Capitolo 1

Numeri complessi

1.1 Numeri complessi e rappresentazione geometrica

Definizione 1.2. *Un numero complesso è un simbolo del tipo* $a + ib$*, dove* $a, b \in \mathbb{R}$*.*

In seguito indicheremo con $\mathbb{C} = \{a + ib : a, b \in \mathbb{R}\}$ l'insieme dei numeri complessi. Un numero complesso $a + ib$, può essere pensato come a moltiplicato per l'unità 1 mentre b sia moltiplicata per i, chiamata *unità immaginaria*. Due numeri complessi $z = a + ib$ e $w = a' + ib'$ si diranno uguali, e scriveremo $z = w$, se $a = a'$ e $b = b'$. Se $b = 0$ il numero complesso si identificherà con il numero reale a. Quindi i numeri reali possono considerarsi come particolari tipi di numeri complessi. Vogliamo definire una somma ed un prodotto su $\mathbb{C}$. La somma converrà definirla ponendo

$$(a + ib) + (a' + ib') := (a + a') + i(b + b');$$

con che si conservano le proprietà commutativa ed associativa. Quindi, per ogni $z, w, k \in \mathbb{C}$, si ha $z + (w + k) = (z + w) + k$ e $z + w = w + z$ (verificare per esercizio). Inoltre, se indichiamo con $0 = 0 + i0$, allora per ogni numero complesso $z = a + ib$ si ha

$$0 + z = z + 0 = z.$$

Infine, se $a+ib \in \mathbb{C}$, allora $(a+ib)+(-a+i(-b)) = (-a+i(-b))+(a+ib) = 0$, ovvero ogni elemento ammette un inverso rispetto alla somma. Quindi, se $z, w, h \in \mathbb{C}$ e

$$z + w = z + h,$$

allora $w = h$ (dimostrare!).

I numeri complessi sono in corrispondenza biunivoca con i punti del piano. Infatti, basta considerare come immagine del numero complesso $z = a + ib$ il punto P del piano di coordinate (a, b), dove a corrisponde alla proiezione di P lungo l'asse delle x e b corrisponde alla proiezione di P lungo l'asse delle y, rispettivamente (vedi Figura 1.2.1) e reciprocamente.

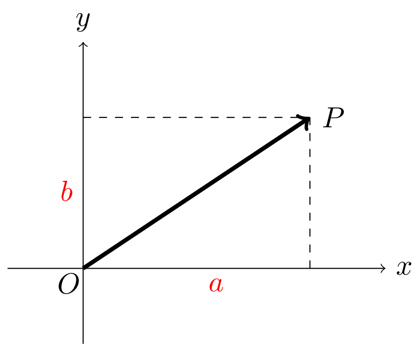


Figura 1.2.1.

Il punto P si può considerare come il segmento di estremi P e O, di componenti a e b. In questa rappresentazione geometrica, hanno considerevole importanza anche le coordinate polari ρ e θ del punto P che prendono rispettivamente il nome di modulo e argomento del numero complesso $a+ib$.

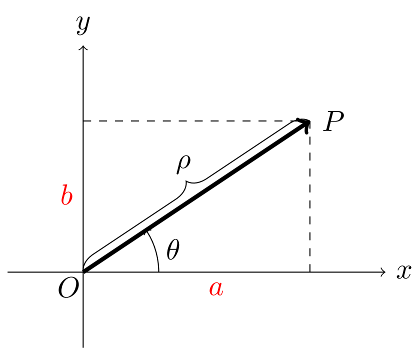


Figura 1.2.2.

In altre parole, il modulo di un numero complesso $z = a+ib$, che denoteremo con $|z|$ è il numero reale positivo $\rho := \sqrt{a^2 + b^2}$ mentre l'argomento di z, se $z \neq 0$, è l'angolo, determinato a meno di multipli di 2π, definito dalle formule:

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

Riferendosi ai moduli e agli argomenti un numero complesso, non nullo, assume la forma $\rho\cos\theta + i\rho\sin\theta$. Il complesso coniugato di un numero complesso $z = a + ib$ è il numero complesso $\overline{z} = a - ib$. L'applicazione

$$\mathbb{C} \longrightarrow \mathbb{C} \qquad z \mapsto \overline{z}$$

si chiama coniugio. Se pensiamo ad un numero complesso z come un punto del piano, il coniugio è la riflessione rispetto all'asse delle x.

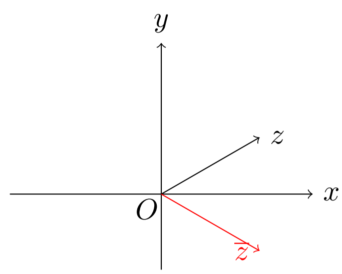


Figura 1.2.3.

In coordinate polari se se $z = \rho\cos\theta + i\rho\sin\theta$, allora $\overline{z} = \rho\cos(-\theta) + i\rho\sin(-\theta)$. Prima di definire un prodotto sui numeri complessi, vogliamo porre l'attenzione che sommare due numeri complessi z e w, corrisponde, geometricamente, alla *somma* dei segmenti corrispondenti nota come regola del parallelogramma

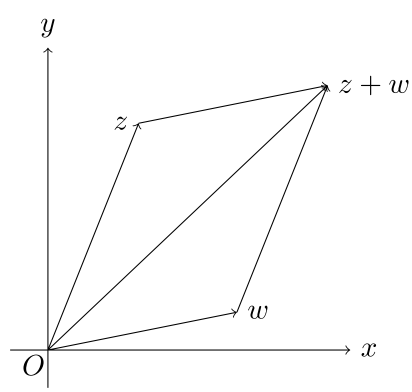


Figura 1.2.4.

Consideriamo due numeri speciali: a e ia, dove a è un numero reale positivo. Vorremmo che moltiplicare i il numero a fosse ia che geometricamente corrisponde a ruotare il vettore di coordinate $(a, 0)$ di $\frac{\pi}{2}$.

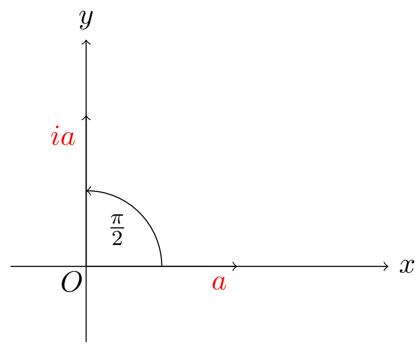


Figura 1.2.5.

Quindi se vogliamo che la stessa legge valga per i dovrà necessariamente porsi

$$i^2 = -1.$$

Quindi se vogliamo che valgono le fondamentali proprietà commutativa, associativa e distributiva della moltiplicazione dei numeri reali, i numeri complessi dovranno necessariamente moltiplicarsi conformemente alla formula:

$$(a + ib)(a' + ib') := (aa' - bb') + i(ab' + a'b).$$

In coordinate polari tale legge ha la seguente interpretazione geometrica: se $z = \rho_1 \cos(\theta_1) + i\rho_1 \sin(\theta_1)$ e $w = \rho_2 \cos(\theta_2) + i\rho_2 \sin(\theta_2)$, allora $zw = \rho_1\rho_2 \cos(\theta_1 + \theta_2) + i\rho_1\rho_2 \sin(\theta_1 + \theta_2)$.

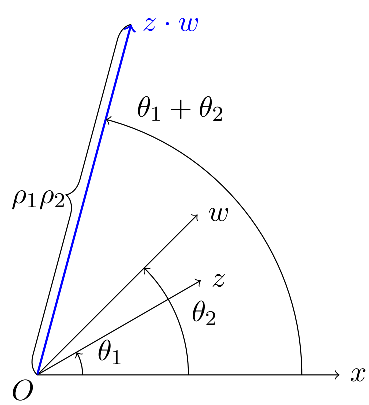


Figura 1.2.6.

Il prodotto gode delle seguenti proprietà.

Proposizione 1.3. *Siano* $z, w, h \in \mathbb{C}$. *Allora*

- $z \cdot (w \cdot h) = (z \cdot w) \cdot h$ *(Associativa);*

- *$z \cdot w = w \cdot z$ (Commutativa);*
- *$(z+w) \cdot h = z \cdot h + w \cdot h$;*
- *$\exists! 1 = 1 + i0$ tale che $1 \cdot z = z \cdot 1 = z$ per ogni $z \in \mathbb{C}$;*
- *se $z = a + ib \in \mathbb{C}$, $z \neq 0$, allora $z' = \frac{a}{a^2+b^2} + i\frac{-b}{a^2+b^2}$ soddisfa $z \cdot z' = z' \cdot z = 1$. L'elemento z' è unico e verrà indicato con z^{-1}.*

Quindi $(\mathbb{C}, +, \cdot)$ è un campo.

Abbiamo visto che possiamo pensare a $\mathbb{R}$ come un sottoinsieme dei numeri complessi. Geometricamente, $\mathbb{R}$ si identifica con l'asse delle x. Inoltre, se $x, y \in \mathbb{R}$, allora il prodotto e la somma, come numeri complessi, coincide con il prodotto e la somma come numeri reali. L'asse delle x è chiamata asse reale mentre l'asse delle y è chiamata asse immaginaria.

Sia $\lambda \in \mathbb{R}$ e sia $z = a + ib$. Allora $\lambda(a + ib) = \lambda a + i\lambda b =$. L'inverso di z rispetto alla somma è dato da $(-1)z$. Da qui in avanti indicheremo con $z - w := z + (-1)w$.

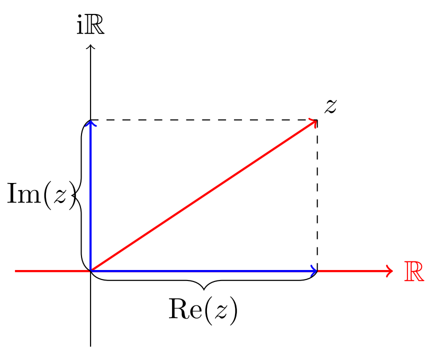
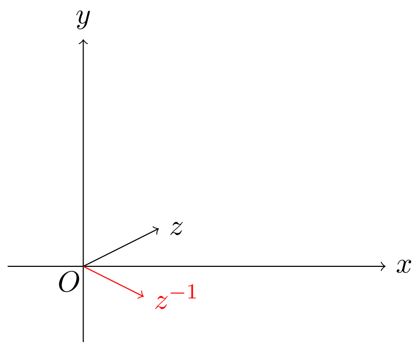


Figura 1.3.1.

Se $z = a + ib$ è un numero complesso, allora la *parte reale* di z, che indicheremo con $\mathrm{Re}(z)$, è il numero reale a, mentre la *parte immaginaria* di z, che indicheremo con $\mathrm{Im}(z)$, è il numero reale b. Un numero complesso è reale se e solamente se $\mathrm{Im}(z) = 0$. Se z è un numero complesso tale che $\mathrm{Re}(z) = 0$, allora z si dice *immaginario puro*. Il modulo di un numero complesso $|z| = \sqrt{\mathrm{Re}(z)^2 + \mathrm{Im}(z)^2}$.

Sia $z \neq 0$. In coordinate polari se $z = \rho \cos\theta + i\rho \sin\theta = \rho(\cos\theta + i \sin\theta)$, allora l'inverso di z ha coordiante polari

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



da cui segue che $z^{-1} = \frac{\overline{z}}{\|z\|^2}$. Geometricamente z^{-1} si ottiene applicando a z la riflessione rispetto all'asse delle x e poi moltiplicando tale elemento per l'inverso del modulo al quadrato di z. Se $\| z \| = 1$, allora $z^{-1} = \overline{z}$.

Proposizione 1.4. *Siano $z_1, z_2 \in \mathbb{C}$. Allora*

a) $\overline{\overline{z_1}} = z_1$;

b) $z_1 + \overline{z_1} = 2\mathrm{Re}(z_1)$;

c) $i(\overline{z_1} - z_1) = 2\mathrm{Im}(z_1)$;

d) $\overline{(z_1 + z_2)} = \overline{z_1} + \overline{z_2}$;

e) $\overline{z_1 \cdot z_2} = \overline{z_1} \cdot \overline{z_2}$;

f) *se* $z_1 \neq 0$, *allora* $\overline{z_1^{-1}} = (\overline{z_1})^{-1}$.

g) $z_1 = \overline{z_1}$ *se e solamente se* $z_1 \in \mathbb{R}$;

h) $z_1 = -\overline{z_1}$ *se e solamente se* z_1 *è immaginario puro.*

i) *se* $z_1 \cdot \overline{z_1} = |z_1|^2 \geq 0$. *Vale* 0 *se e solamente se* $z_1 = 0$;

j) $|\overline{z}_1| = |z_1|$;

k) $|z_1 \cdot z_1| = |z_1||z_2|$;

l) *se* $z_1 \neq 0$, *allora* $|z^{-1}| = \frac{1}{|z|}$;

m) $|\mathrm{Re}(z)| \leq |z|$; $|\mathrm{Im}(z)| \leq |z|$; $|z| \leq |\mathrm{Re}(z)| + |\mathrm{Im}(z)|$;

n) $|z_1 + z_2| \leq |z_1| + |z_2|$;

1.5 Forma Trigonometrica e radici

Sia $z = a + ib \neq 0$.

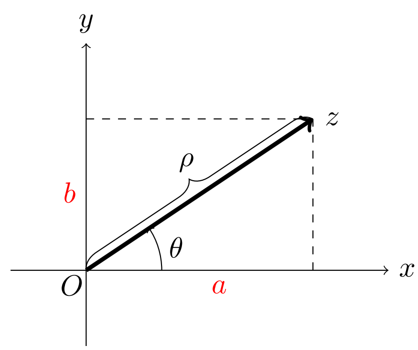


Figura 1.5.1.

In coordinate polari (Figura 1.5.1), $a = \rho \cos\theta$ e $b = \rho \sin\theta$ per cui

$$z = \rho(\cos\theta + i\ \sin\theta) = \rho e^{i\theta},$$

dove $e^{i\theta} := \cos\theta + i\sin\theta$ è un numero complesso di norma $|e^{i\theta}| = 1$ unitaria. Quindi un numero complesso z si può scrivere sia in forma cartesiana, $z = a + ib$ sia in *forma trigonometrica* $z = \rho e^{i\theta} = \rho(\cos\theta + i\sin\theta)$.

Se $z = \rho e^{i\theta}$ e $w = \psi e^{i\psi}$, allora $z = w$ se e solamente se $\rho = \psi$ e $\theta = \psi + 2\pi s$ per un certo $s \in \mathbb{Z}$.

Siano $z, w \in \mathbb{C}$ e siano $z = \rho_1 e^{i\theta_1}$ e $w = \rho_2 e^{i\theta_2}$. Allora

$$z \cdot w = \rho_1 \rho_2 e^{i(\theta_1+\theta_2)}.$$

Se $z \neq 0$, allora

$$z^{-1} = \rho^{-1} e^{-i\theta}.$$

Sia $n \in \mathbb{N}$. Definiamo

$$z^n := \underbrace{z \cdots z}_{n}.$$

È facile verificare che se $z^{n+m} = z^n z^m$ per ogni $n, m \in \mathbb{N}$.

Se $z \neq 0$ possiamo definire le potenze per ogni $n \in \mathbb{Z}$. Poniamo $z^0 = 1$; se $n < 0$, definiamo

$$z^n := (z^{-1})^{-n}.$$

È facile verificare che per ogni $n, m \in \mathbb{Z}$, allora

$$z^{n+m} = z^n z^m$$

Se scriviamo z in forma trigonometrica si ha il seguente risultato.

Lemma 1.6 (Formula di De Moivre)**.** *Sia $z = \rho e^{i\theta} \neq 0$. Allora per ogni $n \in \mathbb{Z}$ si ha*

$$z^n = \rho^n e^{in\theta}.$$

Definizione 1.7. *Sia $w \in \mathbb{C}$ e sia n un intero maggiore oppure uguale a 1. Una radice n-esima di w è un numero complesso z tale che $z^n = w$.*

Proposizione 1.8. *Sia $w \in \mathbb{C}$ differente da zero e sia $n \geq 1$, Allora w ha esattamente n radice distinte $z_0, \ldots, z_{n-1}$ di w. Per la precisione se $w = \rho e^{i\theta}$, allora*

$$z_k = \rho^{1/n} e^{i(\frac{\theta + 2k\pi}{n})},$$

per $k = 0, \ldots, n-1$.

Dimostrazione. La dimostrazione seguirà i seguenti passi: z_k per $k = 0, \ldots, n-1$ sono radici n-esime di w; ogni radice n-esima di w coincide con z_i per un certo $i \in \{0, \ldots, n-1\}$; le radici z_k per $k = 0, \ldots, n-1$ sono a due a due distinte.

Il primo passo è immediato. Applicando il lemma precedente è facile verificare che z_k è una radice n-esime di w per $k = 0, \ldots, n-1$.

Sia $z \in \mathbb{C}$ tale che $z^n = w$. Se $z = \mu e^{i\psi}$, tenendo in mente il Lemma di Moivre, allora $\mu = \rho^{1/n}$ e $n\psi = \theta + 2m\pi$, per un certo $m \in \mathbb{Z}$. Quindi $z = z_i$ per un certo $l \in \{0, \ldots, n-1\}$. Adesso dimostrariamo che i vettori $z_0, \ldots, z_{n-1}$ sono a due a due distinti.

Se $z_i = z_j$ con $i \neq j$, allora

$$\frac{(\theta + 2i\pi)}{n} = \frac{(\theta + 2j\pi)}{n} + 2s\pi$$

per un certo $s \in \mathbb{Z}$. Quindi

$$(i - j) = ns.$$

Poiché

$$|i - j| < n$$

si ha un assurdo. □

Le radici n-esime dell'unità sono, quindi, le soluzioni dell'equazione $z^n = 1$. Abbiamo visto che per ogni naturale n, esistono esattamente n radici distinte dell'unità ed hanno tutte norma uno. Quindi le radici ennesime dell'unità formano un sottoinsieme della circonferenza di raggio uno. Poiché 1 è un numero reale, l'insieme delle radici n-esime dell'unità è invariante per coniugio, ovvero se z è una radice n-esima dell'unità allora anche $\overline{z}$ è una radice n-esima dell'unità. Non vogliamo dilungarci sulle radici n-esime dell'unità poiché non è un argomento che verrà trattato in questo corso ma vogliamo porre l'attenzione, stimolando la curiosità del lettore, su alcune simmetrie che verificano le radici n-esime dell'unità. Per questo, analizzando le radici 3 dell'unità e le radici seste dell'unità.

L'equazione $z^3 = 1$ ha tre soluzioni: tenendo in mente che $1 = e^{i0}$, le radici terze dell'unità sono $z_0 = 1$, $z_1 = e^{i\frac{2\pi}{3}}$ e $z_2 = e^{i\frac{4\pi}{3}}$ e sono i vertici di un triangolo equilatero iscritto nella circonfernenza unitaria.

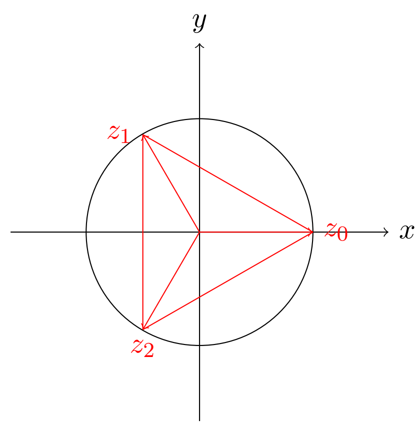


Figura 1.8.1.

Le radici 6-esime dell'unità sono: $z_0 = 1$, $z_1 = e^{i\frac{\pi}{3}}$ e $z_2 = e^{i\frac{2\pi}{3}}$, $z_3 = e^{i\pi}$, $z_4 = e^{i\frac{4\pi}{3}}$, $z_5 = e^{i\frac{5\pi}{3}}$, veritici di un esagano iscritto nella circonferenza di raggio uno.

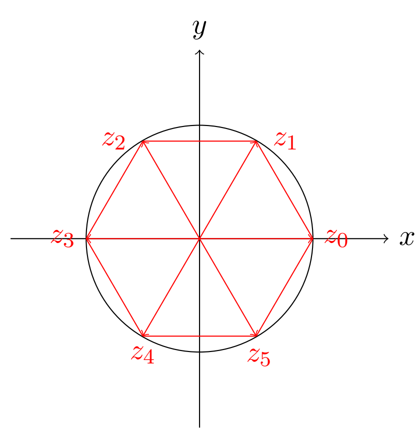


Figura 1.8.2.

1.9 Teorema di Napoleone [Facoltativa]

L'obbiettivo di questa sezione è dimostrare un risultato attribuito a Napoleone.

Teorema 1.10. *Dato un triangolo qualsiasi, si costruisca esternamente su ciascun lato, un triangolo equilatero. I baricentri dei tre triangoli equilateri costruiti sono i vertici di un triangolo equilatero.*

Esistono varie dimostrazioni in letteratura: noi proponiamo una prova che utilizza i numeri complessi. Cominciamo, dimostrando alcuni risultati preliminari che saranno utilizzati nella prova.

Lemma 1.11. *L'equazione* $z^2 + \overline{z} = 0$ *ha esattamente* 4 *soluzioni distinte:* $0, e^{i\frac{\pi}{3}}, e^{i\pi}, e^{i\frac{5}{3}\pi}$.

