 647, 12438, 345, 777, 555$$

Consideriamo una matrice Z quadrata di ordine n (n molto grande) che sia invertibile. Immaginiamo che la matrice Z sia conosciuta soltanto ai due interlocutori che devono scambiare il messaggio cifrato. Nel nostro esempio per ragioni di spazio prendiamo una matrice 3×3 :

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

Allora suddividiamo il messaggio con la codifica elementare in vettori di lunghezza 3 avendo l’accortezza di aggiungere degli spazi finale per ottenere un multiplo del 3.

$$\begin{bmatrix} 5499 \\ 9090 \\ 555 \end{bmatrix}, \begin{bmatrix} 79 \\ 555 \\ 647 \end{bmatrix}, \begin{bmatrix} 12438 \\ 345 \\ 777 \end{bmatrix}, \begin{bmatrix} 555 \\ 647 \\ 647 \end{bmatrix}.$$

Mettiamo tutti questi vettori in una matrice U di dimensione 3×4

$$U = \begin{bmatrix} 5499 & 79 & 12438 & 555 \\ 9090 & 555 & 345 & 647 \\ 555 & 647 & 777 & 647 \end{bmatrix}$$

Consideriamo la matrice prodotto

$$ZU = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \begin{bmatrix} 5499 & 79 & 12438 & 555 \\ 9090 & 555 & 345 & 647 \\ 555 & 647 & 777 & 647 \end{bmatrix} =$$

$$\begin{bmatrix} 25344 & 3130 & 15459 & 3790 \\ 70776 & 6973 & 56139 & 9337 \\ 116208 & 10816 & 96819 & 14884 \end{bmatrix}$$

Allora i numeri che vengono trasmessi sono

25344, 3130, 15459, 3790, 70776, 6973, 56139, 9337, 116208, 10816, 96819, 14884

Una persona che intercetta i numeri non riesce a decodificare il messaggio, mentre il ricevente semplicemente moltiplica a sinistra la matrice ZU per Z^{-1} e recupera U .

10.6 Compressione di Immagini

Questa sezione è presa da [jkhouri/haar.htm](#)



Figure 10.4: Compressione di immagine con differenti metodi

Consideriamo una immagine digitale come una matrice. Ogni componente della matrice corrisponde ad un pixel (elemento elementare della figura). Supponiamo di avere una matrice 256×256 di pixels con valori di ciascuna componente un numero da 0 (nero) a 255 (bianco). Nel mezzo varie sfumature di grigio. La tecnica JPEG divide l'immagine in blocchi 8×8 e assegna una matrice ad ogni blocco. Utilizziamo l'algebra lineare per massimizzare la compressione dell'immagine.

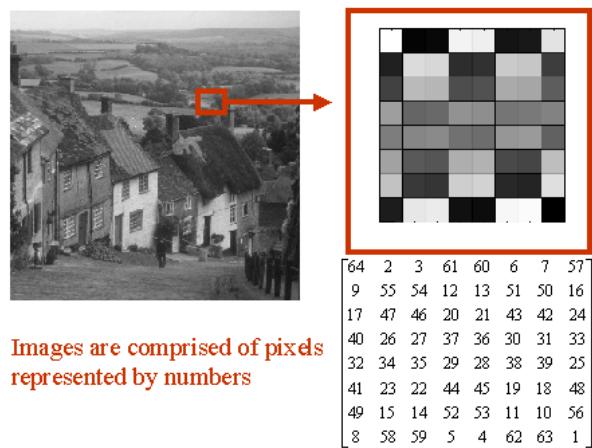


Figure 10.5: Metodo JPEG

Chapter 11

Ortogonalità e Isometrie

In questa capitolo \mathbb{R}^n è lo spazio vettoriale di riferimento.

11.1 Basi ortonormali

Definition 11.1.1. (i) Una *base ortogonale* di \mathbb{R}^n è una base $\mathbf{v}_1, \dots, \mathbf{v}_n$ tale che il prodotto interno $\mathbf{v}_i \cdot \mathbf{v}_j = 0$ per ogni $i \neq j$.

(ii) Una base ortogonale è *ortonormale* se $\mathbf{v}_i \cdot \mathbf{v}_i = 1$.

Una base ortogonale $\mathbf{v}_1, \dots, \mathbf{v}_n$ si trasforma in una base ortonormale considerando i vettori $\frac{\mathbf{v}_1}{\|\mathbf{v}_1\|}, \dots, \frac{\mathbf{v}_n}{\|\mathbf{v}_n\|}$.

Example 62. La base canonica $\mathbf{e}_1 = [1, 0, 0], \mathbf{e}_2 = [0, 1, 0], \mathbf{e}_3 = [0, 0, 1]$ di \mathbb{R}^3 è ortonormale. La base $\mathbf{v}_1 = [\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0], \mathbf{v}_2 = [-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0], \mathbf{v}_3 = [0, 0, 1]$ è anch'essa ortonormale.