Dimostrazione. $z = 0$ è soluzione. Se $z \neq 0$, allora $z^2 + \overline{z} = 0$ se e solamente se $z(z^2 + \overline{z}) = 0$ e quindi se e solamente se $z^3 + |z|^2 = 0$. Se $z = \rho e^{i\theta}$, tenendo in mente che $-1 = e^{i\pi}$, si ha

$$\rho^3 e^{i3\theta} = \rho^2 e^{i\pi},$$

per cui $\rho = 1$ e $\theta = \frac{\pi + 2k\pi}{3}$, per $k = 0, 1, 2$. $\square$

Sia $z \in \mathbb{C}$ tale che $|z| = 1$. Allora $z = e^{i\theta}$ per un certo $\theta \in \mathbb{R}$. Se $w = \rho e^{i\psi}$, allora $zw = \rho e^{i(\theta+\psi)}$. Moltiplicare w per z equivale a ruotare w di una rotazione di angolo θ, per cui l'applicazione

$$\mathbb{C} \longrightarrow \mathbb{C} \qquad w \mapsto zw,$$

è la rotazione del piano di angolo θ. I numeri complessi di norma unitaria rappresentano le rotazioni del piano.

Proposizione 1.12. $w_1, w_2, w_3 \in \mathbb{C}$ *sono i vertici di un triangolo equilatero in senso anti-orario, rispettivamente orario, se e solamente se* $w_3 - w_1 = e^{i\frac{\pi}{3}}(w_2 - w_1)$*, rispettivamente* $w_3 - w_1 = e^{-i\frac{\pi}{3}}(w_2 - w_1)$

Dimostrazione. Siano $w_1, w_2, w_3 \in \mathbb{C}$ orientati in senso anti-orario.

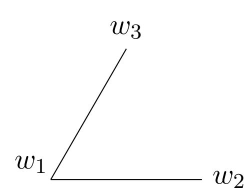


Figura 1.12.1.

Allora w_1, w_2, w_3 sono vertici di un triangolo equilatero se e solamente se $w_3 - w_1$ e $w_2 - w_1$ hanno la stessa lunghezza è l'angolo compreso è $\frac{\pi}{3}$. Quindi w_1, w_2, w_3 sono i vertici di un triangolo equilatero se e solamente se ruotando il numero complesso $w_2 - w_1$ di un angolo pari a $\frac{\pi}{3}$ ottengo il numero complesso $w_3 - w_1$, per cui se e solamente se

$$w_3 - w_1 = e^{i\frac{\pi}{3}}(w_2 - w_1)$$

Se $w_1, w_2, w_3 \in \mathbb{C}$ sono orientati in senso anti-orario,

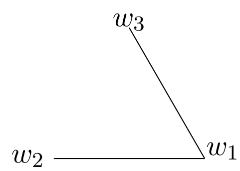


Figura 1.12.2.

allora w_1, w_2, w_3 sono vertici di un triangolo equilatero se e solamente se

$$w_3 - w_1 = e^{-i\frac{\pi}{3}}(w_2 - w_1)$$

□

Corollario 1.13. *Siano w_1 e w_2 vertici di un triangolo equilatero percorso in senso anti-orario, rispettivante orario. Allora il baricentro del triangolo equilatero è*

$$\frac{1}{3}\left((1 + e^{-i\frac{\pi}{3}})w_1 + (1 + e^{i\frac{\pi}{3}})w_2\right),$$

rispettivamente

$$\frac{1}{3}\left((1 + e^{i\frac{\pi}{3}})w_1 + (1 + e^{i\frac{-\pi}{3}})w_2\right),$$

Dimostrazione. Se il triangolo equilatero è percorso in senso antiorario

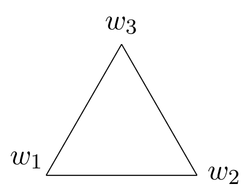


Figura 1.13.1.

applicando la Proposizione 1.12 si ha

$$w_3 = (1 - e^{i\frac{\pi}{3}})w_1 + e^{i\frac{\pi}{3}}w_2.$$

Adesso, tenendo in mente che $\frac{z+\overline{z}}{2} = \mathrm{Re}(z)$ e $\mathrm{Re}(e^{i\frac{\pi}{3}}) = \frac{1}{2}$, si ha $1 - e^{i\frac{\pi}{3}} = e^{-\frac{i\pi}{3}}$. Quindi

$$w_3 = e^{-i\frac{\pi}{3}} w_1 + e^{i\frac{\pi}{3}} w_2.$$

Il baricentro del triangolo di vertici w_1, w_2, w_3 è

$$\frac{1}{3}\left(w_1 + w_2 + w_3\right) = \frac{1}{3}\left((1 + e^{-i\frac{\pi}{3}})w_1 + (1 + e^{i\frac{\pi}{3}})w_2\right).$$

L'altro caso è analogo. Infatti, se il triangolo equilatero è percorso in senso orario

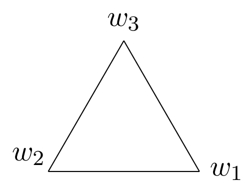


Figura 1.13.2.

applicando la Proposizione 1.12 si ha

$$w_3 = (1 - e^{-i\frac{\pi}{3}})w_1 + e^{-i\frac{\pi}{3}} w_2,$$

da cui segue che

$$w_3 = e^{i\frac{\pi}{3}} w_1 + e^{-i\frac{\pi}{3}} w_2.$$

Il baricentro del triangolo di vertici w_1, w_2, w_3 è

$$\frac{1}{3}\left(w_1 + w_2 + w_3\right) = \frac{1}{3}\left((1 + e^{i\frac{\pi}{3}})w_1 + (1 + e^{-i\frac{\pi}{3}})w_2\right).$$

□

Dimostrazione del Teorema di Napoleone. Sia

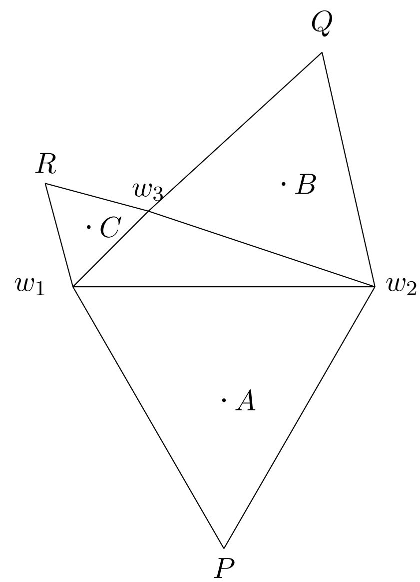


Figura 1.13.3.

un triangolo di vertici w_1, w_2, w_3. Poiché il triangolo di vertici w_1, w_2, P è orientato in senso orario, mentre i triangoli di vertici w_3, w_2, Q e w_1, w_3, R sono orientati in senso antioriario, rispettivamente, applicando il Corollario 1.13 si ha che i baricentri A, B, C dei triangoli equilateri costruiti esternamente sono

$$\begin{aligned} A &= \tfrac{1}{3}\Big((1+e^{i\frac{\pi}{3}})w_1 + (1+e^{-i\frac{\pi}{3}})w_2\Big) \\ B &= \tfrac{1}{3}\Big((1+e^{i\frac{\pi}{3}})w_2 + (1+e^{-i\frac{\pi}{3}})w_3\Big) \\ C &= \tfrac{1}{3}\Big((1+e^{-i\frac{\pi}{3}})w_1 + (1+e^{i\frac{\pi}{3}})w_3\Big) \end{aligned}$$

Per il Lemma 1.12, il triangolo di vertici A, B, C, triangolo orientato in senso anti-orario, è equilatero se e solamente se $C - A = e^{i\frac{\pi}{3}}(B - A)$. Quindi

$$\begin{aligned} C - A = {} & \left(\frac{e^{-\frac{\pi}{3}} - e^{\frac{\pi}{3}}}{3}\right) w_1 \\ & + \left(\frac{-1 - e^{-i\frac{\pi}{3}}}{3}\right) w_2 \\ & + \left(\frac{1 + e^{i\frac{\pi}{3}}}{3}\right) w_3, \end{aligned}$$

che mettendo in evidenza $e^{i\frac{\pi}{}3}$ diventa

$$C - A = e^{i\frac{\pi}{3}}\left(\left(\frac{e^{-i\frac{2\pi}{3}}-1}{3}\right)w_1\left(\frac{-e^{-i\frac{\pi}{3}}-e^{-i\frac{2\pi}{3}}}{3}\right)w_2+\left(\frac{1+e^{-i\frac{\pi}{3}}}{3}\right)w_3\right)$$

□

Adesso calcoliamo $B - A$.

$$\begin{aligned} B - A = &\left(\frac{-1-e^{i\frac{\pi}{3}}}{3}\right)w_1 \\ &+\left(\frac{e^{i\frac{\pi}{3}}-e^{-i\frac{\pi}{3}}}{3}\right)w_2 \\ &+\left(\frac{1+e^{-i\frac{\pi}{3}}}{3}\right)w_3. \end{aligned}$$

Applicando il Lemma 1.11, $e^{i\frac{\pi}{3}}$ verifica l'equazione $z^2+\overline{z}=0$, ovvero $e^{i\frac{2\pi}{3}} = -e^{-i\frac{\pi}{3}}$, dalla quale si deduce

$$e^{-i\frac{2\pi}{3}} = -e^{i\frac{\pi}{3}}.$$

Quindi

$$-1-e^{i\frac{\pi}{3}} = e^{-i\frac{2\pi}{3}}-1$$

e

$$e^{i\frac{\pi}{3}}-e^{-i\frac{\pi}{3}} = -e^{-i\frac{2\pi}{3}}-e^{-i\frac{\pi}{3}}$$

rispettivamente. In particolare

$$B - A = \left(\frac{e^{-i\frac{2\pi}{3}}-1}{3}\right)w_1+\left(-\frac{e^{-i\frac{\pi}{3}}-e^{-i\frac{2\pi}{3}}}{3}\right)w_2+\left(\frac{1+e^{-i\frac{\pi}{3}}}{3}\right)w_3$$

e quindi

$$C - A = e^{i\frac{\pi}{3}}(B - A),$$

concludendo la dimostrazione.

Capitolo 2

Geometria Euclidea

2.1 $\mathbb{R}^3$: Struttura lineare

Sia Σ lo spazio euclideo. Se fissiamo un sistema di riferimento ortogonale, ad ogni puno $P \in \Sigma$ possiamo associare, in maniera univoca, 3 numeri reali che sono determinati proiettando ortogonalmente il segmento di estremi P e O sugli x, y e z (Figura 2.1.1) Indichiamo con $\mathbb{R}^3$ l'insieme ordinato di terne di numeri reali, i.e,

$$\mathbb{R}^3 = \left\{ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} x_1, x_2, x_3 \in \mathbb{R} \right\} :$$

l'applicazione

$$F : \Sigma \longrightarrow \mathbb{R}^3 \qquad P \mapsto \begin{bmatrix} x(P) \\ y(P) \\ z(P) \end{bmatrix}$$

è una corrispondenza biunivoca che identifica lo spazio euclideo con l'insieme $\mathbb{R}^3$.

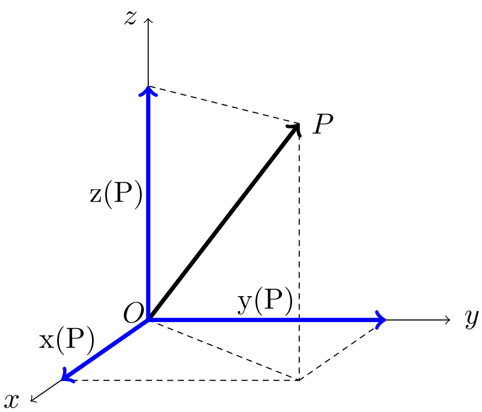


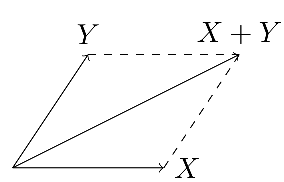
Figura 2.1.1.

Due elementi $X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$, $Y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$ di $\mathbb{R}^3$ si diranno uguali, $X = Y$, se $x_1 = y_1,\ x_2 = y_2,\ x_3 = y_3$. Sullo spazio $\mathbb{R}^3$ definiamo le seguenti operazioni:

a) somma: $\mathbb{R}^3 \times \mathbb{R}^3 \longrightarrow \mathbb{R}^3 \qquad (X, Y) \mapsto X + Y = \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \\ x_3 + y_3 \end{bmatrix};$

b) moltiplicazione per scalare $\mathbb{R} \times \mathbb{R}^3 \longrightarrow \mathbb{R}^3 \qquad (\lambda, Y) \mapsto \lambda X = \begin{bmatrix} \lambda x_1 \\ \lambda x_2 \\ \lambda x_3 \end{bmatrix}.$

La somma di due elementi di $\mathbb{R}^3$, geometricamente, si può calcolare seguendo la regola del parallelogramma. Infatti, dati $X, Y \in \mathbb{R}^3$, si costruisce il parallelogramma di lati X e Y: $X + Y$ coincide con la diagonale del parallelogramma che unisce i vertici opposti a X e Y. La moltiplicazione per



scalare si può pensare come nella seguente figura:

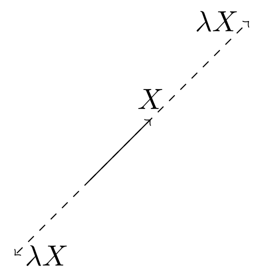


Figura 2.1.2.

La somma e la moltiplicazione per scalari godono di importanti proprietà.

Proposizione 2.2. *Dati $X, Y, Z \in \mathbb{R}^3$, $\lambda, \mu \in \mathbb{R}$ allora*

a) $X + (Y + Z) = (X + Y) + Z$*: proprierà associativa della somma;*

b) $X + Y = Y + X$*: proprietà commutativa della somma;*

c) posto $0_{\mathbb{R}^3} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$*, allora* $0_{\mathbb{R}^3} + X = X + 0_{\mathbb{R}^3} = X$*;*

d) se $X \in \mathbb{R}^3$*, allora* $X + (-1)X = (-1)X + X = 0_{\mathbb{R}^3}$*.* $-X := (-1)X$ *è detto opposto di* X*;*

e) $(\lambda + \mu)X = \lambda X + \mu X$*;*

f) $\lambda(X + Y) = \lambda X + \lambda Y$*;*

g) $(\lambda\mu)X = \lambda(\mu X) = \mu(\lambda X)$*.*

Osservazione 2.3. *Sia $\lambda \in \mathbb{R}$ e $X \in \mathbb{R}^3$. Allora*

- $\lambda 0_{\mathbb{R}^3} = 0_{\mathbb{R}^3}$*;*
- $0X = 0_{\mathbb{R}^3}$

Un elemento $X \in \mathbb{R}^3$ si dice punto o vettore. $0_{\mathbb{R}^3}$ si chiama il vettore nullo. Notazione: $X - Y := X + (-1)Y$.

Definizione 2.4. *Siano $X_1, \dots, X_k \in \mathbb{R}^3$. Si dice combinazione lineare di $X_1, \dots, X_k$ ogni elemento $Z \in \mathbb{R}^3$ per cui esistano $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ tali che $Z = \lambda_1 X_1 + \cdots + \lambda_k X_k$. I numeri reali $\lambda_1, \dots, \lambda_k$ si chiamano i coefficienti della combinazione lineare.*

Siano $X_1, \dots, X_k \in \mathbb{R}^3$. Il vettore nullo è combinazione lineare dei vettori $X_1, \dots, X_k$. Infatti

$$0_{\mathbb{R}^3} = 0X_1 + \cdots + 0X_k.$$

Definizione 2.5. *Siano $X_1, \dots, X_k \in \mathbb{R}^3$. Diremo che:*

- *i vettori $X_1, \dots, X_k \in \mathbb{R}^3$ sono linearmente dipendenti se esitono $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ non tutti nulli, tali che*

$$\lambda_1 X_1 + \cdots + \lambda_k X_k = 0_{\mathbb{R}^3};$$

- *i vettori $X_1, \dots, X_k \in \mathbb{R}^3$ sono linearmente indipendenti se non sono linearmente dipendenti. Quindi se*

$$\lambda_1 X_1 + \cdots + \lambda_k X_k = 0_{\mathbb{R}^3},$$

  allora $\lambda_1 = \cdots = \lambda_k = 0$.

Esempio 2.6. *Siano $Z = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$, $X_1 = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$, $X_2 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$, $X_3 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$. Stabilire se Z è combinazione lineare di X_1, X_2, X_3 significa verificare se esistano $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}$ tale che*

$$\begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} = \lambda_1 \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} + \lambda_2 \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} + \lambda_3 \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix},$$

ovvero

$$\begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} = \begin{bmatrix} \lambda_1 + \lambda_2 + \lambda_3 \\ \lambda_1 + 2\lambda_2 \\ -\lambda_1 + \lambda_2 - \lambda_3 \end{bmatrix}.$$

Quindi Z è combinazione lineare dei vettori X_1, X_2, X_3 se e solamente se il sistema lineare

$$\begin{cases} \lambda_1 + \lambda_2 + \lambda_3 = 1 \\ \lambda_1 + 2\lambda_2 = 2 \\ -\lambda_1 + \lambda_2 - \lambda_3 = -1 \end{cases}$$

ammette soluzioni.

Esempio 2.7. *Siano* $e_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $e_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$, $e_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$. *Ogni vettore di* $\mathbb{R}^3$ *si scrive come combinazioni lineare dei vettori* e_1, e_2, e_3.

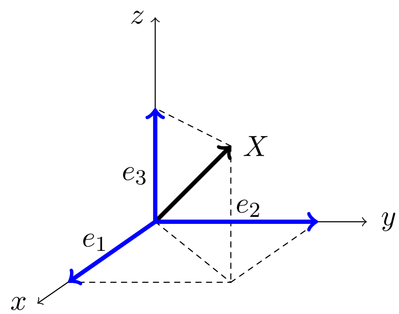


Figura 2.7.1.

Infatti, se $X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^3$, *allora*

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = x_1 e_1 + x_2 e_2 + x_3 e_3.$$

I vettori e_1, e_2, e_3 *sono anche linearmente indipendenti poiché*

$$\alpha_1 e_1 + \alpha_2 e_2 + \alpha_3 e_3 = 0$$

implica

$$\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

ovvero $\alpha_1 = \alpha_2 = \alpha_3 = 0$.

Esempio 2.8. *Siano* $\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 5 \end{bmatrix}$. *Per stabilire se i vettori sono linearmente dipendenti, rispettivamente indipendenti, dobbiamo studiare per quali valori di* $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{R}$, *si ha*

$$\alpha_1 \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} + \alpha_2 \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix} + \alpha_3 \begin{bmatrix} 1 \\ 0 \\ 5 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Sviluppando il termine di sinistra, si ha

$$\begin{bmatrix} \alpha_1 + 2\alpha_2 + \alpha_3 \\ \alpha_1 + 3\alpha_2 \\ 2\alpha_1 + \alpha_2 + 5\alpha_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Quindi stabilire se i vettori $\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 5 \end{bmatrix}$ *sono linearmente indipendenti oppure dipendenti significa studiare il seguente sistema lineare*

$$\begin{cases} \alpha_1 + 2\alpha_2 + \alpha_3 = 0 \\ \alpha_1 + 3\alpha_2 = 0 \\ 2\alpha_1 + \alpha_2 + 5\alpha_3 = 0. \end{cases}.$$

Se l'unica soluzione del sistema è $\alpha_1 = \alpha_2 = \alpha_3 = 0$, allora i vettori sono linearmente indipdenti; se il sistema ammette soluzioni con $\alpha_1, \alpha_2, \alpha_3$ non tutti nulli, allora i vettori sono linearmente dipendenti.

Proposizione 2.9. *Siano $X, Y \in \mathbb{R}^3$. I vettori X e Y sono vettori linearmente dipendenti se e solo se uno è multiplo dell'altro ovvero esiste $\lambda \in \mathbb{R}$ tale che $X = \lambda Y$ oppure $Y = \lambda X$.*

Dimostrazione. Se $X = \lambda Y$ allora $X - \lambda Y = 0_{\mathbb{R}^3}$. Quindi i vettori X e Y sono linearmente dipendenti poiché $0_{\mathbb{R}^3}$ è combinazione lineare a coefficienti non tutti nulli dei vettori X e Y . Il caso $Y = \lambda X$ è analogo.

Viceversa, supponiamo che X e Y sono linearmente dipendenti. Allora esistono α e β non entrambi nulli, tali che

$$\alpha X + \beta Y = 0_{\mathbb{R}^3}.$$

Se $\alpha \neq 0$, allora $X = -\frac{\beta}{\alpha} Y$. Analogamente se $\beta \neq 0$, allora $Y = -\frac{\alpha}{\beta} X$, concludendo la dimostrazione. $\square$

Osservazione 2.10.

- $X \in \mathbb{R}^3$ *è linearmente indipendente se e solamente se* $X \neq 0_{\mathbb{R}^3}$*;*
- $0_{\mathbb{R}^3}$, $X_1, \ldots, X_k$ *sono linearmente dipendenti;*
- *Se $\mathcal{B}$ è un insieme di elementi di $\mathbb{R}^3$ linearmente indipendenti, allora ogni sottoinsieme di $\mathcal{B}$ è costituito da elementi linearmente indipendenti;*
- *se $\mathcal{B}$ è un insieme costituito di elementi di $\mathbb{R}^3$ linearmente dipendenti, allora ogni sovrainsieme di $\mathcal{B}$ è costituito da elementi linearmente dipendenti.*

2.11 $\mathbb{R}^3$: struttura metrica

In questa sezione analizzeremo la struttura metrica dello spazio euclideo. L'obbiettivo è descrivere le nozioni geometriche come lunghezza di un vettore e angolo fra due vettori non nulli attraverso il prodotto scalare canonico di $\mathbb{R}^3$.

Definizione 2.12. *Il prodotto scalare canonico è una funzione che associa ad ogni coppia di vettori un numero reale come segue:*

$$\mathbb{R}^3 \times \mathbb{R}^3 \longrightarrow \mathbb{R} \qquad (X,Y) \mapsto \langle X,Y\rangle = x_1y_1 + x_2y_2 + x_3y_3,$$

per ogni $X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$, $Y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} \in \mathbb{R}^3$

Il prodotto scalare canonico soddisifa alle seguenti proprietà:

Proposizione 2.13. *Se $X, Y, Z \in \mathbb{R}^3$ e $\lambda, \mu \in \mathbb{R}$, allora*

(a) $\langle X,X\rangle \geq 0$ *e* $\langle X,X\rangle = 0$ *se e solamente se* $X = 0_{\mathbb{R}^3}$*;*

(b) $\langle X,Y\rangle = \langle Y,X\rangle$*;*

(c) $\langle X+Z,Y\rangle = \langle X,Y\rangle + \langle Z,Y\rangle$*;*

(d) $\langle \lambda X,Y\rangle = \lambda\langle X,Y\rangle$*;*

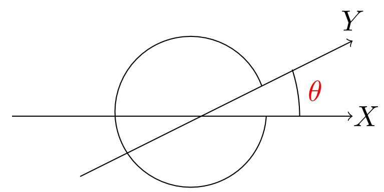
(e) $\langle X,Y+Z\rangle = \langle X,Y\rangle + \langle X,Z\rangle$*;*

(e) $\langle X,\lambda Y\rangle = \lambda\langle X,Y\rangle$*.*

Dimostrazione. Dimostreremo solamente la prima proprietà lascendo le rimanenti per esercizio.

Se $X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$, allora $\langle X,X\rangle = x_1^2 + x_2^2 + x_3^2 \geq 0$. Inoltre $\langle X,X\rangle = 0$ se e solamente se $x_1 = x_2 = x_3 = 0$, quindi se e solamente se $X = 0_{\mathbb{R}^3}$. □

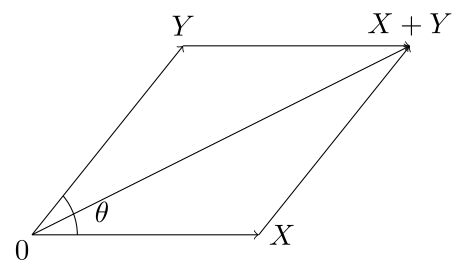
Sia $X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$. La *norma* o *lunghezza* di X è il numero reale non negativo $\parallel X \parallel := \sqrt{\langle X,X\rangle} = \sqrt{x_1^2 + x_2^2 + x_3^2}$. Definiamo la *distanza* fra due vettori X, Y il numero reale non negativo $d(X,Y) = \parallel X - Y \parallel$. In



particolare la lunghezza di un vettore $X \in \mathbb{R}^3$ è la distanza di X dal vettore nullo.

Siano $X, Y \in \mathbb{R}^3$ non entrambi nulli. I vettori X e Y dividono un piano in quattro regioni e due angoli. L'angolo fra X e Y e per definizione l'angolo minore o uguale a π che nella figura anteriore è stato indicato con θ. Affermiamo che Il prodotto scalare determina $\cos\theta$, dove θ è l'angolo fra X e Y.

Il vettore $X + Y$ è la diagonale del parallalegrammo costruito a partire da X e Y, che unisce i vertici opposti a X e Y. Applicando il teorema di



Carnot al triangolo di vertici $0, X, X + Y$, si ha

$$\| X + Y \|^2 = \| X \|^2 + \| Y \|^2 - 2 \| X \| \| Y \| \cos\psi,$$

dove ψ è l'angolo opposto al lato $X + Y$. Poiché $\theta + \psi = \pi$, si ha

$$\| X + Y \|^2 = \| X \|^2 + \| Y \|^2 + 2 \| X \| \| Y \| \cos\theta.$$

Per le proprietà del prodotto scalare, si ha

$$\begin{aligned}
\| X + Y \|^2 &= \langle X + Y, X + Y \rangle \\
&= \langle X, X + Y \rangle + \langle Y, X + Y \rangle \\
&= \langle X, X \rangle + \langle X, Y \rangle + \langle Y, X \rangle + \langle Y, Y \rangle \\
&= \| X \|^2 + 2\langle X, Y \rangle + \| Y \|^2
\end{aligned}$$

da cui segue

$$\langle X, Y\rangle = \cos\theta \parallel X \parallel\parallel Y \parallel .$$

Poiché X e Y sono entrambi non nulli, abbiamo dimostrato la seguente formula:

$$\cos\theta = \frac{\langle X, Y\rangle}{\parallel X \parallel\parallel Y \parallel}.$$

Vediamo alcune conseguenze della formula anteriore.

Corollario 2.14. *Siano $X, Y \in \mathbb{R}^3$ vettori non nulli. Allora*

a) L'angolo fra X e Y è acuto, rispettivamente ottuso, se e solamente se $\langle X, Y\rangle > 0$, rispettivamente $\langle X, Y\rangle < 0$;

b) X e Y sono ortogonali se e solamente se $\langle X, Y\rangle = 0$;

c) (Disuguaglianza di Cauchy-Schwartz)

$$|\langle X, Y\rangle| \leq \parallel X \parallel\parallel Y \parallel .$$

L'uguaglianza vale se e solamente se X e Y sono linearmente dipendenti.

La discussione anteriore suggerisce la seguente definizione.

Definizione 2.15. *Diremo che due vettori $X, Y \in \mathbb{R}^3$ sono ortogonali, oppure perpendicolari, se il loro prodotto scalare è nullo, i.e., $\langle X, Y\rangle = 0$.*

Sia S un sottoinsieme di $\mathbb{R}^3$. Indicheremo con $S^\perp$ l'insieme dei vettori di $\mathbb{R}^3$ ortogonali ad ogni vettore di S. Quindi

$$S^\perp := \{X \in \mathbb{R}^3 : \langle X, s\rangle = 0 \ \forall s \in S\}.$$

Poiché $\langle 0_{\mathbb{R}^3}, X\rangle = 0$ per ogni $X \in \mathbb{R}^3$, il vettore nullo appartiene a $S^\perp$.

Proposizione 2.16. *Se $v, w \in S^\perp$ e $\lambda, \mu \in \mathbb{R}$, allora $\lambda v + \mu w \in S^\perp$.*

Dimostrazione. Per ipotesi $\langle v, s\rangle = \langle w, s\rangle = 0$ per ogni $s \in S$. Noi dobbiamo dimostrare che

$$\langle \lambda v + \mu w, s\rangle = 0$$

per ogni $s \in S$. Per le proprietà del prodotto scalare si ha

$$\begin{aligned}\langle \lambda v + \mu w, s\rangle &= \langle \lambda v, s\rangle + \langle \mu w, s\rangle \\ &= \lambda\langle v, s\rangle + \mu\langle w, s\rangle \\ &= \lambda 0 + \mu 0 \\ &= 0.\end{aligned}$$

$\square$

Teorema 2.17. *Siano $X, Y \in \mathbb{R}^3$. Allora $\| X \|^2 + \| Y \|^2 = \| X + Y \|^2$ se e solamente se X, Y sono ortogonali.*

Dimostrazione. Per le proprietà del prodotto scalare si ha:

$$\begin{aligned} \| X + Y \|^2 &= \langle X + Y, X + Y \rangle \\ &= \langle X, X + Y \rangle + \langle Y, X + Y \rangle \\ &= \langle X, X \rangle + \langle X, Y \rangle + \langle Y, X \rangle + \langle Y, Y \rangle \\ &= \| X \|^2 + 2\langle X, Y \rangle + \| Y \|^2 . \end{aligned}$$

Quindi $\| X + Y \|^2 = \| X \|^2 + \| Y \|^2$ se e solamente se $\langle X, Y \rangle = 0$ quindi se e solamente se X, Y sono vettori perpendicolari. □

2.18 $\mathbb{R}^3$: prodotto vettoriale

Siano $X, Y \in \mathbb{R}^3$ con $X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ e $Y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$. Definiamo il loro prodotto vettoriale, che indicheremo con $X \times Y$, come il vettore

$$X \times Y = \begin{bmatrix} x_2y_3 - x_3y_2 \\ -x_1y_3 + x_3y_1 \\ x_1y_2 - x_2y_1 \end{bmatrix}.$$

Il prodotto vettoriale definisce una applicazione $\mathbb{R}^3 \times \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ che soddisfa alle seguenti proprietà:

Proposizione 2.19. *Siano $X, Y, Z \in \mathbb{R}^3$ e $\lambda, \mu \in \mathbb{R}$. Allora*

- *$X \times Y = -(Y \times X)$, i.e., il prodotto vettoriale è antisimmetrico;*
- *$(X + Y) \times Z = (X \times Z) + (Y \times Z)$;*
- *$(\lambda X) \times Z = \lambda(X \times Z)$;*
- *$X \times (Y + Z) = (X \times Y) + (X \times Y)$;*
- *$X \times (\lambda Y) = \lambda(X \times Y)$;*
- *$X \times Y$ è ortogonale sia al vettore X sia al vettore Y;*
- *$X \times Y = 0$ se e solamente se X e Y sono linearmente dipendenti;*

- $\| X \times Y \| = \| X \| \| Y \| \sin\theta$, *dove* θ *è l'angolo fra* X *e* Y*. Quindi la norma del prodotto vettoriale è l'area del parallelogramma di lati* X *e* Y *;*

Dimostrazione. Dimostreremo solamente le ultime 3 proprietà lasciando per esercizio le altre.

$$\begin{aligned}\langle X, X \times Y\rangle &= x_1(x_2y_3 - x_3y_2) + x_2(-y_3x_1 + x_3y_1) + x_3(x_1y_2 - y_1x_2)\\ &= x_1x_2y_3 + x_2x_3y_1 + x_3x_1y_2\\ &- \ (x_1x_2y_3 + x_2x_3y_1 + x_3x_1y_2)\\ &= 0.\end{aligned}$$

Analogamente è possibile dimostrare che $\langle Y, X \times Y\rangle = 0$.

Siano $X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, Y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} \in \mathbb{R}^3$ tali che $X \times Y = 0$. Allora

$$\begin{cases} x_2y_3 - x_3y_2 = 0 \\ -x_1y_3 + x_3y_1 = 0 \\ x_1y_2 - x_2y_1 = 0 \end{cases}$$

Se $X = 0_{\mathbb{R}^3}$ allora ho finito. Supponiamo quindi $X \neq 0_{\mathbb{R}^3}$. Se $x_1 \neq 0$, allora

$$\begin{cases} y_3 = \frac{y_1}{x_1}x_3 \\ y_2 = \frac{y_1}{x_1}x_2 \\ y_1 = \frac{y_1}{x_1}x_1 \end{cases}$$

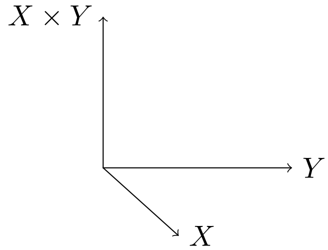
ovvero $Y = \frac{y_1}{x_1}X$. Analogamente gli altri casi.

Siano $X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, Y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} \in \mathbb{R}^3$. Allora

$$\begin{aligned}\| X \times Y \|^2 =& (x_2y_3 - x_3y_2)^2 + (-y_3x_1 + x_3y_1)^2 + (x_1y_2 - y_1x_2)^2\\ =& (x_2y_3)^2 + (x_3y_2)^2 + (y_3x_1)^2 + (x_3y_1)^2 + (x_1y_2)^2 + (y_1x_2)^2\\ &-2\big(x_2y_3x_3y_2 + y_3x_1x_3y_1 + x_1y_2y_1x_2\big)\\ =& x_1^2(y_1^2 + y_2^2 + y_3^2) - x_1^2y_1^2 + x_2^2(y_1^2 + y_2^2 + y_3^2) - x_2^2y_2^2\\ &+x_3^2(y_1^2 + y_2^2 + y_3^2) - x_3^2y_3^2 - 2\big(x_2y_3x_3y_2 + y_3x_1x_3y_1 + x_1y_2y_1x_2\big)\\ =& \| X \|^2 \| Y \|^2 - \big((x_1y_1)(x_1y_1 + x_2y_2 + x_3y_3) + (x_2y_2)((x_1y_1 + x_2y_2 + x_3y_3)\\ &+(x_3y_3)(x_1y_1 + x_2y_2 + x_3y_3)\big)\\ =& \| X \|^2 \| Y \|^2 - \langle X, Y\rangle^2\\ =& \| X \|^2 \| Y \|^2 \ (1 - \cos^2(\theta)) = \| X \|^2 \| Y \|^2 \sin^2\theta.\end{aligned}$$

□

Geometricamente, se X e Y sono vettori linearmente indipendenti, allora la seguente figura



descrive il prodotto vettoriale.

2.20 Rette nello spazio

Possiamo pensare ad una retta nello spazio come ad una particolare traiettora di un punto che si muove, sempre secondo una certa direzione (e verso). Quindi, se pensiamo al parametro t come al tempo, i punti di una retta sono descritti da una equazione parametrica

$$r : X = P + tA,$$

dove $t \in \mathbb{R}$ è il parametro, P è un punto delle retta ed infine A è un vettore non nullo, chiamato *vettore direttore*, che ne indica la direzione (Figura 2.20.1).

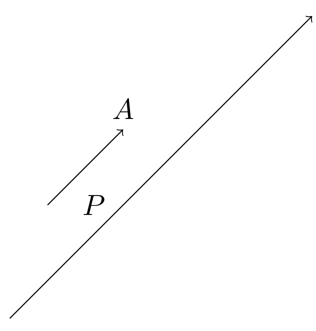


Figura 2.20.1.

Come insieme

$$r = \{X \in \mathbb{R}^3 : X = P + tA, t \in \mathbb{R}\}.$$

Una retta ha molte possibili equazioni parametriche. Se scegliamo un qualsiasi altro punto $Q \in r$, ed come vettore direttore $B = kA$ dove $k \neq 0$, allora $r = Q + sB$, $s \in \mathbb{R}$, per cui i due insiemi

$$\{X \in \mathbb{R}^3 : X = P + tA, t \in \mathbb{R}\} = \{X \in \mathbb{R}^3 : X = Q + sB, s \in \mathbb{R}\}.$$

coincidono. Infatti, sia $t_o \in \mathbb{R}$ tale che $Q = P + t_o A$. Per ogni $s \in \mathbb{R}$ si ha

$$Q + sB = P + t_o A + ksA = P + (t_o + ks)A.$$

Viceversa, per ogni $t \in \mathbb{R}$, si ha

$$\begin{aligned} P + tA &= P + tA + t_o A - t_o A \\ &= Q - t_o A + tA \\ &= Q + (t - t_o)A \\ &= Q + \frac{(t - t_0)}{k} B. \end{aligned}$$

Dato un punto P, le rette che passano per P sono tutte e sole le rette date da equazioni parametriche

$$X = P + tA, \ t \in \mathbb{R},$$

dove A è un vettore non nullo. L'insieme delle rette passanti per P si chiama *stella di rette* di centro P. Per due punti distinti P_1 e P_2 passa una ed una sola retta. Un'equazione parametriche è data da

$$r = P_1 + t(P_1 - P_2), \ t \in \mathbb{R}.$$

Verifichiamo che r è la retta passante per P_1 e P_2. Infatti, se $t = 0$, otteniamo P_1; se $t = -1$ otteniamo P_2. Proviamo che tale retta è unica.

Sia $r : Q + tA$ una retta passante per P_1 e P_2. Allora $P_1 = Q + t_1 A$ e $P_2 = Q + t_2 A$ con $t_1, t_2 \in \mathbb{R}$. Poiché $P_1 \neq P_2$, ne segue che $t_1 \neq t_2$ e $A = \frac{1}{t_1 - t_2}(P_1 - P_2)$. Se $t \in \mathbb{R}$, allora

$$\begin{aligned} Q + tA &= Q + t_1 A + (t - t_1)A \\ &= P_1 + (t - t_1)A \\ &= P_1 + \frac{t - t_1}{t_1 - t_2}(P_1 - P_2). \end{aligned}$$

Siano $r_1 = P_1 + tA_1$ e $r_2 = P_2 + tA_2$ due rette nello spazio. Diremo che le rette sono *ortogonali* o *perpendicolari* se $\langle A_1, A_2 \rangle = 0$. Le posizione reciproca di due rette nello spazio sono le seguenti: r_1 ed r_2 sono coincidenti, e scriveremo $r_1 = r_2$, se ogni elemento che appartiene a r_1 è un elemento di r_2 e viceversa. In particolare i vettori direttori sono linearmente dipendenti da cui segue che $A_\times A_2 = 0_{\mathbb{R}^{\not\vDash}}$. Poiché per due punti passa una ed una sola retta, $r_1 = r_2$ se e solo se hanno due punti in comune. In maniera equivalente è possibile dimostrare che $r_1 = r_2$ se e solamente se $P_1 \in r_2$ e $A_1 \times A_2 = 0_{\mathbb{R}^{\not\vDash}}$.

Le rette r_1 e r_2 si dicono *parallele* se non hanno punti in comune ed hanno la stessa direzione (vedi figura 2.20.2).

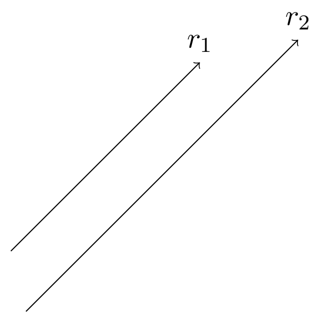


Figura 2.20.2.

I vettori direttori, quindi, di due rette parallele sono proporzionali, ovvero A_1 e A_2 sono linearmente dipendenti. Quindi due rette sono parallele se e solamente se r_1 e r_2 non hanno punti in comune e $A_1 \times A_2 = 0_{\mathbb{R}^3}$.

Le rette r_1 e r_2 sono *incidenti* se hanno esattamente un punto in comune (vedi figura 2.20.3). I vettori direttori sono necessariamente linearmente indipendenti ($A_1 \times A_2 \neq 0_{\mathbb{R}^3}$) altrimenti le due rette sarebbero coincidenti.

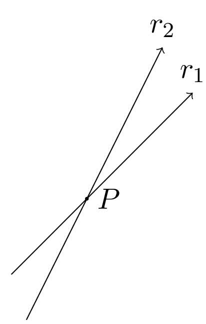


Figura 2.20.3.

Infine, due rette si dicono *sghembe* se non sono nè incidenti nè parallele. Quindi r_1 e r_2 sono sghembe se e solamente se non hanno punti in comune ed i vettori direttori sono linearmente indipendenti (vedi figura 2.20.4).

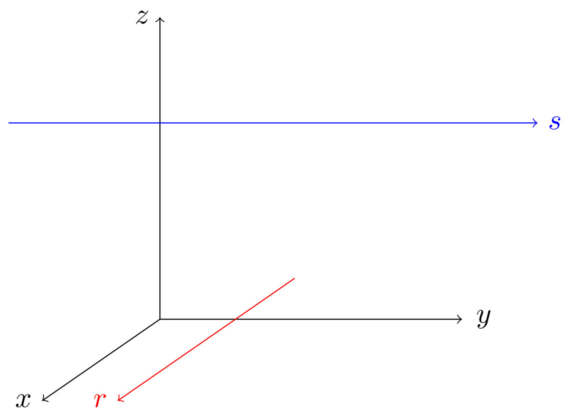


Figura 2.20.4.

2.21 Piani nello spazio

Possiamo pensare ad un piano passante per l'origine come l'insieme dei vettori ortogonali ad un vettore non nullo (vedi figura 2.21.1). Se $n = \begin{bmatrix} a \\ b \\ c \end{bmatrix} \neq 0$, allora il piano passante per l'origine e ortogonale ad n è l'insieme

$$\{n\}^{\perp} = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \in \mathbb{R}^3 : \langle X, n \rangle = 0 \right\} = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \in \mathbb{R}^3 : \ ax + by + cz = 0 \right\}.$$

Il vettore n si chiama *vettore normale al piano*.

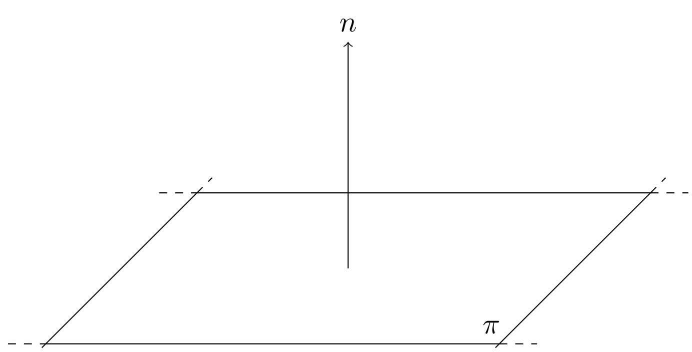


Figura 2.21.1.