Lemma 11.1.1. Sia $\mathbf{v}_1, \dots, \mathbf{v}_n$ una base ortonormale di \mathbb{R}^n e siano $\mathbf{x} = x_1\mathbf{v}_1 + \dots + x_n\mathbf{v}_n, \mathbf{y} = y_1\mathbf{v}_1 + \dots + y_n\mathbf{v}_n$ due vettori. Allora il prodotto interno di \mathbf{x} e \mathbf{y} si esprime come segue:

$$\mathbf{x} \cdot \mathbf{y} = x_1y_1 + \dots + x_ny_n.$$

Se $\mathbf{v}_1, \dots, \mathbf{v}_n$ è una base arbitraria (non necessariamente ortonormale), il Lemma 11.1.1 non vale in generale. Si ha piuttosto il seguente risultato:

$$\mathbf{x} \cdot \mathbf{y} = \sum_{i=1}^n \sum_{j=1}^n x_i y_j (\mathbf{v}_i \cdot \mathbf{v}_j).$$

Lemma 11.1.2. *Le coordinate di un vettore $\mathbf{w} \in \mathbb{R}^n$ rispetto ad una base ortonormale $\mathbf{v}_1, \dots, \mathbf{v}_n$ di \mathbb{R}^n sono calcolate con il prodotto interno:*

$$\mathbf{w} = (\mathbf{w} \cdot \mathbf{v}_1)\mathbf{v}_1 + \cdots + (\mathbf{w} \cdot \mathbf{v}_n)\mathbf{v}_n.$$

Proof. Sia $\mathbf{w} = w_1\mathbf{v}_1 + \cdots + w_n\mathbf{v}_n$. Allora $\mathbf{w} \cdot \mathbf{v}_1 = (w_1\mathbf{v}_1 + w_2\mathbf{v}_2 + \cdots + w_n\mathbf{v}_n) \cdot \mathbf{v}_1 = w_1(\mathbf{v}_1 \cdot \mathbf{v}_1) + w_2(\mathbf{v}_2 \cdot \mathbf{v}_1) + \cdots + w_n(\mathbf{v}_n \cdot \mathbf{v}_1) = w_1(\mathbf{v}_1 \cdot \mathbf{v}_1) + 0 + \cdots + 0 = w_1(\mathbf{v}_1 \cdot \mathbf{v}_1) = w_1$. \square

Example 63. Sia $\mathbf{w} = 2\mathbf{e}_1 + \mathbf{e}_2 + 3\mathbf{e}_3$ un vettore dello spazio. Calcoliamo le coordinate di \mathbf{w} rispetto alla base ortonormale $\mathbf{v}_1 = [\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0]$, $\mathbf{v}_2 = [-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0]$, $\mathbf{e}_3 = [0, 0, 1]$: $\mathbf{w} \cdot \mathbf{v}_1 = \frac{2}{\sqrt{2}} + \frac{1}{\sqrt{2}} = \frac{3}{\sqrt{2}}$; $\mathbf{w} \cdot \mathbf{v}_2 = -\frac{2}{\sqrt{2}} + \frac{1}{\sqrt{2}} = -\frac{1}{\sqrt{2}}$; $\mathbf{w} \cdot \mathbf{e}_3 = 3$. Allora si ha:

$$\mathbf{w} = \frac{3}{\sqrt{2}}\mathbf{v}_1 - \frac{1}{\sqrt{2}}\mathbf{v}_2 + 3\mathbf{e}_3$$

Lemma 11.1.3. *Ogni sottospazio vettoriale di \mathbb{R}^n ammette una base ortonormale.*

Proof. Analizziamo per semplicità il caso $n = 2$. Se i vettori \mathbf{v} e \mathbf{w} sono linearmente indipendenti, allora lo spazio vettoriale $\text{Span}(\mathbf{v}, \mathbf{w})$ è anche generato da \mathbf{v} e $\mathbf{w} - \frac{\mathbf{v} \cdot \mathbf{w}}{\mathbf{v} \cdot \mathbf{v}}\mathbf{v}$ che sono ortogonali: $\mathbf{v} \cdot [\mathbf{w} - \frac{\mathbf{v} \cdot \mathbf{w}}{\mathbf{v} \cdot \mathbf{v}}\mathbf{v}] = (\mathbf{v} \cdot \mathbf{w}) - \frac{\mathbf{v} \cdot \mathbf{w}}{\mathbf{v} \cdot \mathbf{v}}(\mathbf{v} \cdot \mathbf{v}) = 0$. \square

11.2 Sottospazi ortogonali

Due sottospazi X e Y di \mathbb{R}^n sono *ortogonali* se il prodotto interno di vettori dei due spazi è sempre nullo:

$$(\forall \mathbf{x} \in X)(\forall \mathbf{y} \in Y)(\mathbf{x} \cdot \mathbf{y} = 0).$$

Definition 11.2.1. Lo spazio ortogonale ad un sottospazio X è definito come

$$X^\perp = \{\mathbf{v} \in V : (\forall \mathbf{x} \in X)(\mathbf{v} \cdot \mathbf{x} = 0)\}.$$