Per le proprietà del prodotto scalare è facile verificare che se $k \neq 0$, allora $\{kn\}^{\perp} = \{n\}^{\perp}$. Per questo diremo che n è un vettore ortogonale a π. Inoltre, se $r = tn$, $t \in \mathbb{R}$ è la retta passante per l'origine con vettore direttore

n, allora $\pi = \{r\}^{\perp}$, per cui un piano passante per l'origine è l'insieme dei vettori ortogonali ad una retta passante per l'origine.

Sia $P \in \mathbb{R}^3$. L'idea per definire un piano passante per P è quella di traslare lungo il vettore individuato da P, un piano passante per l'origine. Per questo, Il piano passante per P e ortogonale al vettore n non nullo è invece descritto dall'insieme

$$\begin{aligned} \pi : &= \left\{ X \in \mathbb{R}^3 : \langle X - P, n \rangle = 0 \right\} \\ &= \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \in \mathbb{R}^3 : ax + bx + cx + d = 0 \right\}, \end{aligned}$$

dove $d = -\langle P, n \rangle$. Se $P = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$ e $n = \begin{bmatrix} 1 \\ -5 \\ 3 \end{bmatrix}$, allora il piano π passante per P e ortogonale a n è l'insieme dei vettori $\begin{bmatrix} x \\ y \\ z \end{bmatrix} \in \mathbb{R}^3$ che verificano la seguente equazione:

$$\pi : x - 5y + 3z + 2 = 0.$$

Un vettore $v \in \mathbb{R}^3$ appartiene a π se le sue coordinate soddisfano l'equazione precedente. Per esempio, $\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \notin \pi$ poiché $1 - 5 - 3 + 2 \neq 0$; invece $\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \in \pi$ poiché $0 - 5 + 3 + 2 = 0$. Riassumento, un piano è l'insieme dei vettori di $\mathbb{R}^3$ che verificano una equazione cartesiane del tipo

$$\pi : ax + by + cz + d = 0,$$

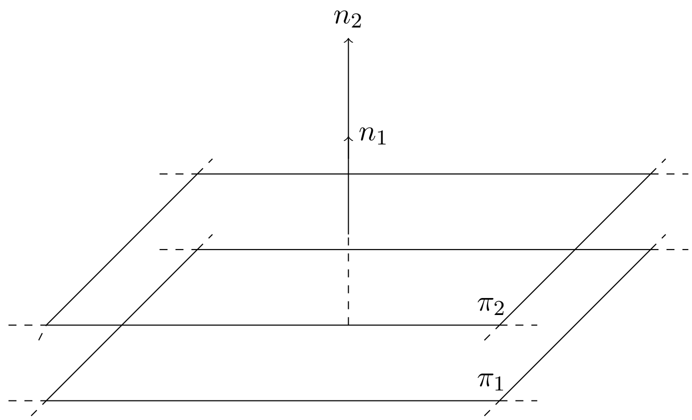
dove $a, b, c \in \mathbb{R}$ non sono tutti nulli. Inoltre, il vettore $n = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$, è chiamato *vettore ortogonale al piano*. Si noti che Il piano π è passante per l'origine se e solamente se $d = 0$.

È abbastanza facile convincersi che per un punto P si possono far passare infiniti piani. Infatti, se $P = \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix}$ allora per ogni $n = \begin{bmatrix} a \\ b \\ c \end{bmatrix} \in \mathbb{R}^3$ non

nullo, il piano
$$a(x-x_0)+b(y-y_0)+c(z-z_0)=0,$$
è un piano passante per P. Sempre in maniera intuitiva è facile convincersi che esistono infiniti piani passanti per due punti distinti. Per il momento non proveremo questo fatto ma vogliamo porre l'attenzione che un piano contiene due punti distinti P e Q, contiene tutta la retta passante per P e Q (verificare!). Invece, per tre punti non allineari passa un unico piano. Ricordiamo che tre punti P_1, P_2 e P_3 si dicono non allineati se non appartengono ad una stessa retta. Caso contrario diremo che sono allineati. Possiamo esprimere la condizione di essere non allineati utilizzando il linguaggio sviluppato ed il prdotto vettoriale. Infatti, P_1, P_2, P_3 sono non allineati, rispettivamente allineati, se e solamente se i vettori $P_2 - P_1$ e $P_3 - P_1$ sono linearmente indipendenti, rispettivamente dipendenti, per cui se e solamente se $(P_2-P_1)\times(P_3-P_1)\neq 0_{\mathbb{R}^3}$, rispettivamente $(P_2-P_1)\times(P_3-P_1)=0_{\mathbb{R}^3}$. Una equazione cartesiana per il piano passante per tre punti distinti è la seguente: se indichiamo con $n=(P_2-P_1)\times(P_3-P_1)$, allora
$$\pi := \left\{X\in\mathbb{R}^3 : \langle X-P_1, n\rangle = 0\right\}.$$
Per esempio, un'equazione cartesiana per il piano passante per $\begin{bmatrix}1\\1\\2\end{bmatrix}, \begin{bmatrix}2\\0\\3\end{bmatrix}, \begin{bmatrix}3\\1\\-1\end{bmatrix}$ è:
$$3x+5y+2z=12.$$
Siano $\pi_1 : a_1x+b_1y+c_1z+d_1=0$ e $\pi_2 : a_2x+b_2y+c_2z+d_2=0$ piani nello spazio. Indichiamo con $n_1=\begin{bmatrix}a_1\\b_1\\c_1\end{bmatrix}$ e $n_2=\begin{bmatrix}a_2\\b_2\\c_2\end{bmatrix}$ i rispettivi vettori normali. Diremo che π_1 e π_2 sono *ortogonali* se n_1 e n_2 sono ortogonali, per cui se $\langle n_1, n_2\rangle = 0$. Adesso studiamo la mutua posizione di due piani nello spazio.

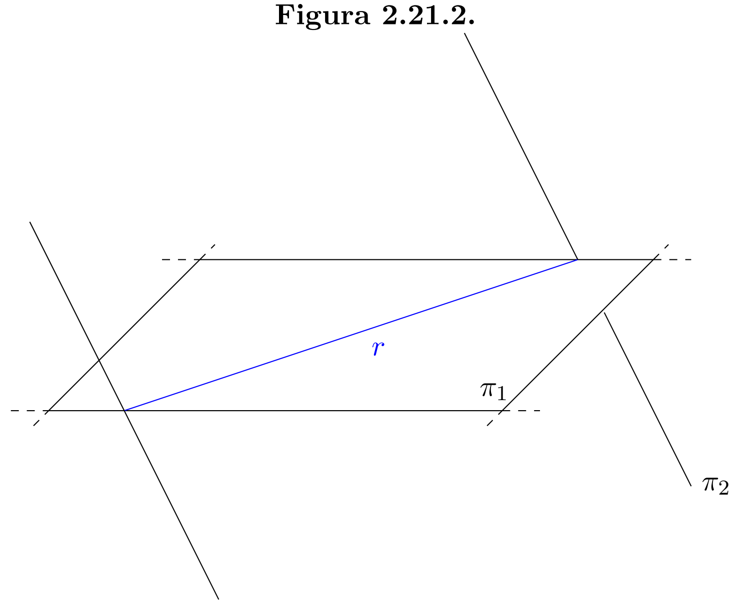
Diremo che i piani π_1 e π_2 sono:

- *coincidenti*, ed indicheremo $\pi_1=\pi_2$, se π_1 e π_2 sono lo stesso sottoinsiemi di $\mathbb{R}^3$. Algebricamente, significa che esiste $k\in\mathbb{R}$ non nullo tale che $n_1=kn_2$, in particolare $n_1\times n_2=0_{\mathbb{R}^3}$, e $d_1=kd_2$. Quindi equazioni cartesiane dei due piani differiscono per un multiplo non nullo. Forniremo una prova di questa equivalenza dopo aver sviluppato la teoria dei sistemi lineari.



- *paralleli* se π_1 e π_2 non si intersecano. È facile convincersi che se π_1 e π_2 sono paralleli allora $n_1 \times n_2 = 0_{\mathbb{R}^3}$.
- *incidenti* se si intersecano lungo una retta. Questo succede se e solamente se n_1 e n_2 sono vettori linearmente indipendenti, ovvero se e solamente se $n_1 \times n_2$ è un vettore non nullo.

Figura 2.21.2.



Calcolare equazioni parametriche per un piano significa risolvere un sistema lineare.

Sia $\pi : ax + by + cz + d = 0$ un piano nello spazio. Poiché il vettore $n = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$ è non nullo, allora a, b, c non sono tutti nulli. Supponiamo che $a \neq 0$. Possiamo ricavare la variabile x in funzione dell'altre: $x = -\frac{d}{a} - \frac{b}{a}y - \frac{c}{a}z$. Quindi

$$\begin{aligned} \pi &= \left\{ \begin{bmatrix} -\frac{d}{a} - \frac{b}{a}y - \frac{c}{a}z \\ y \\ z \end{bmatrix} : y, z \in \mathbb{R} \right\} \\ &= \left\{ \begin{bmatrix} -d/a \\ 0 \\ 0 \end{bmatrix} + y \begin{bmatrix} -b/a \\ 1 \\ 0 \end{bmatrix} + z \begin{bmatrix} -c/a \\ 0 \\ 1 \end{bmatrix} \right\}. \end{aligned}$$

Il piano π è l'insieme dei vettori della forma: $P + yv + zw$, $y, z \in \mathbb{R}$, dove: $P = \begin{bmatrix} -d/a \\ 0 \\ 0 \end{bmatrix} \in \pi$; $v = \begin{bmatrix} -b/a \\ 1 \\ 0 \end{bmatrix}$ e $w = \begin{bmatrix} -c/a \\ 0 \\ 1 \end{bmatrix}$ sono vettori linearmenti indipendenti e ortogonali a n, che sono chiamate *equazioni parametriche*. Analogamente gli altri casi: se $b \neq 0$, allora $y = -\frac{d}{b} - \frac{a}{b}x - \frac{c}{b}z$ per cui

$$\pi = \left\{ \begin{bmatrix} 0 \\ -d/b \\ 0 \end{bmatrix} + x \begin{bmatrix} 1 \\ -a/b \\ 0 \end{bmatrix} + z \begin{bmatrix} 0 \\ -c/b \\ 1 \end{bmatrix} \right\};$$

se $c \neq 0$, allora $z = -\frac{d}{c} - \frac{a}{c}x - \frac{b}{c}y$, per cui

$$\pi = \left\{ \begin{bmatrix} 0 \\ 0 \\ -d/c \end{bmatrix} + x \begin{bmatrix} 1 \\ 0 \\ -a/c \end{bmatrix} + y \begin{bmatrix} 0 \\ 1 \\ -b/c \end{bmatrix} \right\}.$$

Esempio 2.22. *Determinare equazioni parametrica per il piano $\pi : 2x - y + z = 4$. Possiamo ricavare $y = 2x + z - 4$, per cui si ha*

$$\begin{aligned} \pi &= \left\{ \begin{bmatrix} x \\ 2x + z - 4 \\ z \end{bmatrix} : x, z \in \mathbb{R} \right\} \\ &= \left\{ \begin{bmatrix} 0 \\ -4 \\ 0 \end{bmatrix} + x \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} + z \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \ x, z \in \mathbb{R} \right\}; \end{aligned}$$

Quindi π *ha equazioni paramentriche:* $\pi : P + yv + zw,\ y, z \in \mathbb{R}$, *dove* $P = \begin{bmatrix} 0 \\ -4 \\ 0 \end{bmatrix} \in \pi,\ v = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, w = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$. *Analogamente,* $z = 4 - 2x + y$ *per cui*

$$\begin{aligned} \pi &= \left\{ \begin{bmatrix} x \\ y \\ 4 - 2x + y \end{bmatrix} : x, y \in \mathbb{R} \right\} \\ &= \left\{ \begin{bmatrix} 0 \\ 0 \\ 4 \end{bmatrix} + x \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix} + y \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} x, y \in \mathbb{R} \right\}; \end{aligned}$$

Se $P_1, P_2, P_3 \in \mathbb{R}^3$ sono tre punti non allineati, per cui $(P_2 - P_1) \times (P_3 - P_1) \neq 0$, allora l'unico piano passante per P_1, P_2, P_3 ha equzioni parametriche (verificare!)

$$\pi : P_1 + t(P_2 - P_1) + s(P_3 - P_1).$$

Sia π un piano di equazioni parametriche

$$\pi : P + tv + sw, \forall s, t \in \mathbb{R},$$

con v e w linearmente indipendenti. Allora un'equazione cartesiana per il piano π è $\pi : ax + by + cx + d = 0$, dove $n = \begin{bmatrix} a \\ b \\ c \end{bmatrix} = v \times w$ e $d = -\langle P, n \rangle$.

Esempio 2.23. *Determinare un'equazione cartesiana per il piano* $\pi : \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} + t \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} + s \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix},\ s, t \in \mathbb{R}.$

$$n = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} \times \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}$$

Quindi

$$\pi : 2x + y + z = d.$$

Imponendo il passaggio per $\begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$ *si ha*

$$2x + y + z = 3.$$

Esempio 2.24. *Determinare equazioni parametriche e una equazione cartesiane per il piano* π *contenente la retta*

$$r : \begin{bmatrix} 1 \\ 2 \\ -2 \end{bmatrix} + t \begin{bmatrix} -4 \\ 3 \\ 2 \end{bmatrix}$$

e passante per $P = \begin{bmatrix} 3 \\ 3 \\ -3 \end{bmatrix}$.

Il vettore $Q = \begin{bmatrix} 1 \\ 2 \\ -2 \end{bmatrix} \in r$, *per cui* Q *deve appartenere al piano* π *cercato. Quindi* $P, Q \in \pi$, *per cui* $P - Q = \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}$ *è ortogonale al vettore normale al piano. La stessa condizione vale per un qualsiasi vettore direttore di* r. *Quindi il piano cercato ha equazioni parametriche*

$$\pi : \begin{bmatrix} 3 \\ 3 \\ -3 \end{bmatrix} + t \begin{bmatrix} -4 \\ 3 \\ 2 \end{bmatrix} + s \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}.$$

Per determinare un'equazione cartesiane basta osservare che un vettore normale al piano è dato da

$$n = (P - Q) \times A = \begin{bmatrix} 5 \\ 0 \\ 10 \end{bmatrix}.$$

Imponendo il passaggio per P *si ha*

$$5x + 10z = -15.$$

2.25 Rette e Piani nello spazio

Nella sezione precedente, abbiamo affermato che due piani non paralleli si intersecano lungo una retta. Se consideriamo un'equazione cartesiana per ciascuno piano, la retta così ottenuta è definita da due equazioni cartesiane. Vogliamo provare che ogni retta è intersezione di due piani non paralleli per cui ogni retta è definita da due equazioni caretsiane.

Sia $r : P + tA$ una retta. Se $P = \begin{bmatrix} x_o \\ y_o \\ z_0 \end{bmatrix}$ e $A = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}$ allora

$$\begin{cases} x = x_o + ta_1 \\ y = y_o + ta_2 \\ z = z_o + ta_3 \end{cases}$$

Se a_1 è differente da zero, allora possiamo ricavare

$$t = \frac{x - x_o}{a_1},$$

e sostituendo l'espressione di t nelle altre due equazione, si ha:

$$r = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \in \mathbb{R}^3 : y - y_o = \frac{a_2(x - x_o)}{a_1}, z - z_o = \frac{a_3(x - x_o)}{a_1}, \right\}.$$

Analogamente se $a_2 \neq 0$, allora

$$r = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \in \mathbb{R}^3 : x - x_o = \frac{a_1(y - y_o)}{a_2}, z - z_o = \frac{a_3(y - y_o)}{a_2}, \right\};$$

se $a_3 \neq 0$, allora

$$r = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \in \mathbb{R}^3 : y - y_o = \frac{a_2(z - z_o)}{a_3}, y - y_o = \frac{a_3(z - z_o)}{a_1}, \right\}.$$

Le equazioni precedenti si chiamano *equazioni cartesiane.* Si noti che una stessa retta ha infinite possibili equazioni cartesiane, nonché ci sono infinite coppie di piani che si intersecano lungo una retta.

Esempio 2.26.
$$r : \begin{cases} x = 1 + t \\ y = 2 + 2t \\ z = 3t \end{cases}$$
Posso ricavare $t = x - 1$ *e quindi*
$$r = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \in \mathbb{R}^3 : y - 2x = 0,\ z - 3x = -3 \right\}.$$

Si r una retta nello spazio di equazioni cartesiane
$$r := \begin{cases} ax + by + cz + d = 0 \\ a'x + b'y + c'z + d' = 0 \end{cases}$$
Determinare equazioni parametriche equivale a risolvere un sistema di 2 equazioni in 3 incognite. Al momento non abbiamo ancora sviluppato un prodecimento per risolvere un sistema lineare. Verrà affrontato completamente nel prossimo capitolo. Quello che possiamo calcolare, in questo momento, è un vettore direttore di r. Se una retta r è contenuto in un piano π allora un vettore direttore di r è ortogonale ad un vettore normale al piano (verificare). Quindi un vettore direttore di r è ottenuto calcolando il prodotto vettoriale dei vettori normali ai piani:
$$A = \begin{bmatrix} a \\ b \\ c \end{bmatrix} \times \begin{bmatrix} a' \\ b' \\ c' \end{bmatrix}.$$

Esempio 2.27. *Sia*
$$r : \begin{cases} x - y + z = 1 \\ y - 2z = 2 \end{cases}.$$
Allora un vettore direttore di r *è il vettore*
$$\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \times \begin{bmatrix} 0 \\ 1 \\ -2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}.$$

L'insieme dei piani passanti per una retta r si chiama *fascio di piani* di asse r. Se la retta r ha equazioni cartesiane
$$\begin{cases} ax + by + cz + d = 0 \\ a'x + b'y + c'z + d' = 0 \end{cases},$$

allora si può dimostrare che un piano π appartiene al fascio di piani di asse r se e solamete se esistano $\lambda, \mu \in \mathbb{R}$, non entrambi nulli, tali che π ha equazioni cartesiane

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

Il fascio dei piani di asse r può essere utilizzato per risolvere alcuni tipi di esercizi.

Esempio 2.28. *Determinare un'equazione cartesiana del piano π contenente la retta*

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

e passante per $P = \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix}$.

Un piano contenente r ha equazione cartesiana

$$\alpha(x - y - z - 1) + \beta(x + y + z - 2) = 0,$$

al variare di $\alpha, \beta \in \mathbb{R}$. Imponedo il passaggio per P si ha

$$\alpha + 2\beta = 0,$$

ovvero $\alpha = -2\beta$. Quindi se scelgo $\beta = 1$, allora $\alpha = -2$ ed il piano cercato ha equazione cartesiana

$$-x + 3y + 3z = 0.$$

Esempio 2.29. *Determinare la retta r passante per* $P = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$ *ed incidente alle rette*

$$s_1 : \begin{cases} x + z = 0 \\ 2x - y = 0 \end{cases} \qquad s_2 : \begin{cases} x - y = 1 \\ y - z = 3 \end{cases}$$

L'obbiettivo è determinare due piani non paralleli che contengono la retta cercata.

Una equazione cartesiana per il piano contenente s_1, applichiamo il metodo del fascio, è:

$$\alpha(x + z) + \beta(2x - y) = 0.$$

Imponendo il passaggio per P si ha:

$$2\alpha + 4\beta = 0,$$

ovvero $\alpha = -2\beta$. Se $\beta = -1$, allora $\alpha = 2$ ed il piano cercato è $\pi_1 : y+2z = 0$. Analogamente il fascio di piani di asse s_2 è dato da

$$\alpha(x - y - 1) + \beta(y - z - 3) = 0.$$

al variare di $\alpha, \beta \in \mathbb{R}$ non tutti nulli. Imponendo il passaggio per P si ha

$$2\alpha - 6\beta = 0,$$

otteniamo il piano $\pi_2 : 3x-2y-z = 6$. Quindi la retta la retta r ha equazioni cartesiane

$$r : \begin{cases} y + 2z = 0 \\ 3x - 2y - z = 6 \end{cases} .$$

Esempio 2.30. *Determinare un'equazione cartesiana della retta r incidente ed ortogonale alle rette*

$$s_1 : \begin{cases} x - z = 0 \\ 2x - y = 2 \end{cases} \qquad s_2 : \begin{cases} x + y = 2 \\ y + z = 4 \end{cases}$$

Poiché la retta r è ortogonale sia a s_1, sia s_2, un vettore direttore della retta r è dato dal prodotto vettoriale dei vettori direttori di s_1 e s_2, rispettivamente, ovvero $A = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$ è un vettore direttore di r. L'obbiettivo è determinare due piani non paralleli che contengono la retta cercata.

Il piano contenente s_1 e con vettore normale ortogonale al vettore direttore A di r contiene tutta la retta r. Per determinare tale piano applichiamo il metodo del fascio:

$$\alpha(x - z) + \beta(2x - y - 2) = 0.$$

Il vettore normale al piano è il vettore:

$$n_{\alpha,\beta} = \begin{bmatrix} \alpha + 2\beta \\ -\beta \\ -\alpha \end{bmatrix}.$$

Imponendo la condizione

$$\langle n_{\alpha,\beta}, A\rangle = 0 \iff \alpha + 2\beta + \alpha = 2(\alpha + \beta) = 0,$$

otteniamo il piano $\pi_1 : x - y + z = 2$. Analogamente il fascio di piani di asse s_2 è dato da

$$\alpha(x + y - 2) + \beta(y + z - 4) = 0.$$

al variare di $\alpha, \beta \in \mathbb{R}$ non tutti nulli. Imponendo la condizione

$$\langle n_{\alpha,\beta}, A\rangle = 0 \iff \alpha - \beta = 0,$$

otteniamo il piano $\pi_2 : x + 2y + z = 6$, ovvero la retta cercata ha equazione cartesiane

$$r : \begin{cases} x - y + z = 2 \\ x + 2y + z = 6 \end{cases}.$$

Definizione 2.31. *Due rette r_1 ed r_2 si dicono* complanari *se esiste un piano π che le contiene entrambe.*

Due rette sghembe non sono complanari. Se due rette sghembe fossero complanari sarebbero contenute in uno stesso piano, ma due rette contenute in un piano sono coincidenti, incidenti oppure parallele. Una contraddizione. In generale si può dimostrare il seguente risutato.

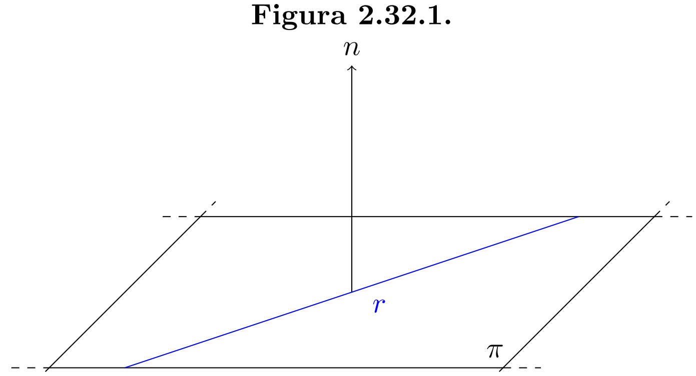
Proposizione 2.32. *Due rette sono complanari se e solamente se sono parallele oppure incidenti.*

Terminiamo la sezione analizzando la mutua posizione di una retta ed un piano nello spazio.

Sia $r : P + tA$, $t \in \mathbb{R}$ e $\pi : ax + by + cz + d = 0$. Indichiamo con $n = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$ il vettore normale al piano. Allora:

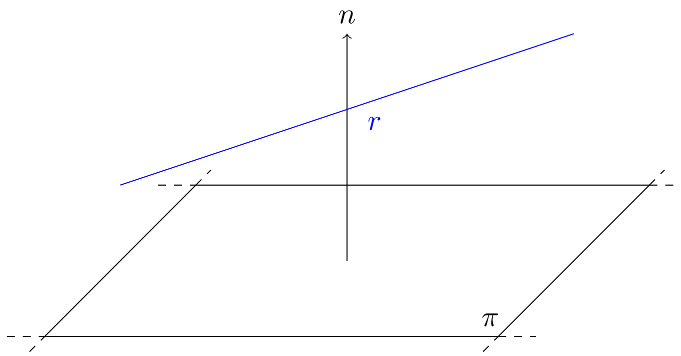
- $r \subset \pi$ se e solamente se $P \in \pi$ e $\langle A, n\rangle = 0$;

Figura 2.32.1.



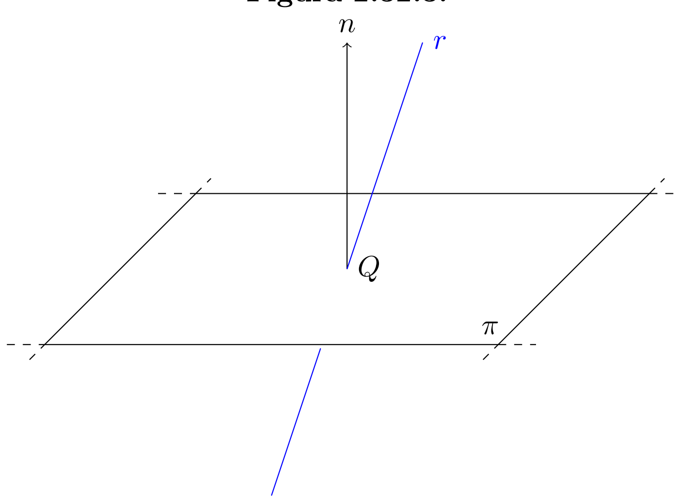
- diremo che la retta r è parallela a π se $r \cap \pi = \emptyset$. In particolare $\langle A, n \rangle = 0$;

Figura 2.32.2.



- diremo che r e π sono *incidenti* se $r \cap \pi = \{Q\}$. Si può dimostrare che r ed π sono incidenti se e solamente se $\langle A, n \rangle \neq 0$;

Figura 2.32.3.



Esempio 2.33. *Determinare equazioni cartesiane della retta r passante per* $P = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$*, parallela al piano $x + y - z = 6$ ed incidente l'asse delle x.*

La retta cercata, passa per P ed è parallela al piano $x + y - z = 6$, per cui è contenuta nel piano $\pi : x + y - z = 2$. L'intersezione del piano π e *l'asse delle x è il vettore*

$$\pi \cap \{\text{asse x}\} = \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix} = Q.$$

Quindi

$$r : P + t(P - Q),$$

ovvero

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

Possiamo ricavare t, per esempio, dalla 1 equazione ed ottenere equazioni cartesiane

$$r : \begin{cases} 2x + y = 4 \\ x + z = 2 \end{cases} .$$

Vediamo un altro metodo. Determiniamo due piani π_1 e π_2 non paralleli che contengono la retta r. Un piano parallelo a $x + y - z = 6$ ha equazioni cartesiane

$$x + y - z = d.$$

Imponendo il passaggio per P troviamo il piano che contiene la retta cercata, ovvero $x+y-z = 2$. Applicando il medoto del fascio troviamo che un piano che contiene l'asse delle x ha equzioni cartesiane

$$\alpha y + \beta z = 0.$$

Imponendo il passaggio per P, troviamo il piano contenente sia l'asse delle x e la retta r, ovvero $-y + 2z = 0$. Quindi

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

2.34 Distanze in $\mathbb{R}^3$

Consideriamo lo spazio vettoriale $\mathbb{R}^3$ dotato del prodotto scalare standard $\langle \cdot, \cdot \rangle$.

Definizione 2.35. *Dati due punti $P_1 = \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix}$ e $P_2 = \begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix}$ di $\mathbb{R}^3$, ricordiamo che la distanza $d(P_1, P_2)$ tra P_1 e P_2 è la lunghezza del vettore $P_2 - P_1$, ovvero*

$$d(P_1, P_2) = \|P_2 - P_1\| = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}.$$

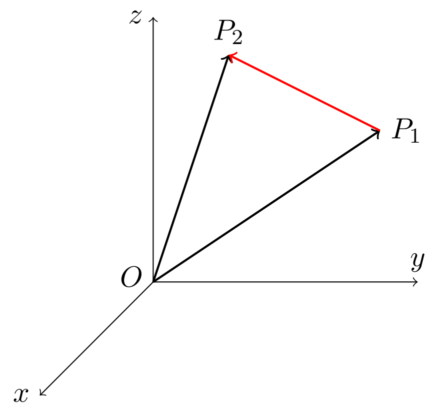


Figura 2.35.1: Distanza tra due punti.

Definizione 2.36. *Siano P un punto e r una retta in $\mathbb{R}^3$. La* proiezione ortogonale *di P su r è il punto H che si ottiene intersecando la retta r con l'unico piano passante per P e ortogonale a r.*

Definizione 2.37. *La distanza tra un punto P e una retta r in $\mathbb{R}^3$ è la distanza tra il punto P e la sua proiezione ortogonale H su r, ovvero*

$$d(P, r) = \|P - H\| .$$

Per determinare la distanza tra un punto e una retta non è necessario calcolare esplicitamente le coordinate della proiezione ortogonale, vale infatti il seguente risultato.

Proposizione 2.38. *Siano P un punto e r una retta in $\mathbb{R}^3$. Se Q è un punto di r e A è un suo vettore direttore, allora*

$$d(P, r) = \frac{\|(Q - P) \times A\|}{\|A\|}.$$

Facoltativa. Sia H la proiezione ortogonale di P su r. Il triangolo HQP è rettangolo. Indichiamo con ϑ l'ampiezza dell'angolo $H\widehat{Q}P$ (cf. Figura 2.38.1). A seconda del verso del vettore A, l'ampiezza dell'angolo compreso tra A e il vettore $\overline{QP}$ è ϑ o $\pi - \vartheta$. In entrambi i casi si ha

$$\|(Q - P) \times A\| = \|(Q - P)\| \, \|A\| \sin(\vartheta).$$

Inoltre, essendo HQP un triangolo rettangolo, vale la seguente relazione

$$\|(P - H)\| = \|P - Q\| \sin(\vartheta).$$

Dunque

$$d(P, r) = \|P - H\| = \|P - Q\| \sin(\vartheta) = \frac{\|(P - Q) \times A\|}{\|A\|}.$$

$\square$

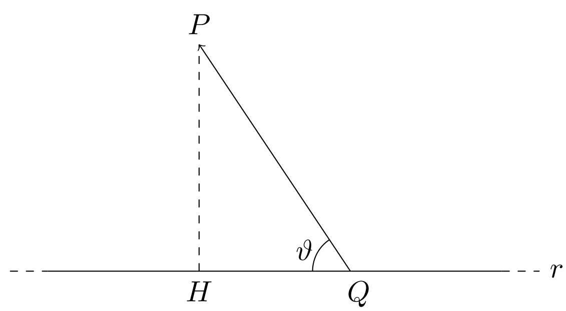


Figura 2.38.1: Situazione descritta nella dimostrazione della Proposizione 2.38.

Definizione 2.39. *Siano P un punto e π un piano in $\mathbb{R}^3$. La proiezione ortogonale di P su π è il punto H che si ottiene intersecando il piano π con l'unica retta passante per P e ortogonale a π.*

Definizione 2.40. *La distanza tra un punto P e un piano π in $\mathbb{R}^3$ è la distanza tra P e la sua proiezione ortogonale H su π, ovvero*

$$d(P, \pi) = \|P - H\|.$$

Anche in questo caso non è necessario determinare esplicitamente le coordinate di H per calcolare la distanza tra un punto e un piano, ma è sufficiente conoscere un punto del piano e un suo vettore normale.

Proposizione 2.41. *Siano Q un punto del piano π e N un suo vettore normale. Allora*

$$d(P, \pi) = \frac{|\langle P - N, N\rangle|}{\|N\|}.$$

Facoltativa. Sia H la proiezione ortogonale di P su π, il triangolo QHP è rettangolo. Sia $\vartheta \in \left(0, \frac{\pi}{2}\right)$ l'ampiezza dell'angolo $Q\widehat{P}H$ (si veda la Figura 2.41.1). Osserviamo che il vettore N ortogonale al piano π è parallelo al vettore $\overline{HP}$. Di conseguenza, l'ampiezza dell'angolo compreso tra N e $\overline{QP}$ è pari a ϑ se i vettori N e $\overline{QP}$ hanno lo stesso verso, mentre è pari a $\pi - \vartheta$ se N e P_H hanno verso opposto. Nel primo caso

$$\langle P - Q, N\rangle = \|P - Q\| \, \|N\| \cos(\vartheta),$$

mentre nel secondo caso

$$\langle P - Q, N\rangle = \|P - Q\| \, \|N\| \cos(\pi - \vartheta) = - \|P - Q\| \, \|N\| \cos(\vartheta).$$

Dunque

$$\left|\langle \overline{QP}, N\rangle\right| = \left\|\overline{QP}\right\| \|N\| \left|\cos(\vartheta)\right| = \left\|\overline{QP}\right\| \|N\| \cos(\vartheta).$$

Nel triangolo rettangolo QHP vale inoltre la seguente relazione

$$\|P - H\| = \|P - Q\| \cos(\vartheta).$$

Pertanto

$$d(P, \pi) = \left\|\overline{HP}\right\| = \frac{|\langle P - Q, N\rangle|}{\|N\|}.$$

□

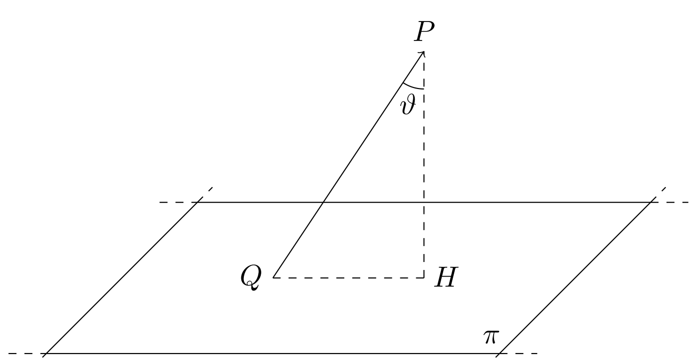


Figura 2.41.1: Situazione descritta nella dimostrazione della Proposizione 2.41.

Per determinare la distanza tra un punto e un piano di cui è nota l'equazione cartesiana, è possibile ricorrere alla seguente formula.

Proposizione 2.42. *Se il piano π ha equazione cartesiana $ax+by+cz+d = 0$ e il punto P ha coordinate $\begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix}$, allora*

$$d(P, \pi) = \frac{|ax_0 + by_0 + cz_0 + d|}{\sqrt{a^2 + b^2 + c^2}}.$$

Facoltativa. È sufficiente determinare le coordinate della proiezione ortogonale H di P su π e calcolare la lunghezza del vettore $P - H$. Il punto H si ottiene intersecando la retta r passante per P e ortogonale al piano π con il piano stesso. Un vettore normale a π è dato ad esempio da $\begin{pmatrix} a \\ b \\ c \end{pmatrix}$, dunque una possibile equazione parametrica per r è

$$\begin{cases} x = x_0 + ta \\ y = y_0 + tb \\ z = z_0 + tc \end{cases}.$$

Mettendo a sistema con l'equazione cartesiana del piano e risolvendo rispetto a t si ottengono le coordinate di H

$$H = \begin{pmatrix} x_0 - a\frac{ax_0+by_0+cz_0+d}{a^2+b^2+c^2} \\ y_0 - b\frac{ax_0+by_0+cz_0+d}{a^2+b^2+c^2} \\ z_0 - c\frac{ax_0+by_0+cz_0+d}{a^2+b^2+c^2} \end{pmatrix}.$$

Ora

$$P - H = \frac{ax_0 + by_0 + cz_0 + d}{a^2 + b^2 + c^2} \begin{pmatrix} a \\ b \\ c \end{pmatrix}.$$

Di conseguenza abbiamo

$$\|P - H\| = \left| \frac{ax_0 + by_0 + cz_0 + d}{a^2 + b^2 + c^2} \right| \sqrt{a^2 + b^2 + c^2} = \frac{|ax_0 + by_0 + cz_0 + d|}{\sqrt{a^2 + b^2 + c^2}}.$$

$\square$

Definizione 2.43. *Siano r_1 e r_2 due rette in $\mathbb{R}^3$. La distanza $d(r_1, r_2)$ tra r_1 e r_2 si definisce come segue:*

a) se r_1 e r_2 sono parallele, allora $d(r_1, r_2)$ è la distanza tra un punto qualsiasi di r_1 e la retta r_2;

b) se r_1 e r_2 non sono parallele, allora $d(r_1, r_2)$ è la distanza tra un punto qualsiasi di r_1 e l'unico piano contenente r_2 e parallelo a r_1.

Dalla definizione precedente si ha che

a) se r_1 e r_2 sono coincidenti, allora $d(r_1, r_2) = 0$;

b) se r_1 e r_2 sono incidenti, allora $d(r_1, r_2) = 0$. Infatti in tal caso il piano contenente r_2 e parallelo a r_1 contiene anche la retta r_1.

Dunque la distanza tra due rette è non nulla solo quando le due rette sono o parallele non coincidenti o sghembe.

Osservazione 2.44. *Se r_1 e r_2 sono due rette sghembe in $\mathbb{R}^3$, allora ogni segmento che unisce due punti qualsiasi delle rette, uno su r_1 e uno su r_2, ha lunghezza maggiore o uguale alla distanza $d(r_1, r_2)$ che si ottiene ricorrendo alla Definizione 2.43. La lunghezza di tale segmento è pari a $d(r_1, r_2)$ solo quando i due punti considerati sono quelli che si ottengono intersecando le rette r_1 e r_2 con l'unica retta perpendicolare e incidente ad entrambe. Per questo motivo $d(r_1, r_2)$ è anche detta* minima distanza tra due rette sghembe*.*

Capitolo 3

Spazi Vettoriali

3.1 Spazi vettoriali, sottospazi vettoriali ed esempi

Definizione 3.2. *Uno spazio vettoriale V su un campo $\mathbb{K}$, per noi $\mathbb{K} = \mathbb{R}$ oppure $\mathbb{K} = \mathbb{C}$, è un insieme su cui sono definite due operazioni: una di somma e un prodotto per scalare*

$$V \times V \longrightarrow V \qquad \mathbb{K} \times V \longrightarrow V$$

$$(v, w) \mapsto v + w \qquad (\lambda, v) \mapsto \lambda v,$$

che soddifano alle seguenti proprietà: se $u, v, w \in V$ e $\lambda, \mu \in \mathbb{K}$, allora

a) $u + (v + w) = (u + v) + w$;

b) $v + w = w + w$;

c) $\exists 0_V \in V$ *tale che* $0_V + v = v + 0_V = v$;

d) $\forall v \in V$, $\exists v' \in V$ *tale che* $v + v' = v' + v = 0_V$;

e) $\lambda(v + w) = \lambda v + \lambda w$;

f) $(\lambda + \mu)v = \lambda v + \mu v$;

g) $\lambda(\mu v) = \mu(\lambda v) = (\lambda\mu)v$;

h) $1v = v$, *per ogni* $v \in V$;

Gli elementi di V si chiamano vettori.

Proposizione 3.3. *Sia V uno spazio vettoriale su $\mathbb{K}$. Allora:*

- *siano $v, w, z \in V$. Se $v + w = v + z$ allora $w = z$;*
- *$\exists! 0_V \in V$ tale che per ogni $v \in V$ si ha $v + 0_V = 0_V + v = v$;*
- *sia $0 \in \mathbb{K}$. Allora $0v = 0_V$ per ogni $v \in V$;*
- *$\forall v \in V$ esiste un unico $v' \in V : v + v' = v' + v = 0_V$. Inoltre $v' = (-1)v$;*
- *per ogni $\lambda \in \mathbb{K}$ si ha $\lambda 0_V = 0_V$;*
- *sia $v \neq 0_V$. Allora $\lambda v = 0_V$ se e solamente se $\lambda = 0$.*

Dimostrazione. Se
$$v + w = v + z,$$
sommanda, in entrambi i menbri, l'opposto di v rispetto alla somma si ha
$$w = z.$$
Supponiamo che esistano $0_V, 0'_V$ tale che $x + 0_V = 0_V + x = x + 0'_V = 0'_V + x = x$ per ogni $x \in V$. Allora
$$0_V = 0_V + 0'_V = 0'_V.$$
Sia $0 \in \mathbb{K}$ e sia $v \in V$. Allora
$$0v = (0 + 0)v = 0v + 0v.$$
Sommanda l'opposto de $0v$ in entrambi i membri si ha
$$0v = 0_V.$$
Poiché $v + (-1)v = (1 - 1)v = 0v = 0_V$, $(-1)v$ è un inverso rispetto alla somma. Vediamo l'unicità. Siano $v', v'' \in V$ tale che $v' + v = v'' + v = 0_V$. Allora
$$v' = v' + 0_V = v' + (v + v'') = (v' + v) + v'' = 0_V + v'' = v''.$$
Sia $\lambda \in \mathbb{K}$. Allora
$$\lambda 0_V = \lambda(0_V + 0_V) = \lambda 0_V + \lambda 0_V,$$
ovvero $\lambda 0 = 0_V$. Infine sia $v \neq 0_V$ e sia $\lambda \neq 0$. Supponiamo che $\lambda v = 0$. Moltiplicando per λ^{-1} a destra e sinistra si ha $v = 0_V$. Assurdo. Quindi $\lambda v = 0_V$ se e solamente se $\lambda = 0$. $\square$

Da qui in avanti, denoteremo $v-w := v+(-1)w$. Vediamo alcuni esempi di spazi vettoriali.

Sia

$$\mathbb{R}^n = \left\{ \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \; x_1, \ldots, x_n \in \mathbb{R} \right\},$$

l'insieme delle n-eplu ordinate di numeri reali. Se $X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$, $Y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$, allora $X = Y$ se $x_1 = y_1, \ldots,\ x_n = y_n$. Somma e moltiplicazione per scalare sono definite come segue:

a) $X + Y = \begin{bmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{bmatrix}$;

b) $\lambda X = \begin{bmatrix} \lambda x_1 \\ \vdots \\ \lambda x_n \end{bmatrix}$,

Proposizione 3.4. $(\mathbb{R}^n, +, \cdot)$ *è uno spazio vettoriale su* $\mathbb{R}$.