Proposition 11.2.1. *Sia X un sottospazio di uno spazio vettoriale reale V . Allora $\dim X + \dim X^\perp = \dim V$.*

Proof. Sia $\mathbf{v}_1, \dots, \mathbf{v}_k$ una base ortonormale di X e $\mathbf{w}_1, \dots, \mathbf{w}_r$ una base ortonormale di X^\perp . Consideriamo il sottospazio Y generato dai vettori $\mathbf{v}_1, \dots, \mathbf{v}_k, \mathbf{w}_1, \dots, \mathbf{w}_r$. È sufficiente provare che $Y = V$. Supponiamo per assurdo che esista $\mathbf{u} \in V \setminus Y$. Consideriamo il vettore $\mathbf{t} = (\mathbf{u} \cdot \mathbf{v}_1)\mathbf{v}_1 + \dots + (\mathbf{u} \cdot \mathbf{v}_k)\mathbf{v}_k$. Allora $\mathbf{0} \neq \mathbf{u} - \mathbf{t}$ (altrimenti $\mathbf{u} \in X$) e si ha $(\mathbf{u} - \mathbf{t}) \cdot \mathbf{v}_i = \mathbf{u} \cdot \mathbf{v}_i - \mathbf{t} \cdot \mathbf{v}_i = 0$ per ogni $i = 1, \dots, k$. Quindi $\mathbf{u} - \mathbf{t} \in X^\perp$ e si scrive come combinazione lineare di $\mathbf{w}_1, \dots, \mathbf{w}_r$. Ne segue la conclusione $\mathbf{u} \in Y$, che contraddice l'ipotesi su \mathbf{u} . \square

Example 64. Sia $A = [3, 3]$ un vettore del piano. Sia X il sottospazio vettoriale generato da A , che è la retta passante per l'origine e per il punto A . Lo spazio ortogonale X^\perp a X è la retta di equazione $3x + 3y = 0$. Allora una base del piano è determinata dal vettore A e da un vettore $B \in X^\perp$, per esempio $B = [1, -1]$.

Example 65. Sia $A = [3, 3, 1]$ un vettore dello spazio. Sia X il sottospazio vettoriale generato da A , che è la retta passante per l'origine e per il punto A . Lo spazio ortogonale X^\perp a X è il piano di equazione $3x + 3y + z = 0$. Allora una base dello spazio è determinata dal vettore A e da due vettori $B, C \in X^\perp$ non collineari, per esempio $B = [1, -1, 0]$ e $C = [1, 1, -6]$.

11.3 Rango di una matrice

Sia A una matrice di tipo $m \times n$ a coefficienti reali e sia $f_A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ la trasformazione lineare generata da A rispetto alle basi canoniche di \mathbb{R}^n e \mathbb{R}^m rispettivamente. Le colonne di A generano un sottospazio $\text{Im}(f_A) = \text{Span}(A^1, \dots, A^n)$ di \mathbb{R}^m e le righe di A generano un sottospazio $\text{Span}(A_1, \dots, A_m)$ di \mathbb{R}^n . Abbiamo la seguente uguaglianza per la dimensione di $\text{Span}(A^1, \dots, A^n)$:

$$n = \dim \text{Ker}(f_A) + \dim \text{Span}(A^1, \dots, A^n)$$

mentre il sottospazio $\text{Span}(A_1, \dots, A_m)$ di \mathbb{R}^n è ortogonale a $\text{Ker}(f_A)$ perché, se $A\mathbf{x} = \mathbf{0}$, si ha $A_i \cdot \mathbf{x} = 0$. Quindi dalla Proposizione 11.2.1 si ricava:

$$n = \dim \text{Ker}(f_A) + \dim \text{Span}(A_1, \dots, A_m).$$

In conclusione, $\dim \text{Span}(A^1, \dots, A^n) = \dim \text{Span}(A_1, \dots, A_m)$.

Il *rango* $r(A)$ della matrice A è la dimensione comune di questi due sottospazi. È chiaro che $r(A) \leq \min(n, m)$. Se estraiamo $r(A)$ colonne linearmente indipendenti $c_1, \dots, c_{r(A)}$ e $r(A)$ righe linearmente indipendenti $d_1, \dots, d_{r(A)}$ otteniamo una matrice quadrata $E = (e_{ij})$ di ordine $r(A)$ con $e_{ij} = A_{c_i d_j}$ il cui determinante è non nullo.

Se consideriamo il sistema lineare omogeneo $A\mathbf{x} = \mathbf{0}$, allora la dimensione dello spazio delle soluzioni è $n - r(A)$, dove $r(A)$ è il rango di A .

Se il sistema lineare $A\mathbf{x} = \mathbf{b}$ ammette almeno una soluzione \mathbf{v}_0 , allora l'insieme delle soluzioni è $\{\mathbf{v}_0 + \mathbf{x} : A\mathbf{x} = \mathbf{0}\}$.