Osserviamo che Il vettore nullo è $\begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$.

Sia

$$\mathbb{C}^n = \left\{ \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \; x_1, \ldots, x_n \in \mathbb{K} \right\},$$

i.e., l'insieme delle n-ple ordinate di numeri complessi. Se $X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$, $Y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \in \mathbb{C}^n$, allora $X = Y$ se $x_1 = y_1, \ldots,\ x_n = y_n$. Possiamo definire una somma ed una moltiplicazione per scalare come segue:

a) $\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{bmatrix}$;

b) $\lambda \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \lambda x_1 \\ \vdots \\ \lambda x_n \end{bmatrix}$.

Proposizione 3.5. $(\mathbb{C}^n, +, \cdot)$ *è uno spazio vettoriale su* $\mathbb{C}$.

Sia $\mathbb{K} = \mathbb{R}$ oppure $\mathbb{K} = \mathbb{C}$. Poniamo

$$\mathbb{K}[x] := \{a_0 + a_1 x + \cdots + a_n x^n : a_0, \ldots, a_n \in \mathbb{K},\ n \in \mathbb{N}\}.$$

Siano $v, w \in \mathbb{K}[x]$. Se $v = a_0 + a_1 x + \cdots + a_n x^n$ e $w = b_0 + b_1 x + \cdots + b_m x^m$, allora diremo che $v = w$ se e solamente se $n = m$ e $a_0 = b_0, \ldots, a_n = b_n$. Somma e moltiplicazione per scalare sono così definite: se $m > n$, allora possiamo scrivere $v = a_0 + a_1 x + \cdots + a_n x^n + 0x^{n+1} + \cdots + 0x^m$; se $m < n$, allora possiamo scrivere $w = b_0 + b_1 x + \cdots + b_m x^m + 0x^{m+1} + \cdots 0x_n$ per cui possiamo assumere che $n = m$ e definiamo

$$v + w := (a_0 + b_0) + \cdots + (a_n + b_n)x^n \qquad \lambda v := \lambda a_0 + \cdots + \lambda a_n x^n.$$

Per esempio:

$$(x+2x^3+x^4+5x^5)+(1-3x+2x^2-x^3+x^4+x^6) = 1-2x+2x^2+x^3+2x^4+5x^5+x^6;$$

$$3(1 - x + 2x^2 - 3x^3) = 3 - 3x + 6x^2 - 9x^3.$$

L'insieme $\mathbb{K}[x]$ con le operazioni appena definite è uno spazio vettoriale su $\mathbb{K}$. Il vettore nullo è il polinomio che ha tutti i coefficienti uguali a zero chiamato il polinomio nullo.

Sia $\mathbb{K}_n[x] = \{a_0 + a_1 x + \cdots + a_n x^n : a_0, \ldots, a_n \in \mathbb{K}\}$ l'insieme dei polinomi di grado minore oppure uguale a n. Le operazione

$$v + w := (a_0 + b_0) + \cdots + (a_n + b_n)x^n \qquad \lambda v := \lambda a_0 + \cdots + \lambda a_n x^n.$$

definiscono una struttura di spazio vettoriale su $\mathbb{K}$.

Se X un insieme e sia $V = \{f : X \longrightarrow \mathbb{K}\}$ l'insieme di tutte le applicazioni di X a valori nel campo $\mathbb{K}$. V ammette una struttura di spazio vettoriale come segue:

$$(f + g)(p) := f(p) + g(p) \qquad (\lambda f)(p) := \lambda f(p).$$

Il vettore nullo è l'applicazione che associa ad ogni elemento di X l'elemento $0 \in \mathbb{K}$, ovvero l'applicazione che vale costantemente $0 \in \mathbb{K}$.

Definizione 3.6. *Un sottoinsieme $W \subseteq V$ non vuoto si dice un sottospazio vettoriale di V se*

- *per ogni coppia di vettori $w, z \in W$ si ha $w + z \in W$ (chiuso rispetto alla somma);*
- *per ogni $w \in W$ e per ogni $\lambda \in \mathbb{K}$, si ha $\lambda w \in W$ (chiuso rispetto alla moltiplicazione per scalare).*

Osservazione 3.7. *Se W è un sottospazio vettoriale di V, allora $0_V \in W$.*

Esistono sottoinsiemi di uno spazio vettoriale che verificano la prima condizione ma non la seconda e viceversa. Per esempio, sia

$$W_1 = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} : x \geq 0,\, y \geq 0 \right\}.$$

W_1 verifica la prima proprietà ma non è chiuso rispetto alla moltiplicazione per scalara.

Sia

$$W_2 = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} : x = 0 \,\text{oppure}\, y = 0 \right\}.$$

W_2 è chiuso rispetto la moltiplicazione per scalare ma non è chiuso rispetto alla somma. Quindi affinché un sottoinsieme sia un sottospazio vettoriale deve essere chiuso rispetto alla somma e rispetto alla moltiplicazione per scalare. La prossima proposizione unifica le due condizioni ed è lasciata per esercizio.

Proposizione 3.8. *Sia $W \subseteq V$. W è un sottospazio vettoriale se e solamente se per ogni $w_1, w_2 \in W$ e per ogni $\alpha_1, \alpha_2 \in \mathbb{K}$, allora $\alpha_1 w_1 + \alpha_2 w_2 \in W$.*

Esempio 3.9.

a) sia $S \subseteq \mathbb{R}^3$. Allora $S^{\perp}$ è un sottospazio vettoriale di $\mathbb{R}^3$;

b) $W = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^3 : x + y = 0 \right\}$ è un sottospazio vettoriale di $\mathbb{R}^3$;

c) $W = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \in \mathbb{R}^3 : x + y + z = 1 \right\}$ non è un sottospazio vettoriale di $\mathbb{R}^3$;

d) *sia* $\mathbb{K}[x]$ *l'insieme dei polinomi a coefficienti in* $\mathbb{K}$ *e sia* $p = a_0 + a_1 x + \cdots + a_n x^n \in \mathbb{K}[x]$. *Possiamo pensare a* p *come una funzione come segue:*

$$p : \mathbb{K} \longrightarrow \mathbb{K} \qquad \alpha \mapsto p(\alpha) = a_0 + a_1\alpha + \cdots + a_n\alpha^n.$$

È facile verificare che $(p+q)(\alpha) = p(\alpha) + q(\alpha)$ *ed* $(\lambda p)(\alpha) = \lambda p(\alpha)$.

Se $\beta \in \mathbb{K}$, *allora* $W = \{p \in \mathbb{K}[x] : p(\beta) = 0\}$ *è un sottospazio vettoriale di* $\mathbb{K}[x]$*;*

e) *sia* $V = \{f : \mathbb{R} \longrightarrow \mathbb{R}$. *Allora* $W = \{f \in V : f(1) = 0\}$ *è un sottospazio vettoriale di* V*;*

Definizione 3.10. *Siano* $v_1, \ldots, v_s \in V$. *Si dice combinazione lineare di* $v_1, \ldots, v_s \in V$ *ogni elemento* $w \in V$ *esprimibile nella forma*

$$w = \lambda_1 v_1 + \cdots + \lambda_s v_s.$$

I numeri $\lambda_1, \ldots, \lambda_s$ *si dicono i coefficienti della combinazione lineare. L'insieme di tutte le combinazioni lineari di* $v_1, \ldots, v_s$ *verrà indicato con*

$$\mathcal{L}(v_1, \ldots, v_s) := \{\lambda_1 v_1 + \cdots + \lambda_s v_s : \lambda_1, \ldots, \lambda_s \in \mathbb{K}\}.$$

$\mathcal{L}(v_1, \ldots, v_s)$ *si chiama lo* spazio generato *da* $v_1, \ldots, v_s$.

Esempio 3.11. *In* $\mathbb{R}^3$ *si consideri una retta passante per l'origine:*

$$r : X = tA, \ t \in \mathbb{R}, \ A \neq 0.$$

Allora $r = \mathcal{L}(A)$.

Un piano passante per l'origine ha equazioni parametriche:

$$\pi : X = tv + sw, s, t \in \mathbb{R}, \ v \times w \neq 0,$$

ovvero $\pi = \mathcal{L}(v, w)$.

Proposizione 3.12. *Siano* $v_1, \ldots, v_s \in V$. *Allora* $\mathcal{L}(v_1, \ldots, v_s)$ *è un sottospazio vettoriale di* V.

Dimostrazione. Dobbiamo dimostrare che se $v, w \in \mathcal{L}(v_1, \ldots, v_s)$ e $\lambda \in \mathbb{K}$, allora $v + w \in \mathcal{L}(v_1, \ldots, v_s)$ e $\lambda v \in \mathcal{L}(v_1, \ldots, v_s)$.

Se $v, w \in \mathcal{L}(v_1, \ldots, v_s)$, allora esistono $\alpha_1, \ldots \alpha_s$, rispettivamente $\beta_1, \ldots, \beta_s$, tali che $v = \alpha_1 v_1 + \ldots + \alpha_s v_s$, rispettivamente $w = \beta_1 v_1 + \cdots + \beta_s v_s$. Allora

$$\begin{aligned} v + w &= \alpha_1 v_1 + \cdots + \alpha_s v_s + \beta_1 v_1 + \cdots + \beta_s v_s \\ &= (\alpha_1 + \beta_1) v_1 + \cdots + (\alpha_s + \beta_s) v_s \in \mathcal{L}(v_1, \ldots, v_s) \end{aligned}$$

e

$$\lambda v = \lambda\alpha_1 v_1 + \cdots + \lambda\alpha_s v_s \in \mathcal{L}(v_1, \ldots, v_s).$$

□

Le rette ed i piani passanti per l'origine sono sottospazi vettoriale di $\mathbb{R}^3$, vedi esempio 3.11, ed è facile convincersi che sono tutti i sottospazi vettoriali di $\mathbb{R}^3$ oltre a $\{0_{\mathbb{R}^3}\}$ e $\mathbb{R}^3$ stesso.

Proposizione 3.13. *Sia W un sottospazio vettoriale di V e siano $w_1, \ldots w_s \in W$. Allora $\mathcal{L}(w_1, \ldots, w_s) \subseteq W$.*

Dimostrazione. Siano $\alpha_1, \ldots, \alpha_k \in \mathbb{K}$. I vettori $\alpha_1 w_1, \cdots, \alpha_k w_k \in W$, poiché W è chiuso rispetto alla moltiplicazione per scalare. W è anche chiuso rispetto alla somma per cui

$$\alpha_1 w_1 + \cdots + \alpha_k w_k \in W.$$

Questo significa che ogni combinazione lineare di $w_1, \ldots, w_k$ appartiene ancora a W per cui

$$\mathcal{L}(w_1, \ldots, w_s) = \{\alpha_1 w_1 + \cdots + \alpha_k w_k : \alpha_1, \ldots, \alpha_k \in \mathbb{K}\} \subseteq W.$$

□

Osservazione 3.14. *Sia $w, w_1, \ldots, w_k \in V$. Il vettore w è combinazione lineare dei vettori $w_1, \ldots, w_k$ se e solamente se $w \in \mathcal{L}(w_1, \ldots, w_k)$.*

Definizione 3.15. *Siano $v_1, \ldots, v_s \in V$. Diremo che:*

a) $v_1, \ldots, v_s$ *sono* linearmente dipendenti *se esistono $\alpha_1, \ldots, \alpha_s$ non tutti nulli tali che*
$$\alpha_1 v_1 + \cdots + \alpha_s v_s = 0_V;$$

b) $v_1, \ldots, v_s$ linearmente indipendenti *se non sono linearmente dipendenti, ovvero se*
$$\alpha_1 v_1 + \cdots + \alpha_s v_s = 0_V,$$
allora necessariamente $\alpha_1 = \cdots = \alpha_s = 0$;

c) $v_1, \ldots, v_s$ *formano un* sistema di generatori *se $\mathcal{L}(v_1, \ldots, v_s) = V$. Se V è generato da un numero finito di elementi, diremo che V è* finitamente generato.

Esempio 3.16. *In* $\mathbb{R}^4$ *si considerino i vettori* $\begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -3 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ -1 \\ 1 \end{bmatrix}$.
I vettori sono linearmente dipendenti se esistono $\alpha_1, \ldots, \alpha_4 \in \mathbb{R}$ *non tutti nulli tali che*

$$\alpha_1 \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} + \alpha_2 \begin{bmatrix} 1 \\ 1 \\ -1 \\ 0 \end{bmatrix} + \alpha_3 \begin{bmatrix} 0 \\ 1 \\ -3 \\ 1 \end{bmatrix} + \alpha_4 \begin{bmatrix} 2 \\ 1 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix},$$

per cui se e solamente se esistono $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ *non tutti nulli tali che*

$$\begin{bmatrix} \alpha_1 + \alpha_2 + 2\alpha_4 \\ \alpha_2 + \alpha_3 + \alpha_4 \\ -\alpha_2 - 3\alpha_3 - \alpha_4 \\ \alpha_1 + \alpha_3 + \alpha_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

Quindi, se il sistema lineare

$$\begin{cases} \alpha_1 + \alpha_2 + 2\alpha_4 = 0 \\ \alpha_2 + \alpha_3 + \alpha_4 = 0 \\ -\alpha_2 - 3\alpha_3 - \alpha_4 = 0 \\ \alpha_1 + \alpha_3 + \alpha_4 = 0 \end{cases}.$$

ammette solzioni "non banali", ovvero $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ *non tutti nulli che verificano le equazioni anteriori, allora i vettori sono linearmente dipendenti. Altrimenti i vettori sono linearmente indipendenti.*

Esempio 3.17. *Siano* $e_1 = \begin{bmatrix} 1 \\ \vdots \\ 0 \end{bmatrix}, \ldots, e_n = \begin{bmatrix} 0 \\ \vdots \\ 1 \end{bmatrix} \in \mathbb{K}^n$. *Ogni vettore* $X \in \mathbb{K}^n$ *è combinazione lineare di* $e_1, \ldots, e_n$. *Infatti*

$$\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1 e_1 + \cdots + x_n e_n.$$

Quindi $\mathcal{L}(e_1, \ldots, e_n) = \mathbb{K}^n$. *Si osservi inoltre che i vettori* $e_1, \ldots, e_n$ *sono linearmente indipendenti.*

Esempio 3.18. *Siano* $1, 1+x, 1-x+x^2 \in \mathbb{R}_2[x]$. *Il vettore* $1+x+3x^2 \in \mathcal{L}(1, 1+x, 1-x+x^2)$ *se esistono* $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{R}$ *tali che*

$$\alpha_1 1 + \alpha_2(1+x) + \alpha_3(1-x+x^2) = 1+x+3x^2,$$

quindi se e solamente se

$$(\alpha_1+\alpha_2+\alpha_3) + (\alpha_2-\alpha_3)x + \alpha_3 x^2 = 1+x+3x^2.$$

Quindi, se il sistema lineare

$$\begin{cases} \alpha_1+\alpha_2+\alpha_3 = 1 \\ \alpha_2 - \alpha_3 = 1 \\ \alpha_3 = 3 \end{cases},$$

di tre equazioni in tre incognite, ammette soluzioni allora il vettore $1+x+3x^2 \in \mathcal{L}(1, 1+x, 1-x+x^2)$, *altrimenti il vettore non appartiene al sottospazio* $\mathcal{L}(1, 1+x, 1-x+x^2)$.

Esempio 3.19. *Sia* $1, x, \ldots, x^n \in \mathbb{K}_n[x]$. *È facile verificare che* $\mathcal{L}(1, x, \ldots, x^n) = \mathbb{K}_n[x]$. *Inoltre se*

$$\alpha_0 1 + \alpha_1 x + \cdots + \alpha_n x^n = 0_{\mathbb{K}_n[x]},$$

allora $\alpha_0 = \cdots = \alpha_n = 0$.

Osservazione 3.20. *Sia* V *uno spazio vettoriale su* $\mathbb{K}$. *Allora*

- $v \in V$ *è linearmente indipendente se e solamente se* $v \neq 0_V$*;*
- *siano* $v_1, \ldots, v_n \in V$. *I vettori* $0_V, v_1, \ldots, v_n$ *sono linearmente dipendenti.*
- $v, w \in V$ *sono linearmenti dipendenti se e solamente se* $v = \lambda w$ *oppure* $w = \lambda v$ *per un certo* $\lambda \in \mathbb{K}$*;*
- *sia* $\mathcal{B} = \{v_1, \ldots, v_n\}$. *Allora:*
  - *se* $v_1, \ldots v_n$ *sono linearmente indipendenti, allora ogni sottoinsieme di* $\mathcal{B}$ *è costituito da vettori linearmente indipendenti;*
  - *se* $\mathcal{B}$ *è un insieme formato da vettori linearmente dipendenti, allora ogni sovrainsieme di* $\mathcal{B}$ *è costituito da vettori linearmente dipendenti;*

Capitolo 4

Matrici

4.1 Matrici

Definizione 4.2. *Una matrice, reale o complessa, di formato $m \times n$ è una tabella rettangolare di numeri reali oppure complessi con m righe e n colonne. Quindi mn elementi del tipo:*

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} = (a_{ij})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}},$$

dove $a_{ij} \in \mathbb{R}$ oppure $a_{ij} \in \mathbb{C}$.

Esempio 4.3.

$$A = \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & 3 \end{pmatrix} \in M_{2\times 3}(\mathbb{R})$$

$a_{11} = 1$, $a_{12} = 0$, $a_{13} = 3$, $a_{21} = 0$, $a_{22} = 1$, $a_{23} = 3$

Sia $\mathbb{K} = \mathbb{R}$ oppure $\mathbb{K} = \mathbb{C}$. L'insieme delle matrici di formato $m \times n$ a coefficienti in $\mathbb{K}$ verrà indicato con $M_{m\times n}(\mathbb{K})$.

Siano $A, B \in M_{m\times n}(\mathbb{K})$. Se $A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$ e $B = \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & & \vdots \\ b_{m1} & \cdots & b_{mn} \end{pmatrix}$, diremo che $A = B$ se $a_{ij} = b_{ij}$ per ogni $1 \leq i \leq m$, e $1 \leq j \leq n$.

Sia $A \in M_{m\times n}(\mathbb{K})$. Allora

$$A^1 = \begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix}, \ldots, A^n = \begin{bmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix}$$

sono le colonne di A; analogamente

$$A_1 = [a_{11}, \cdots, a_{1n}], \ \ldots \ , A_m = [a_{m1}, \ldots, a_{mn}],$$

sono le righe di A. Si osservi che $A^j \in \mathbb{K}^m$. Possimo quindi pensare ad una matrice $A \in \mathrm{M}_{m\times n}(\mathbb{K})$ formata da n colonne $A = (A^1, \ldots, A^n)$ oppure formata da m righe $A = \begin{bmatrix} A_1 \\ \vdots \\ A_m \end{bmatrix}$.

Esempio 4.4.

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

$$A^1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, A^2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, A^3 = \begin{bmatrix} 3 \\ 3 \end{bmatrix}.$$

$$A_1 = \begin{bmatrix} 1 & 0 & 3 \end{bmatrix}, A_2 = \begin{bmatrix} 0 & 1 & 3 \end{bmatrix}$$

Se $n = m$, il numero delle righe coincide con il numero delle colonne, allora la matrice A si dirà *matrice quatrata* di ordine n.

Esempio 4.5.

$$A = \begin{pmatrix} 1 & -56 & 3 \\ 5 & 11 & 3 \\ 2 & 1 & -99 \end{pmatrix} \in M_{3\times 3}(\mathbb{R}).$$

4.5.1 Struttura Lineare

Siano $A, B \in M_{m\times n}(\mathbb{K})$ e sia $\lambda \in \mathbb{K}$. Possiamo definire la somma e la moltiplicazione per scalare come segue: se $A = (a_{ij})_{\substack{1 \le i \le m \\ 1 \le j \le n}}$ e $B = (b_{ij})_{\substack{1 \le i \le m \\ 1 \le j \le n}}$, definiamo:

$$A + B := (a_{ij} + b_{ij})_{\substack{1 \le i \le m \\ 1 \le j \le n}}$$

$$\lambda A := (\lambda a_{ij})_{\substack{1 \le i \le m \\ 1 \le j \le n}}$$

Esempio 4.6. *Siano* $A = \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & 3 \end{pmatrix}, B = \begin{pmatrix} 2 & 1 & 3 \\ 9 & 0 & -3 \end{pmatrix} \in M_{2\times 3}(\mathbb{R})$, *allora*

$$A+B = \begin{pmatrix} 1+2 & 0+1 & 3+3 \\ 0+9 & 1+0 & 3-3 \end{pmatrix} = \begin{pmatrix} 3 & 1 & 6 \\ 9 & 1 & 0 \end{pmatrix}, \quad -3A = \begin{pmatrix} -3 & 0 & -9 \\ 0 & -3 & -9 \end{pmatrix}.$$

Proposizione 4.7. *$M_{m\times n}(\mathbb{K})$ è uno spazio vettoriale su $\mathbb{K}$.*

Il vettore nullo è la matrice $0_{M_{m\times n}(\mathbb{K})} = \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{bmatrix}$ chiamata anche matrice nulla.