11.4 Matrici ortogonali

Definition 11.4.1. Una matrice quadrata A di ordine n è *ortogonale* se la matrice inversa di A è la matrice trasposta A^t di A .

Se A è ortogonale allora $\det(A) = \pm 1$: $\det(A)^2 = \det(A)\det(A) = \det(A)\det(A^t) = \det(AA^t) = \det(I_n) = 1$, perché la trasposta di una matrice ha lo stesso determinante della matrice stessa.

Proposition 11.4.1. *Sia A una matrice quadrata di ordine n . Allora le seguenti condizioni sono equivalenti:*

- (i) A è ortogonale;
- (ii) Le colonne di A costituiscono una base ortonormale di \mathbb{R}^n (cioé, $A^i \cdot A^i = 1$ e $A^i \cdot A^j = 0$ per ogni $i \neq j$);
- (iii) La matrice A preserva il prodotto interno di vettori: $A\mathbf{x} \cdot A\mathbf{y} = \mathbf{x} \cdot \mathbf{y}$ (e quindi le lunghezze dei vettori $\|A\mathbf{x}\| = \|\mathbf{x}\|$ per ogni \mathbf{x}, \mathbf{y}).

Proof. (i) \Leftrightarrow (ii) $A^t A = I_n$ sse $A^i \cdot A^j = \delta_{ij}$, il delta di Kronecker.

Sia $B = AA^t$. Allora $B_{ii} = A_i \cdot (A^t)^i = A_i \cdot A_i = 1$ e, per $i \neq j$, $B_{ij} = A_i \cdot (A^t)^j = A_i \cdot A_j = 0$. Quindi B è la matrice identica. La stessa prova funziona per $A^t A$.

- (i) \Rightarrow (iii) $(A\mathbf{x}) \cdot (A\mathbf{y}) = (A\mathbf{x})^t (A\mathbf{y}) = \mathbf{x}^t A^t A\mathbf{y} = \mathbf{x}^t I_n \mathbf{y} = \mathbf{x} \cdot \mathbf{y}$.
- (iii) \Rightarrow (i) Se $\mathbf{x}^t A^t A\mathbf{y} = \mathbf{x}^t \mathbf{y}$ per ogni \mathbf{x} e \mathbf{y} , allora scegliendo opportunamente i valori di \mathbf{x} e \mathbf{y} si ricava $A^t A = I_n$.

□

Proposition 11.4.2. *Le matrici quadrate ortogonali di ordine n costituiscono un gruppo $O(n)$ rispetto alla moltiplicazione di matrici: il prodotto di due matrici ortogonali è una matrice ortogonale, l'inversa di una matrice ortogonale è ortogonale e la matrice identica è ortogonale.*

Proof. Siano A e B ortogonali. Allora, $(AB)(AB)^t = ABB^tA^t = AA^t = I_n$. \square

Le matrici ortogonali con determinante uguale ad 1 costituiscono un sottogruppo del gruppo $O(n)$.

11.5 Il teorema spettrale

Definition 11.5.1. Un endomorfismo lineare $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ è *simmetrico* se la matrice di f rispetto alla base canonica è simmetrica.

Proposition 11.5.1. *Sia $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ un endomorfismo lineare. Allora le seguenti condizioni sono equivalenti:*

(i) f è simmetrico;

(ii) $f(\mathbf{x}) \cdot \mathbf{y} = \mathbf{x} \cdot f(\mathbf{y})$ per ogni vettore \mathbf{x} e \mathbf{y} ;

(iii) La matrice di f rispetto ad una qualsiasi base ortonormale è simmetrica.

Proof. Sia $\mathbf{v}_1, \dots, \mathbf{v}_n$ una base ortonormale. Sia $A = (a_{ij})$ la matrice di f rispetto alla data base e siano $\mathbf{x} = x_1\mathbf{v}_1 + \dots + x_n\mathbf{v}_n$, $\mathbf{y} = y_1\mathbf{v}_1 + \dots + y_n\mathbf{v}_n$ due vettori. Dal Lemma 11.1.1 si ha:

$$\mathbf{x} \cdot f(\mathbf{y}) = [x_1 \dots x_n]A[y_1 \dots y_n]^t \quad (11.1)$$

e

$$\begin{aligned} f(\mathbf{x}) \cdot \mathbf{y} &= [y_1 \dots y_n]A[x_1 \dots x_n]^t \\ &= ([y_1 \dots y_n]A[x_1 \dots x_n]^t)^t \\ &= [x_1 \dots x_n]A^t[y_1 \dots y_n]^t \end{aligned} \quad (11.2)$$

(iii) \Rightarrow (i) Ovvia.

(i) \Rightarrow (ii) Segue da (11.1) e (11.2) considerando che la matrice A è simmetrica (i.e., $A = A^t$).

(ii) \Rightarrow (iii) Sia $\mathbf{v}_1, \dots, \mathbf{v}_n$ una base ortonormale. Sia $A = (a_{ij})$ la matrice di f rispetto alla data base e siano $\mathbf{x} = x_1\mathbf{v}_1 + \dots + x_n\mathbf{v}_n$, $\mathbf{y} = y_1\mathbf{v}_1 + \dots + y_n\mathbf{v}_n$ due vettori. Da (11.1), (11.2) e l'ipotesi si ha:

$$[y_1 \dots y_n]A[x_1 \dots x_n]^t = [x_1 \dots x_n]A^t[y_1 \dots y_n]^t.$$

Scegliendo opportunamente \mathbf{x} e \mathbf{y} si verifica che A è simmetrica. Per esempio, se $x_2 = 1$, $y_1 = 1$, $x_j = 0$ ($j \neq 2$) e $y_k = 0$ ($k \neq 1$), allora $[0, 1, 0, \dots, 0]A = A_2$ seleziona la seconda riga di A , mentre $[0, 1, 0, \dots, 0]A^t = A^2$ seleziona la seconda colonna di A . Moltiplicando scalarmente per $[1, 0, \dots, 0]^t$ da (11.2) ricaviamo $a_{21} = a_{12}$.

□

Theorem 11.5.2. (Teorema Spettrale) *Le seguenti condizioni sono equivalenti per un endomorfismo lineare $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$:*

- (i) f è simmetrico;
- (ii) Esiste una base ortonormale costituita da autovettori di f ;
- (iii) La matrice di f rispetto alla base canonica è diagonalizzabile.

Proof. (ii) \Rightarrow (i) La matrice di f rispetto ad una base di autovettori è diagonale, quindi simmetrica. La conclusione segue dalla Proposizione 11.5.1.

(i) \Rightarrow (ii) Limitiamo la prova alle dimensioni $n = 2, 3$. Analizziamo prima il caso $n = 2$. Dalla Proposizione 11.5.1 la matrice A di f rispetto alla base canonica è simmetrica:

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

Se $b = 0$, la matrice A è diagonale e le colonne A^1, A^2 costituiscono una base ortogonale di autovettori. Dividendo entrambi i vettori per la loro lunghezza otteniamo una base ortonormale di autovettori. Supponiamo quindi $b \neq 0$. Il polinomio caratteristico di A è: $\lambda^2 - (a+c)\lambda + (ac - b^2) = 0$, da cui si ricava:

$$\lambda = \frac{(a+c) \pm \sqrt{(a+c)^2 - 4ac + b^2}}{2} = \frac{(a+c) \pm \sqrt{(a-c)^2 + b^2}}{2}$$

Quindi abbiamo due radici λ_1 e λ_2 reali e distinte. Dalla Proposizione 10.1.1 due autovettori \mathbf{a} di λ_1 e \mathbf{b} di λ_2 sono linearmente indipendenti e costituiscono

una base di \mathbb{R}^2 . Dall'ipotesi che f è simmetrico si ha $\lambda_1(\mathbf{a} \cdot \mathbf{b}) = (\lambda_1\mathbf{a}) \cdot \mathbf{b} = f(\mathbf{a}) \cdot \mathbf{b} = \mathbf{a} \cdot f(\mathbf{b}) = \mathbf{a} \cdot (\lambda_2\mathbf{b}) = \lambda_2(\mathbf{a} \cdot \mathbf{b})$, che implica $\mathbf{a} \cdot \mathbf{b} = 0$.

Analizziamo ora il caso $n = 3$. Sia A la matrice simmetrica di f rispetto alla base canonica. Il polinomio caratteristico di f è di terzo grado e quindi ammette una radice reale λ_1 . Sia $\mathbf{v} = v_1\mathbf{e}_1 + v_2\mathbf{e}_2 + v_3\mathbf{e}_3$ un autovettore di λ_1 di lunghezza unitaria (cioé, $f(\mathbf{v}) = \lambda\mathbf{v}$ e $\|\mathbf{v}\| = 1$). Consideriamo il piano \mathbf{v}^\perp ortogonale a \mathbf{v} di equazione $v_1x + v_2y + v_3z = 0$. Siano $\mathbf{w}_1, \mathbf{w}_2 \in \mathbf{v}^\perp$ due vettori del piano, ortogonali tra loro e di lunghezza unitaria. Allora, i vettori $\mathbf{v}, \mathbf{w}_1, \mathbf{w}_2$ costituiscono una base ortonormale di \mathbb{R}^3 . Dalla Proposizione 11.5.1 la matrice B di f rispetto alla base $\mathbf{v}, \mathbf{w}_1, \mathbf{w}_2$ è simmetrica. Inoltre, per $i = 1, 2$, si ha: $\mathbf{0} = \lambda_1(\mathbf{v} \cdot \mathbf{w}_i) = f(\mathbf{v}) \cdot \mathbf{w}_i = \mathbf{v} \cdot f(\mathbf{w}_i)$. Quindi $f(\mathbf{w}_i)$ è ortogonale a \mathbf{v} , da cui si ricava $f(\mathbf{v}^\perp) \subseteq \mathbf{v}^\perp$. Quindi, da $f(\mathbf{v}) = \lambda_1\mathbf{v}$, $f(\mathbf{w}_1) = a\mathbf{w}_1 + b\mathbf{w}_2$ e $f(\mathbf{w}_2) = d\mathbf{w}_1 + c\mathbf{w}_2$ si ricava

$$B = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & a & b \\ 0 & d & c \end{bmatrix}$$