Sia $A \in M_{m\times n}(\mathbb{K})$. La trasposta di $A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$ è una matrice $A^T \in M_{n\times m}(\mathbb{K})$ così definita:

$$A^T = (a_{ji})_{\substack{1 \le j \le n \\ 1 \le i \le m}} .$$

La matrice A^T si ottiene a partire da A scambiando le righe con le colonne.

Esempio 4.8. *Sia $A = \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & 3 \end{pmatrix}$. Poiché $A^1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $A^2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, $A^3 = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$, si ha*

$$A^T = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 3 & 3 \end{pmatrix}.$$

Proposizione 4.9. *Siano $A, B \in M_{m\times n}(\mathbb{K})$ e sia $\lambda \in \mathbb{K}$. Allora*

a) $(A+B)^T = A^T + B^T$ *;*

b) $(\lambda A)^T = \lambda A^T$*;*

c) $(A^T)^T = A$.

Se $A = (a_{ij})_{\substack{1 \le i \le m \\ 1 \le j \le n}} \in M_{m\times n}(\mathbb{C})$, definiamo:

$$\overline{A} = (\overline{a}_{ij})_{\substack{1 \le i \le m \\ 1 \le j \le n}} ,$$

$$A^* = (\overline{a}_{ji})_{\substack{1 \le i \le m \\ 1 \le j \le n}} = (\overline{A})^T = \overline{(A^T)}$$

La matrice A^* si chiama *l'aggiunta* di A.

Esempio 4.10. *Sia* $A = \begin{pmatrix} 1+i & i & 3-i \\ i & 1 & i \end{pmatrix}$. *Allora*

$$\overline{A} = \begin{pmatrix} 1-i & -i & 3+i \\ -i & 1 & -i \end{pmatrix}.$$

L'aggiunto di A è data da

$$A^* = \begin{pmatrix} 1-i & -i \\ -i & 1 \\ 3+i & -i \end{pmatrix}.$$

Proposizione 4.11. *Siano* $A, B \in M_{m\times n}(\mathbb{C})$ *e* $\lambda \in \mathbb{C}$. *Allora*

a) $\overline{(A+B)} = \overline{A} + \overline{B}$;

b) $\overline{(\lambda A)} = \overline{\lambda A}$;

c) $\overline{\overline{A}} = A$;

d) $(A+B)^* = A^* + B^*$;

e) $(\lambda A)^* = \overline{\lambda} A^*$;

f) $(A^*)^* = A$;

Da qui in avanti considereremo solamente matrici quadrate.

Sia $A = (a_{ij})_{\substack{1 \le i \le n \\ 1 \le j \le n}} \in M_{n\times n}(\mathbb{K})$. Gli elementi $a_{11}, \cdots, a_{nn}$ si dicono elementi sulla *diagonale principale*.

$$\begin{pmatrix} a_{11} & * & \cdots & * \\ \vdots & a_{22} & & \vdots \\ \vdots & & \ddots & \vdots \\ * & \cdots & \cdots & a_{nn} \end{pmatrix}$$

Una matrice A si dice *diagonale* se tutti i coefficienti al di fuori dalla diagonale principale sono nulli: se $A = (a_{ij})_{\substack{1 \le i \le n \\ 1 \le j \le n}}$ è diagonale allora $a_{ij} = 0$ se $i \neq j$.

$$\begin{pmatrix} * & 0 & \cdots & 0 \\ 0 & * & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & * \end{pmatrix}$$

La matrice nulla è una matrice diagonale. La *matrice identità di ordine* n, che indicheremo con Id_n oppure Id ove fosse chiaro il formato, è la matrice diagonale che ha tutti gli elementi uguali ad 1 sulla diagonale principale:

$$\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & 1 \end{pmatrix}$$

La somma di due matrici diagonali è ancora una matrice diagonale per cui l'insieme delle matrici diagonali è chiuso rispetto alla somma. Infatti,

$$\begin{pmatrix} * & 0 & \cdots & 0 \\ 0 & * & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & * \end{pmatrix} + \begin{pmatrix} * & 0 & \cdots & 0 \\ 0 & * & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & * \end{pmatrix} = \begin{pmatrix} * & 0 & \cdots & 0 \\ 0 & * & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & * \end{pmatrix}.$$

Analogomente se moltiplichiamo una matrice diagonale per uno scalare il risultato è ancora una matrice diagonale, quindi l'insieme delle matrici diagonali è chiuso rispetto alla moltipliazione per scalare. Riassumendo, l'insieme delle matrici diagonali è un sottospazio vettoriale di $M_{n\times n}(\mathbb{K})$.

Una matrice A si dice *triangolare superiore*, rispettivamente *triangolare inferiore*, se tutti gli elementi sotto la diagonale principale sono nulli, rispettivamente sopra la diagonale principale sono nulli. Quindi, se $A = (a_{ij})_{\substack{1\leq i\leq n\\ 1\leq j\leq n}}$, è una matrice triangolare superiore, allora

$$A = \begin{pmatrix} * & * & \cdots & * \\ 0 & * & \cdots & * \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & * \end{pmatrix}:$$

$a_{ij} = 0$ quando $i > j$. Se $A = (a_{ij})_{\substack{1\leq i\leq n\\ 1\leq j\leq n}}$ è triangolare inferiore, allora

$$A = \begin{pmatrix} * & 0 & \cdots & 0 \\ * & * & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ * & \cdots & * & * \end{pmatrix}:$$

$a_{ij} = 0$ quando $i < j$. L'insieme della matrici triangolari superiori, rispettivamente triangolari inferiori, è chiuso rispetto alla somma ed è chiuso rispetto alla moltiplicazione per scalare $M_{n\times n}(\mathbb{K})$ per cui sono sottospazi vettoriali. Verificheremo che la somma di matrici triangolari superiori è ancora una matrice triangolare superiore e motiplicando per uno scalare una matrice triangolare inferiore si ottiene una matrice triangolare inferiore, lasciando le altre verifiche per esercizio.

$$\begin{pmatrix} * & * & \cdots & * \\ 0 & * & \cdots & * \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & * \end{pmatrix} + \begin{pmatrix} * & * & \cdots & * \\ 0 & * & \cdots & * \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & * \end{pmatrix} = \begin{pmatrix} * & * & \cdots & * \\ 0 & * & \cdots & * \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & * \end{pmatrix}.$$

Sia $\lambda \in \mathbb{K}$. Allora

$$\lambda \begin{pmatrix} * & 0 & \cdots & 0 \\ * & * & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ * & \cdots & * & * \end{pmatrix} = \begin{pmatrix} \lambda * & 0 & \cdots & 0 \\ \lambda * & \lambda * & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ \lambda * & \cdots & \lambda * & \lambda * \end{pmatrix}.$$

Sia $A \in M_{n\times n}(\mathbb{K})$. Allora $A^T \in \mathrm{M}_{n\times n}(\mathbb{K})$. In dettaglio, se

$$A = \begin{pmatrix} {\color{red}*} & {\color{blue}*} & \cdots & {\color{blue}*} \\ * & {\color{red}*} & \cdots & {\color{blue}*} \\ \vdots & & & \vdots \\ \cdot & & & {\color{blue}*} \\ * & \cdots & * & {\color{red}*} \end{pmatrix},$$

allora

$$A^T = \begin{pmatrix} {\color{red}*} & * & \cdots & * \\ {\color{blue}*} & {\color{red}*} & \cdots & * \\ \vdots & & & \vdots \\ \cdot & & & * \\ {\color{blue}*} & \cdots & {\color{blue}*} & {\color{red}*} \end{pmatrix}.$$

Quindi la diagonale principale di A coincide con la diagonale principale di A^T.

Sia $A \in \mathrm{M}_{n\times n}(\mathbb{R})$. Diremo che A è *simmetrica* se $A = A^T$. Diremo che A è *antisimmetrica* se $A = -A^T$. Se $A = (a_{ij})_{\substack{1\le i\le n\\ 1\le j\le n}}$, allora A è simmetrica, rispettivamente antisimmetrica, se e solamente se $a_{ij} = a_{ji}$, rispettivamente $a_{ij} = -a_{ji}$, per ogni $1 \le i, j \le n$. Poiché la diagonale principale di A coincide con la diagonale principale di A^T, gli elementi sulla diagonale di una matrice antisimmetrica sono tutti nulli.

Esempio 4.12. *La matrice* $\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$ *è simmetrica mentre* $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ *è antisimmetrica.*

L'insieme delle matrici simmetriche, rispettivamente antisimmetriche, è un sottospazio vettoriale di $M_{n\times n}(\mathbb{R})$. Proveremo, solamente, che la somma di due matrici simmetriche è ancora una matrice simmetrica e che moltiplicando una matrice antisimmetrica per uno scalare ottengo ancora una matrice antisimmetriche, lasciando per esercizio le altre verifiche.

Siano $A, B \in \mathrm{M}_{n\times n}(\mathbb{R})$ tali che $A = A^T$ e $B = B^T$. La tesi è che $(A+B)^T = A+B$. Applicando le proprietà della trasposta si ha

$$(A+B)^T = A^T + B^T = A + B.$$

Sia $A \in \mathrm{M}_{n\times n}(\mathbb{R})$ antisimmetrica e sia $\lambda \in \mathbb{R}$. L'ipotesi è $A = -A^T$; la tesi è $(\lambda A)^T = -\lambda A$.

$$(\lambda A)^T = \lambda A^T = -\lambda A.$$

Diremo che una matrice a coefficienti complessi $A \in M_{n\times n}(\mathbb{C})$ è *Hermitiana* (*auto-aggiunta*), rispettivamente *anti-Hermitiana* (*anti-autoggiunta*), se $A = A^*$, rispettivamente $A = -A^*$. Se $A = (a_{ij})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}}$ è Hermitiana allora $a_{ij} = \overline{a_{ji}}$ per ogni $1 \leq i, j \leq n$, mentre se $A = (a_{ij})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}}$ è anti-Hermitiana, allora $a_{ij} = -\overline{a_{ji}}$ per ogni $1 \leq i, j \leq n$. In particolare se A è Hermitiana, rispettivamente anti-Hermitiana, gli elementi sulla diagonale principale sono numeri reali, rispettivamente immaginari puri.

Esempio 4.13. *La matrice* $\begin{pmatrix} 1 & 1+i \\ 1-i & -4 \end{pmatrix}$ *è Hermitiana mentre la matrice* $\begin{pmatrix} -i & 1-3i \\ -1-3i & 7i \end{pmatrix}$ *è anti-Hermitiana.*

La somma di matrici Hermitiane, rispettivamente anti-Hermitiane, è ancora una matrice Hermitiana, rispettivamente anti-Hermitiana (verificare per esercizio). Tuttavia, l'insieme delle metrici Hermitiane (anti-Hermitiane) non è chiuso rispetto alla moltiplicazione per scalare. Infatti è possibile dimostrare che una matrice A è Hermitiana se e solamente se iA è anti-Hermitiana.

Definizione 4.14. *Sia* $A \in M_{n\times n}(\mathbb{K})$. *La traccia di* A *è la somma degli elementi sulla diagonale principale. Se* $A = (a_{ij})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}}$ *, allora Tr(A)=* $\sum_{i=1}^{n} a_{ii} \in \mathbb{K}$.

Proposizione 4.15. *Siano A, B matrici quadrate di ordine n e sia $\lambda \in \mathbb{K}$. Allora*

a) $\mathrm{Tr}(A+B) = \mathrm{Tr}(A) + \mathrm{Tr}(B)$*;*

b) $\mathrm{Tr}(\lambda A) = \lambda \mathrm{Tr}(A)$*;*

c) $\mathrm{Tr}(A) = \mathrm{Tr}(A^T)$*;*

d) se $A \in M_{n\times n}(\mathbb{C})$*, allora* $\mathrm{Tr}(A^*) = \overline{\mathrm{Tr}(A)}$*.*

4.15.1 Prodotto di matrici

Sia $\mathbb{K} = \mathbb{R}$ oppure $\mathbb{C}$. Date le matrici $A \in M_{m\times p}(\mathbb{K})$ e $B \in M_{p\times n}(\mathbb{K})$, dove il numero delle colonne di A è uguale al numero delle righe di B, il prodotto $AB \in M_{m\times n}(\mathbb{K})$, chiamato prodotto righe per colonna è così definito: se

$$A = (a_{ij})_{\substack{1 \leq j \leq m \\ 1 \leq j \leq p}}, B = (b_{ij})_{\substack{1 \leq i \leq p \\ 1 \leq j \leq m}}$$

allora $AB = C = (c_{ij})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} \in M_{m\times n}(\mathbb{K})$, dove

$$c_{ij} = \sum_{m=1}^{p} a_{im} b_{mj}.$$

Esempio 4.16.

$$\begin{pmatrix} 1 & 0 & 2 \\ 0 & 2 & -3 \end{pmatrix} \begin{pmatrix} -1 & 2 \\ 2 & 0 \\ -1 & -4 \end{pmatrix} = \begin{pmatrix} -3 & -6 \\ 7 & 12 \end{pmatrix}.$$

Il prodotto definisce un'applicazione

$$M_{m\times p}(\mathbb{K}) \times M_{p\times n}(\mathbb{K}) \longrightarrow M_{m\times n}(\mathbb{K}) \qquad (A, B) \mapsto AB$$

Può succedere che la matrice AB sia definita mentre BA non sia definita. Per esempio se $A \in M_{3\times 4}(\mathbb{R})$ e $B \in M_{4\times 2}(\mathbb{R})$, allora AB è definita mentre BA no. Se $A \in M_{n\times m}(\mathbb{K})$ e $B \in M_{m\times n}(\mathbb{K})$, con $n \neq m$, allora le matrici AB e BA non sono confrontabili. La domanda se il produtto di due matrici è commutativo ha senso solo se consideriamo matrici quadrate dello stesso ordine. Il prossimo esempio dimostra che il prodotto di matrici non è commutativo per cui, generalmente, $AB \neq BA$. È anche possibile che $AB = 0_{M_{n\times n}(\mathbb{K})}$ benché le matrici A e B non siano la matrice nulla.

Esempio 4.17. *Siano* $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, $C = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$. *Allora:*

- $AC = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$;
- $CA = A$;
- $AA = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$;
- $CC = C$;

Sia $A = (a_{ij})_{\substack{1 \le i \le n \\ 1 \le j \le n}}$ e sia $p \in \mathbb{N}$. Definiamo $A^p = \underbrace{A \cdots A}_{p}$ se $p > 0$. Il prodotto di matrici gode delle seguenti proprietà.

Proposizione 4.18. *Sia* $A \in M_{m \times n}(\mathbb{K})$, $B, C \in M_{n \times p}(\mathbb{K})$, $D \in M_{p \times q}(\mathbb{K})$ *e* $\lambda \in \mathbb{K}$. *Allora*

a) $A \,\mathrm{Id}_n = A$ *e* $\mathrm{Id}_n B = B$;

b) $A(BD) = (AB)D$;

c) $A(B + C) = AB + AC$;

d) $(B + C)D = BD + CD$;

e) $A(\lambda B) = (\lambda A)B = \lambda(AB)$;

f) $(AB)^T = B^T A^T$;

g) *se* $\mathbb{K} = \mathbb{C}$, *allora* $(AB)^* = B^* A^*$.

Dimostrazione. Dimostriamo solamente che $(AB)^T = B^T A^T$. Poiché $B^T \in M_{p \times n}(\mathbb{K})$ e $A^T \in M_{n \times m}(\mathbb{K})$ il prodotto $B^T A^T$ è ben definito ed il risultato è una matrice di formato $p \times m$ come $(AB)^T$. Poniamo $B^T = ({b^T}_{ij})_{\substack{1 \le i \le p \\ 1 \le j \le n}}$ e $A^T = ({a^T}_{ij})_{\substack{1 \le i \le n \\ 1 \le j \le m}}$. Sia $C = (AB)^T$. Poiché

$$(AB)_{ij} = \sum_{l=1}^{n} a_{il} b_{lj},$$

dove $(AB)_{ij}$ è l'elemento che si trova sulla i-esima riga e sulla j-esima colonna della matrice AB, si ha

$$c_{ij} = (AB)_{ji} = \sum_{l=1}^{n} a_{jl} b_{li} = \sum_{l=1}^{n} b^T_{il} a^T_{lj} = (B^T A^T)_{ij},$$

dove $(B^TA^T)_{ij}$ è l'elemento che si trova sulla j riga e sulla i colonna della matrice B^TA^T, per cui $(AB)^T = B^TA^T$. $\square$

Osservazione 4.19. *Molti studenti confondono la proprietà (f) con l'uguaglianza $(AB)^T = A^TB^T$. Vorrei porre all'attenzione che oltre ad essere sbagliata, la matrice A^TB^T non è sempre definita. Infatti, se $A \in M_{4\times 2}(\mathbb{K})$ e $B \in M_{2\times 3}(\mathbb{K})$, allora A^TB^T non è definita poiché il numero di colonne della matrice A^T è 2 mentre il numero delle righe della matrice B^T è 3.*

Osservazione 4.20. *Sia $A \in M_{m\times p}(\mathbb{K})$ e sia $B \in M_{p\times n}(\mathbb{K})$. Allora $AB \in M_{m\times n}(\mathbb{K})$. Sia $1 \leq k \leq n$. Allora*

$$(AB)^k = AB^k.$$

Infatti

$$(AB)^k = \begin{bmatrix} \sum_{l=1}^p a_{1l}b_{lk} \\ \vdots \\ \sum_{l=1}^p a_{il}b_{lk} \\ \vdots \\ \sum_{l=1}^p a_{ml}b_{lk} \end{bmatrix} = \begin{bmatrix} a_{11} & \cdots & \cdots & a_{1p} \\ \vdots & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ a_{m1} & \cdots & \cdots & a_{mp} \end{bmatrix} \begin{bmatrix} b_{1k} \\ \vdots \\ \vdots \\ b_{pk} \end{bmatrix} = AB^k$$

Analogamente, se $1 \leq k \leq m$, si ha

$$(AB)_k = A_kB.$$

Infatti

$$\begin{aligned} (AB)_k &= \begin{bmatrix} \sum_{l=1}^p a_{kl}b_{l1} & \cdots & \sum_{l=1}^p a_{kl}b_{li} & \cdots & \sum_{l=1}^p a_{kl}b_{ln} \end{bmatrix} \\ &= \begin{bmatrix} a_{k1} & \cdots & \cdots & a_{kp} \end{bmatrix} \begin{bmatrix} b_{11} & \cdots & \cdots & b_{1n} \\ \vdots & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ b_{p1} & \cdots & \cdots & b_{pn} \end{bmatrix} = A_kB \end{aligned}$$

Siano $e_1 = \begin{bmatrix} 1 \\ \vdots \\ 0 \end{bmatrix}, \ldots, e_n = \begin{bmatrix} 0 \\ \vdots \\ 1 \end{bmatrix} \in \mathbb{K}^n$ e sia $A \in \mathrm{M}_{m\times n}(\mathbb{K})$. Allora

$$Ae_i = \begin{bmatrix} a_{11} & \cdots & a_{1i} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2i} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mi} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} a_{1i} \\ \vdots \\ \vdots \\ \vdots \\ a_{mi} \end{bmatrix} = A^i$$

Analogamente, se $B \in \mathrm{M}_{n\times m}(\mathbb{K})$, allora

$$e_i^T B = B_i,$$

per $i = 1, \dots, m$.

Proposizione 4.21. *Sia $A \in M_{m\times p}(\mathbb{K})$ e $B \in M_{p\times m}(\mathbb{K})$. Allora $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$.*

Dimostrazione.

$$\begin{aligned}\mathrm{Tr}(AB) &= \sum_{k=1}^{m}\sum_{j=1}^{p} a_{kj}b_{jk} = \sum_{j=1}^{p}\sum_{k=1}^{m} b_{kj}a_{jk} \\ &= \mathrm{Tr}(BA).\end{aligned}$$

□

4.22 Matrici invertibili

In questa sezione tratteremo solamente matrici quadrate.

Definizione 4.23. *Una matrice A quadrata di ordine n si dice invertibile se esiste un matrice quadrata B di ordine n tale che $AB = BA = \mathrm{Id}_n$.*

Proposizione 4.24.

a) esistono matrici diverse da zero che non sono invertibili;

b) se A è invertibile, allora esiste una unica B tale che $AB = BA = \mathrm{Id}_n$. B si dice l'inversa di A e si pone $B = A^{-1}$;

c) $(A^{-1})^{-1} = A$

d) se A, B sono matrici invertibili, tali sono AB e BA e $(AB)^{-1} = B^{-1}A^{-1}$, $(BA)^{-1} = A^{-1}B^{-1}$.

e) se A è invertibile, allora A^T è invertibile e la sua iversa è $(A^T)^{-1} = (A^{-1})^T$;

f) se $A \in M_{n\times n}(\mathbb{C})$ è invertibile, allora A^* è invertibile e $(A^*)^{-1} = (A^{-1})^*$.

Dimostrazione. Sia $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Allora $A^2 = 0$. Se esistesse B tale che $AB = \mathrm{Id}_2$, allora

$$0 = A^2 B = A(AB) = A,$$

assurdo.