Infine dalla simmetria di B si ha $b = d$, per cui

$$B = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & a & b \\ 0 & b & c \end{bmatrix}$$

Se $b = 0$, la matrice B è diagonale e dalla Proposizione 10.2.1 $\mathbf{v}, \mathbf{w}_1, \mathbf{w}_2$ costituiscono una base di autovettori. Supponiamo $b \neq 0$. Il polinomio caratteristico di B è: $(\lambda - \lambda_1)(\lambda^2 - (a+c)\lambda + (ac-b^2))$. Quindi B , oltre all'autovalore λ_1 già conosciuto, ha due autovalori reali distinti (si veda il caso $n = 2$) λ_2 e λ_3 . Consideriamo un autovettore \mathbf{t}_i di λ_i ($i = 2, 3$) di lunghezza unitaria. Allora come nel caso $n = 2$, si ottiene $\lambda_2(\mathbf{t}_2 \cdot \mathbf{t}_3) = f(\mathbf{t}_2) \cdot \mathbf{t}_3 = \mathbf{t}_2 \cdot f(\mathbf{t}_3) = \lambda_3(\mathbf{t}_2 \cdot \mathbf{t}_3)$, da cui $\mathbf{t}_2 \cdot \mathbf{t}_3 = 0$. La base di autovettori è $\mathbf{v}, \mathbf{t}_2, \mathbf{t}_3$. \square

Corollary 11.5.3. *Sia A una matrice quadrata di ordine n a coefficienti reali. Allora le seguenti condizioni sono equivalenti:*

- (i) A è simmetrica;
- (ii) Esiste una base ortonormale di \mathbb{R}^n costituita da autovettori di A ;
- (iii) Esiste una matrice ortogonale U (cioé, le colonne U^1, \dots, U^n di U costituiscono una base ortonormale di \mathbb{R}^n) tale che $U^t A U$ è diagonale.

11.6 Isometrie lineari

Una *isometria lineare* è un isomorfismo lineare $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ che preserva il prodotto interno dei vettori: $f(\mathbf{x}) \cdot f(\mathbf{y}) = \mathbf{x} \cdot \mathbf{y}$ e quindi anche la lunghezza dei vettori. Sia A la matrice della isometria lineare rispetto alla base canonica costituita dai vettori colonna $\mathbf{e}_1, \dots, \mathbf{e}_n$ di \mathbb{R}^n . Da $f(\mathbf{e}_i) = A\mathbf{e}_i = A^i$ (la colonna i della matrice A) si ricava che $\|A^i\| = 1$. Inoltre, per $i \neq j$, si ha: $0 = \mathbf{e}_i \cdot \mathbf{e}_j = f(\mathbf{e}_i) \cdot f(\mathbf{e}_j) = A^i \cdot A^j$. Quindi la matrice A è ortogonale, la sua inversa $A^{-1} = A^t$ è la matrice trasposta e $\det(A) = \pm 1$.

11.6.1 Rotazioni e simmetrie assiali del piano

Consideriamo il piano cartesiano con la base canonica $\mathbf{e}_1 = [1, 0]$ e $\mathbf{e}_2 = [0, 1]$. Una isometria lineare è determinata (rispetto alla base canonica) da una matrice ortogonale

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

Rotazioni del piano

Analizziamo prima il caso in cui $\det(A) = 1$. Valgono le seguenti proprietà: $a^2 + b^2 = 1$, $c^2 + d^2 = 1$, $ac + bd = 0$ ed infine $ad - bc = 1$. Le colonne di A costituiscono una base ortonormale ed inoltre l'orientazione di questa base coincide con quella canonica $\mathbf{e}_1, \mathbf{e}_2$, perché il determinante di A è 1. Se utilizziamo le coordinate polari, esiste un angolo θ tale che $[a, b] = [\cos \theta, \sin \theta]$ e $[c, d] = [\cos(\theta + \frac{\pi}{2}), \sin(\theta + \frac{\pi}{2})] = [-\sin \theta, \cos \theta]$. Quindi la matrice A è uguale a

$$A = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}.$$

La matrice A corrisponde ad una rotazione R_θ in senso antiorario di un angolo θ . Il vettore \mathbf{e}_1 viene trasformato nel vettore

$$R_\theta(\mathbf{e}_1) = (\cos \theta)\mathbf{e}_1 + (\sin \theta)\mathbf{e}_2,$$

mentre il vettore \mathbf{e}_2 viene trasformato nel vettore

$$R_\theta(\mathbf{e}_2) = (-\sin \theta)\mathbf{e}_1 + (\cos \theta)\mathbf{e}_2 = \cos(\theta + \pi/2)\mathbf{e}_1 + \sin(\theta + \pi/2)\mathbf{e}_2.$$

Ogni rotazione è invertibile. La matrice inversa, che corrisponde ad una rotazione oraria di un angolo θ , è la sua trasposta:

$$\begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} = \begin{bmatrix} \cos(-\theta) & \cos(\frac{\pi}{2} - \theta) \\ \sin(-\theta) & \sin(\frac{\pi}{2} - \theta) \end{bmatrix}$$

Vi è una corrispondenza bigettiva tra rotazioni e matrici

$$\begin{bmatrix} a & b \\ -b & a \end{bmatrix}$$

con determinante $a^2 + b^2 = 1$.

Si ricorda che un altro modo di rappresentare le rotazioni è con la moltiplicazione dei numeri complessi. Se $z = \cos \theta + i \sin \theta$ e $v = v_1 + v_2 i$, allora

$$zv = (v_1 \cos \theta - v_2 \sin \theta) + (v_1 \sin \theta + v_2 \cos \theta)i,$$

che corrisponde alla moltiplicazione di matrici:

$$\begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} v_1 \cos \theta - v_2 \sin \theta \\ v_1 \sin \theta + v_2 \cos \theta \end{bmatrix}$$

La matrice A della rotazione di angolo $\theta \neq 0$ ammette autovalori reali soltanto per l'angolo $\theta = \pi$:

$$\begin{vmatrix} \cos \theta - \lambda & -\sin \theta \\ \sin \theta & \cos \theta - \lambda \end{vmatrix} = (\cos \theta - \lambda)^2 + \sin \theta^2$$

$(\cos \theta - \lambda)^2 + \sin \theta^2 = 0$ sse $\theta = \pi$. In tal caso l'autovalore $\lambda = -1$ è doppio.

Riflessioni (Simmetrie assiali) lineari del piano

Se la matrice

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

ha determinante uguale a -1 , allora le colonne di A corrispondono ad una base ortonormale con orientazione diversa da quella della base canonica e_1, e_2 . Quindi esiste un angolo θ tale che $[a, b] = [\cos \theta, \sin \theta]$ e $[c, d] = [\cos(\theta - \frac{\pi}{2}), \sin(\theta - \frac{\pi}{2})] = [\sin \theta, -\cos \theta]$. Quindi la matrice A è uguale a

$$A = \begin{bmatrix} \cos \theta & \sin \theta \\ \sin \theta & -\cos \theta \end{bmatrix}$$

ed è simmetrica. Quindi, essendo ortogonale, abbiamo $AA = I_2$.

Calcoliamo gli autovalori della matrice A :

$$\begin{vmatrix} \cos \theta - \lambda & \sin \theta \\ \sin \theta & -\cos \theta - \lambda \end{vmatrix} = (\lambda - \cos \theta)(\lambda + \cos \theta) - \sin^2 \theta = \lambda^2 - 1$$

L'equazione $\lambda^2 - 1 = 0$ ammette come soluzioni $\lambda = \pm 1$. Quindi esistono due rette R_1 e R_2 passanti per l'origine e perpendicolari tra di loro con le seguenti proprietà: la prima retta R_1 è l'autospazio dell'autovalore $\lambda = 1$, mentre la seconda retta R_2 è l'autospazio dell'autovalore $\lambda = -1$. La matrice A agisce sulla prima retta R_1 come la matrice identica, mentre agisce sulla seconda retta R_2 mandando ogni vettore \mathbf{v} che ha la direzione di R_2 in $-\mathbf{v}$.

La trasformazione lineare determinata dalla matrice A si chiama *riflessione (simmetria assiale) lineare* di asse la retta R_1 passante per l'origine. La trasformazione geometrica lascia invariata la retta R_1 e associa ad ogni punto P del piano non appartenente ad R_1 il punto Q in modo tale che il segmento PQ sia perpendicolare alla retta R_1 e abbia come punto medio H , che è il piede della perpendicolare condotta da P a R_1 .

Il prodotto di due riflessioni lineari A e B è una rotazione, perché $\det(AB) = \det(A)\det(B) = (-1)(-1) = 1$.

Example 66. Una riflessione lineare attorno all'asse delle y si rappresenta con la matrice

$$\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$$

Infatti,

$$\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix}; \quad \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ 0 \end{bmatrix} = \begin{bmatrix} -x \\ 0 \end{bmatrix}$$

mentre per ogni altro vettore:

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

Example 67. La matrice ortogonale:

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

rappresenta la riflessione lineare rispetto alla bisettrice $y = x$:

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} y \\ x \end{bmatrix}$$

Il prodotto delle due riflessioni lineari è una rotazione di un angolo $\theta = -\frac{\pi}{2}$:

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

Il prodotto delle due riflessioni lineari in ordine inverso è una rotazione di un angolo $\theta = \frac{\pi}{2}$:

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

Nella parte restante di questa sezione calcoliamo, rispetto alla base canonica $\mathbf{e}_1, \mathbf{e}_2$, la matrice A di una simmetria assiale lineare rispetto alla retta R che forma un angolo ϕ con l'asse positivo delle x . Consideriamo il vettore unitario $\mathbf{v}_1 = (\cos \phi) \mathbf{e}_1 + (\sin \phi) \mathbf{e}_2$ che giace nella retta R . La retta Q passante per l'origine e perpendicolare ad R ha equazione lineare $(\cos \phi)x + (\sin \phi)y = 0$. Un vettore unitario giacente sulla retta Q è $\mathbf{v}_2 = (\sin \phi) \mathbf{e}_1 - (\cos \phi) \mathbf{e}_2$. La matrice della simmetria assiale rispetto alla base $\mathbf{v}_1, \mathbf{v}_2$ è la matrice diagonale degli autovalori:

$$B = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

Infatti, $[1, 0]$ sono le coordinate del vettore \mathbf{v}_1 rispetto alla base ortonormale $\mathbf{v}_1, \mathbf{v}_2$, mentre $[0, 1]$ sono le coordinate del vettore \mathbf{v}_2 rispetto alla base $\mathbf{v}_1, \mathbf{v}_2$.

Per calcolare la matrice A , dobbiamo applicare alla matrice B un cambio di base dalla base canonica $\mathbf{e}_1, \mathbf{e}_2$ alla base $\mathbf{v}_1, \mathbf{v}_2$. Allora abbiamo:

$$A = \begin{bmatrix} \cos \phi & \sin \phi \\ \sin \phi & -\cos \phi \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} \cos \phi & \sin \phi \\ \sin \phi & -\cos \phi \end{bmatrix} = \begin{bmatrix} \cos \phi & -\sin \phi \\ \sin \phi & \cos \phi \end{bmatrix} \begin{bmatrix} \cos \phi & \sin \phi \\ \sin \phi & -\cos \phi \end{bmatrix} =$$

$$\begin{bmatrix} \cos \phi^2 - \sin \phi^2 & 2 \cos \phi \sin \phi \\ 2 \cos \phi \sin \phi & \sin \phi^2 - \cos \phi^2 \end{bmatrix} = \begin{bmatrix} \cos(2\phi) & \sin(2\phi) \\ \sin(2\phi) & -\cos(2\phi) \end{bmatrix}$$

La matrice A è ortogonale con determinante -1 . Verifichiamo che $A\mathbf{v}_1^t = \mathbf{v}_1^t$:

$$\begin{bmatrix} \cos(2\phi) & \sin(2\phi) \\ \sin(2\phi) & -\cos(2\phi) \end{bmatrix} \begin{bmatrix} \cos(\phi) \\ \sin(\phi) \end{bmatrix} = \begin{bmatrix} \cos(\phi) \cos(2\phi) + \sin(\phi) \sin(2\phi) \\ \cos(\phi) \sin(2\phi) - \sin(\phi) \cos(2\phi) \end{bmatrix}$$

Verifichiamo che il vettore ottenuto coincide con \mathbf{v}_1^t . Per fare questo ci avvaliamo dei numeri complessi e scriviamo $\mathbf{v}_1 = \cos \phi + i \sin \phi$ come numero complesso. Ricordiamo dalla Proposizione 1.1.7 che $\mathbf{v}_1 \mathbf{v}_1^t = \mathbf{v}_1^2 =$

$\cos(2\phi) + i \sin(2\phi)$ e che il coniugato di \mathbf{v}_1 è $\bar{\mathbf{v}}_1 = \cos \phi - i \sin \phi$. Allora, si ha (ricordando che la norma $|\mathbf{v}_1| = 1$):

$$\mathbf{v}_1 = \mathbf{v}_1 |\mathbf{v}_1|^2 = \mathbf{v}_1 \mathbf{v}_1 \bar{\mathbf{v}}_1 = (\cos(\phi) \cos(2\phi) + \sin(\phi) \sin(2\phi)) + i(\cos(\phi) \sin(2\phi) - \sin(\phi) \cos(2\phi))$$

ed abbiamo ottenuto la conclusione.

11.7 Rotazioni nello spazio

Le coordinate polari possono essere estese in tre dimensioni anche utilizzando le coordinate (ρ, θ, φ) , in cui ρ è la distanza dall'origine degli assi, θ è l'angolo formato con l'asse z e φ l'angolo formato dalla proiezione sul piano-xy con l'asse x . Le tre coordinate cartesiane di un punto si ottengono dalle tre coordinate polari di quel punto con le formule:

$$x = \rho \sin \theta \cos \varphi; \quad y = \rho \sin \theta \sin \varphi; \quad z = \rho \cos \theta.$$

Consideriamo nello spazio vettoriale \mathbb{R}^3 una matrice ortogonale A di tipo 3×3 con determinante uguale ad 1:

$$A = \begin{bmatrix} a & d & g \\ b & e & h \\ c & f & i \end{bmatrix}$$

I tre vettori colonna A^1, A^2, A^3 costituiscono una base ortonormale di \mathbb{R}^3 con lo stesso orientazione della base canonica $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$.