Sia $A \in M_{n\times n}(\mathbb{K})$ e siano $B, B' \in M_{n\times n}(\mathbb{K})$ tali che

$$AB = BA = AB' = B'A = Id_n.$$

Allora

$$B = BId_n = B(AB') = (BA)B' = B'.$$

Le rimanenti proprietà sono lasciate per esercizio. □

Osservazione 4.25. *Nonostante che il prodotto fra matrici non sia commutativo, si può dimostrare che se A, B sono matrici quadrate di ordine n tali che $AB = \mathrm{Id}_n$, allora A è invertibile e B è l'inversa di A.*

Dalla proposizione anteriore segue che il prodotto di matrici invertibili è ancora una matrice invertibile per cui l'insieme $GL(n, \mathbb{K}) = \{A \in M_{n\times n}(\mathbb{K}) :$ A è invertibile $\}$ è chiuso rispetto al prodotto e l'inverso, i.e., è un *gruppo* non commutativo, ed è chiamto gruppo lineare generale.

Sia A una matrice invertibile e sia $n \in \mathbb{Z}$ negativo. Definiamo: $A^0 = \mathrm{Id}_n$; se $n < 0$, definiamo

$$A^n := (A^{-1})^{-n}.$$

È facile verificare che per ogni $n, m \in \mathbb{Z}$ si ha $A^{n+m} = A^n A^m = A^m A^n$.

Definizione 4.26. *Una matrice $A \in M_{n\times n}(\mathbb{R})$ si dice* ortogonale *se $A^T = A^{-1}$ per cui $AA^T = A^T A = \mathrm{Id}_n$*

Definizione 4.27. *Una matrice $A \in M_{n\times n}(\mathbb{C})$ si dice* unitaria *se $AA^* = A^*A = \mathrm{Id}_n$.*

Si può dimostrare che il prodotto di matrici ortogonali, rispettivamente unitarie, è ancora una matrice ortogonale, rispettivamente una matrice unitaria. Dalla definzione segue direttamente che l'inversa di una matrice ortogonale, rispettivamente unitaria, è ancora una matrice ortogonale, rispettivamente unitaria. Invece la somma di matrici ortogonali, rispettivamente unitarie, non è un generale una matrice ortogonale, rispettivamente unitaria.

4.28 Determinante

Ad ogni matrice quadrata $A \in M_{n\times n}(\mathbb{K})$, dove $\mathbb{K} = \mathbb{R}$ oppure $\mathbb{C}$, possiamo associare un scalare, chiamato il *determinante di A*, definito come segue: se $A = (a) \in M_{1\times 1}(\mathbb{K})$, allora $\det(A) = a$. Supponiamo di averlo definito per matrici di ordine $n-1$. Definiamo

$$\det(A) = \sum_{j=1}^{n} (-1)^{j+1} a_{j1} \det(A_{j1}) \in \mathbb{K},$$

dove A_{j1} è una matrice di ordine $n-1\times n-1$ che si ottiene da A eliminando la $j-$esima riga e la 1 colonna. Se

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

Allora

$$A_{11} = \begin{bmatrix} 3 & 1 \\ 3 & 0 \end{bmatrix},\ A_{21} = \begin{bmatrix} -1 & 4 \\ 3 & 0 \end{bmatrix},\ A_{31} = \begin{bmatrix} -1 & 4 \\ 3 & 1 \end{bmatrix}.$$

Questa formula è detta *lo sviluppo di Laplace secondo la prima colonna.*

Esempio 4.29.

a) Se $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, *allora* $\det A = a_{11}a_{22} - a_{12}a_{21}$*;*

b) se $\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$*, allora*

$$\begin{aligned} \det A = {} & a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} \\ & - a_{21}a_{12}a_{33} + a_{21}a_{13}a_{32} \\ & + a_{31}a_{12}a_{23} - a_{31}a_{13}a_{22}. \end{aligned}$$

c) Se

$$A = \begin{pmatrix} a_{11} & * & \cdots & * \\ 0 & a_{22} & \cdots & * \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & a_{nn} \end{pmatrix},$$

è triangolare superiore, allora $\det A = a_{11} \cdot \cdots \cdot a_{nn}$.

Proprietà del Determinante

1 $\det(A) = \det(A^T)$;

2 $A = (A^1, \dots, A^n)$. Allora $\det(A^1, \dots, \lambda A^i, \dots, A^n) = \lambda \det(A^1, \dots, A^n)$, per ogni $1 \leq i \leq n$ e per ogni $\lambda \in \mathbb{K}$;

3

$$\begin{aligned}\det(A^1, \dots, A^i + B^i, \dots, A^n) &= \det(A^1, \dots, A^i, \dots A^n) \\ &\quad + \det(A^1, \dots, B^i, \dots, A^n),\end{aligned}$$

per ogni $1 \leq i \leq n$, i.e., è additivo su ogni colonna;

4 $\det(A) = 0$ se la matrice A ha due colonne uguali;

5 dalla [2] segue che $\det(A) = 0$ se la matrice A ha una colonna fatta tutta da zeri;

6 [3] e [4] implicano che il determinante di una matrice cambia di segno se si scambiano due colonne;

7 [3] e [2] implicano che il valore del determinante non cambia sommando ad una colonna un multiplo di un altra colonna. Ovvero, se $A = (A^1, \dots, A^n)$ allora per ogni $1 \leq i, j \leq n$, $i \neq j$ e $\lambda \in \mathbb{K}$ si ha

$$\det(A) = \det(A^1, \dots, A^i + \lambda A^j, \dots, A^n);$$

8 $\det(AB) = \det(A)\det(B)$ (Formula di Binet).

Osservazione 4.30.

- *da [1] segue che il determinante di una matrice triangolare inferiore è il prodotto degli elementi sulla diagonale principale,*
- $\det(\mathrm{Id}_n) = 1$*;*
- *si può dimostrare che le proprietà* $[1], \dots, [7]$ *valgano anche per le righe;*
- *da [1] segue che il determinante si può sviluppare rispetto alla 1 riga. Si può dimostrare che il determinante si può sviluppare rispetto alla $k-$esima colonna, i.e.,*

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

oppure rispetto alla $k-$esima riga

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

dove $A_{\alpha\beta}$ è una matrice di ordine $n-1 \times n-1$ che si ottiene da A eliminando la $\alpha-$esima riga e la β-esima colonna.

Definizione 4.31. *Una matrice $A \in \mathrm{M}_{n\times n}(\mathbb{K})$ si dice* singolare*, rispettivamente* non singolare*, se* $\det(A) = 0$*, rispettivamente* $\det(A) \neq 0$*.*

Proposizione 4.32. *Una matrice A è invertibile se e solamente se* $\det(A) \neq 0$ *ovvero se e solamente se A è non singolare. Inoltre* $\det(A^{-1}) = \frac{1}{\det(A)}$.

Dimostrazione. Se A è invertibile, allora esiste A^{-1} tale che $AA^{-1} = \mathrm{Id}_n$. Applicando la formula di Binet si ottiene che

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

da cui segue $\det(A) \neq 0$ e $\det(A^{-1}) = \frac{1}{\det(A)}$. Viceversa, supponiamo che $\det(A) \neq 0$. Definiamo la matrice B di ordine n come segue:

$$b_{ij} = (-1)^{i+j} \det(A_{ji}) / \det(A),$$

dove, nuovamente, A_{ji} è la matrice che ottengo da A eliminando la j-esima riga e le i-esima colonna. Si può dimostrare che $AB = BA = \mathrm{Id}_n$, ovvero $B = A^{-1}$. □

Corollario 4.33. *Se A è una matrice ortogonale, allora* $|\det(A)| = 1$

Dimostrazione. Sia $A \in M_{n\times n}(\mathbb{R})$ una matrice ortogonale. Allora $AA^T = 1$. Applicando il teorema di Binet, otteniamo

$$1 = \det(A)\det(A^T) = [\det(A)]^2,$$

dove l'ultima uguaglianza segue dalla proprietà $\det(A) = \det(A^T)$, concludendo la dimostrazione. □

Corollario 4.34. *Se A è una matrice unitaria, allora* $|\det(A)| = 1$*.*

Dimostrazione.

$$1 = \det(A)\det(A^*) = \det(A)\overline{\det(A)} = |\det(A)|^2$$

□

4.35 Rango di una matrice

Sia $A \in M_{m\times n}(\mathbb{K})$ dove $\mathbb{K} = \mathbb{R}$ oppure $\mathbb{K} = \mathbb{C}$. Un minore di ordine p di A è una matrice di formato $p \times p$ che si ottiene cancellando $m-p$ righe e $n-p$ colonne dalla matrice A. Analogamente, possiamo definire un minore di A selezionando p righe e p colonne. Se

$$A = \begin{bmatrix} 1 & -1 & 4 & -1 \\ 4 & -2 & 3 & 1 \\ 0 & 3 & 0 & 5 \end{bmatrix} \in \mathrm{M}_{3\times 4}(\mathbb{R}).$$

eliminando la 1 riga e la 3 e 4 colonna, oppure selezionando la 2 e 3 riga e la 1 e 2 colonna, otteniamo il minore

$$\begin{bmatrix} 4 & -2 \\ 0 & 3 \end{bmatrix} \in \mathrm{M}_{2\times 2}(\mathbb{R}).$$

Eliminando la 2 colonna, oppure selzionando la 1, 3 e 4 colonna e la 1, 2 e 3 colonna, otteneniamo il minore

$$\begin{bmatrix} 1 & 4 & -1 \\ 4 & 3 & 1 \\ 0 & 0 & 5 \end{bmatrix} \in \mathrm{M}_{3\times 3}(\mathbb{R}).$$

Diremo che il *rango per minori*, che indicheremo con $rg(A)$, è r se:

a) $\exists$ un minore di A di ordine r con determinante $\neq 0$;

b) $r = m$ o $r = n$ oppure tutti i minori di ordine $r+1$ sono singolari;

Dalla definizione di rango per minori segue che $rg(A) \leq \min(m, n)$. Se A è una matrice quadrata, allora vale il seguente risultato.

Proposizione 4.36. *$A \in M_{n\times n}(\mathbb{K})$ è non singolare se e solamente se $rg(A) = n$.*

Sia A' una minore di A ottenuta cancellando $m-p$ righe e $n-p$ colonne dalla matrice A. Se ad A' aggiungiamo un altra riga ed un altra colonna di A diremo che stiamo orlando A'. È possibile dimostrare il seguente fatto.

Teorema 4.37 (orlati)**.** *Il rango per minori della matrice A è uguale ad r se e solamente se esiste una minore M di ordine r non singolare ed $r = \min(m, n)$ oppure tutti minori di A di ordine $r+1$ che contengono M, ovvero tutti i minori di A che ottengo orlando M, sono singolari.*

Come per il determinate, effettuando operazioni sulle righe e/o sulle colonne di una matrice A, il rango non cambia.

Definizione 4.38. *Sia $A \in \mathrm{M}_{m\times n}(\mathbb{K})$. Si chiamano operazione elementari di riga (colonna):*

a) scambiare di posto due righe (colonna);

b) sommare ad una riga (colonna) un multiplo di un altra;

c) moltiplicare una riga (colonna) per uno scalare non nullo.

Proposizione 4.39. *Il rango di una matrice non cambia se si effetuano operazioni elementari di righe, rispettivamente colonne.*

Idea. Sia $A \in \mathrm{M}_{n\times m}(\mathbb{K})$ e sia $\tilde{A}$ la matrice ottenuta attraverso una operazione elementare sulle righe, rispettivamente colonne, di A. Sia M un minore di $\tilde{A}$. Allora M è anche un minore di A oppure è un minore di A sul quale è stato effettuata una operazione elementare di riga, rispettiavemente colonna. Quindi il valore del determinante o rimane inalterato oppure cambia di segno, per cui il rango per minore non cambia. □

Il prossimo risultato caratterizza il rango per minori di una matrice A in termini delle righe, rispettivamente colonne, della matrice A.

Teorema 4.40. *Sia $A \in M_{m\times n}(\mathbb{K})$ dove $\mathbb{K} = \mathbb{R}$ oppure $\mathbb{K} = \mathbb{C}$. Il rango per minori di A coincide con il massimo numero di colonne linearmente indipendenti, rispettivamente massimo numero di righe linearmente indipendenti. In particolare $rg(A) = rg(A^T)$.*

Possiamo, quindi, unificare le nozioni precedenti nella seguente definizione.

Definizione 4.41. *Sia $A \in \mathrm{M}_{m\times n}(\mathbb{K})$. Il rango di A, che indicheremo con $\mathrm{rg}(A)$, è il rango per minori di A, ovvero il massimo numero di colonne linearmente indipendenti, rispettivamente massimo numero di righe linearmente indipendenti*

La dimostrazione del Teorema 4.40 è facoltativa. Prima di dimostrare il Teorema 4.40 proveremo tre Lemmi.

Abbiamo già osservato che un minore M di formato $p \times p$ di A è completamente determinato da p colonne e p righe distinte di A. In dettaglio, selezionando p colonne distinte di A, $A^{j_1}, \ldots, A^{j_p}$, e successimante scegliendo p righe distinte $\alpha_1, \cdots, \alpha_p$ della matrice $(A^{j_1}, \ldots, A^{j_p}) \in \mathrm{M}_{m\times p}(\mathbb{K})$ otteniamo un minore di A di ordine p. Lo stesso minore lo possiamo costruire considerando le p righe $\alpha_1, \ldots, \alpha_p$ di A e poi della matrice $\begin{bmatrix} A_{\alpha_1} \\ \vdots \\ A_{\alpha_p} \end{bmatrix} \in \mathrm{M}_{p\times n}(\mathbb{K})$ selezionado le p colonne distinte $j_1, \ldots, j_p$. Viceversa è facile dimostrare che ogni minore di A di ordine p si ottiene in questo modo. In seguito indicheremo con $A^{j_1,\ldots,j_p}_{\alpha_1,\ldots,\alpha_p}$ il minore di A ottenuto come in precedenza, ovvero considerando le p colonne $A^{j_1}, \ldots, A^{j_p}$ e le p righe $A_{\alpha_1}, \ldots, A_{\alpha_p}$ di A.

Lemma 4.42. *Siano $v_1, \ldots, v_k \in \mathbb{K}^n$ vettori linearmente indipendenti. Siano $\lambda_2, \ldots, \lambda_k \in \mathbb{K}$. Allora i vettori $v_1, v_2 - \lambda_1 v_1, \ldots, v_k - \lambda_k v_1$ sono linearmente indipendenti.*

Dimostrazione. Siano $\alpha_1, \ldots, \alpha_k \in \mathbb{K}$ tale che

$$\alpha_1 v_1 + \alpha_2(v_2 - \lambda_2 v_1) + \cdots + \alpha_k(v_k - \lambda_k v_1) = 0.$$

Vogliamo dimostrare che $\alpha_1 = \cdots = \alpha_k = 0$. Possiamo riscrivere la combinazione lineare come segue:

$$(\alpha_1 - \lambda_2\alpha_2 - \cdots - \lambda_k\alpha_k)v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k = 0.$$

I $v_1, \ldots, v_k$ sono linearmente indipendenti per cui $\alpha_2 = \cdots = \alpha_k = \alpha_1 - \lambda_2\alpha_2 - \cdots - \lambda_n\alpha_n = 0$, ovvero $\alpha_2 = \cdots = \alpha_k = \alpha_1 = 0$. □

Lemma 4.43. *Se $v_1, \ldots, v_k \in \mathbb{K}^n$ sono linearmente dipendenti, allora esiste $j \in \{1, \ldots, k\}$ tale che v_j è combinazione lineare dei rimanenti.*

Dimostrazione. Siano $\alpha_1, \ldots, \alpha_k \in \mathbb{K}$ tale che

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k = 0,$$

con $\alpha_1, \ldots, \alpha_k$ non tutti nulli. Sia $j \in \{1, \ldots, k\}$ tale che $\alpha_j \neq 0$. Allora

$$v_j = \sum_{i=1, i\neq j}^{k} (-\frac{\alpha_i}{\alpha_j}) v_i,$$

da cui la tesi. □

Lemma 4.44. *Siano $v_1, \ldots, v_n \in \mathbb{K}^n$. Sia $A = (v_1, \ldots, v_n) \in \mathrm{M}_{n\times n}(\mathbb{K})$ la matrice le cui colonne sono i vettori $v_1, \ldots, v_n$. I vettori $v_1, \ldots, v_n$ sono linearmente indipendente se e solamente se* $\det(v_1, \ldots, v_n) \neq 0$.

Dimostrazione. Se $\det(v_1, \ldots, v_n) \neq 0$, allora $v_1, \ldots, v_n$ sono linearmente indipendenti. Infatti, se $v_1, \ldots, v_n$ fossero dipendenti, per il Lemma 4.43, allora esisterebbe $1 \leq j \leq n$ tale che $v_j = \sum_{m=1, m\neq j}^{n} \beta_m v_m$. Quindi

$$\det(v_1, \ldots, v_j, \ldots, v_n) = \sum_{m=1, m\neq j}^{n} \beta_m \det \underbrace{(v_1, \ldots, v_m, \ldots v_n)}_{j}.$$

Poiché la matrice $\underbrace{(v_1, \ldots, v_m, \ldots, v_n)}_{j}$ ha due colonne uguali, per $m = 1, \ldots, j-1, j+1, \ldots n$, allora il suo determinante è nullo e quindi anche $\det(v_1, \ldots, v_n)$ sarebbe nullo. Assurdo.

Viceversa, supponiamo che $v_1, \ldots, v_n$ sono linearmente indipendenti. Dimostriamo per induzione su n che $\det(v_1, \ldots, v_n) \neq 0$.

Se $n = 1$, allora $v_1 = (a_1) \neq 0$ e quindi $\det v_1 = a_1 \neq 0$. Supponiamo vero per $n-1$ e dimostriamolo per n. Consideriamo la matrice

$$A = (v_1, \ldots, v_n) = \begin{pmatrix} a_{11} & \cdots & \cdots & a_{1n} \\ a_{21} & \cdots & \cdots & a_{2n} \\ \vdots & \cdots & \cdots & \vdots \\ a_{n1} & \cdots & \cdots & a_{nn} \end{pmatrix}$$

A meno di scambiare due righe della matrice A possiamo suppore che $v_1 = \begin{bmatrix} a_{11} \\ \vdots \\ a_{nn} \end{bmatrix}$ con $a_{11} \neq 0$. Infatti, se scambio due righe il determinante cambia di segno ed è facile verificare che le colonne della nuova matrice continuano ad essere linearmente indipendenti.

Sia $\lambda_j = \frac{a_{j1}}{a_{11}}$, per $2 \leq j \leq n$. Per il lemma 4.44, i vettori $v_1, v_2 - \lambda_2 v_1, \ldots, v_n - \lambda_n v_1$ sono lineramente indipendenti. Inoltre

$$(v_1, v_2 - \lambda_2 v_1, \ldots, v_n - \lambda_n v_1) = \left(\begin{array}{c|ccc} a_{11} & 0 & \cdots & 0 \\ \hline a_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & b_{n2} & \cdots & b_{nn} \end{array}\right)$$

Per le proprietà del determinante si ha

$$\det(v_1, \ldots, v_n) = \det(v_1, v_2 - \lambda_2 v_1, \ldots, v_n - \lambda_n v_1) = a_{11} \det B,$$

dove $B = A_{11}$. L'ultima uguaglianza segue dallo sviluppo di Laplace rispetto alla 1 riga. Poiché i vettori $v_2 - \lambda_2 v_1, \ldots, v_n - \lambda_n v_1$ sono linearmente indipendenti, le colonne dlla matrice B sono linearmente indipendenti. Per ipotesi induttiva $\det B \neq 0$ e quindi $\det(v_1, \ldots, v_n) = a_{11} \det B \neq 0$. $\square$

Dimostrazione del Teorema 4.40. Sia r il rango per minori di A e sia $\mathrm{rg}_c(A)$, rispettivamente $\mathrm{rg}_r(A)$, il massimo numero di colonne linearmente indipendenti, rispettivamente il massimo numero di righe linearmente indipendenti. Per definizione di rango per minori esiste un minore $M = A^{j_1,\ldots,j_r}_{\alpha_1,\ldots,\alpha_r}$ di A di ordine r tale che $\det M \neq 0$. Per il Lemma 4.44 le colonne di M sono linearmente indipendenti. Poiché $\det M = \det M^T$ anche le righe di M sono linearmente indipendenti. Quindi le colonne $A^{j_1}, \ldots A^{j_r}$, rispettivamente le righe $A_{\alpha_1}, \ldots, A_{\alpha_p}$ sono linearmente indipendenti, ovvero $\mathrm{rg}_c(A) \geq r$, rispettivamente $\mathrm{rg}_r(A) \geq r$. Adesso dimostriamo la disuguaglianza inversa, ovvero $\mathrm{rg}_c(A) \leq r$, rispettivamente $\mathrm{rg}_r(A) \leq r$. Dimostreremo solamente $\mathrm{rg}_c(A) \leq r$ poiché l'altra disuguaglianza è identica.

Sia $k = \mathrm{rg}_c(A)$ e siano $A^{j_1}, \ldots, A^{j_k}$ colonne linearmente indipendenti. Sia $B = (A^{j_1}, \ldots, A^{j_k}) \in \mathrm{M}_{m\times k}(\mathbb{K})$. Vogliamo dimostrare che esiste un minore di B di ordine k con determinante differente di zero. La dimostrazione la faremo per induzione su k.

Se $k = 1$, allora è vero. Supponiamo vera per $k-1$ e proviamola per k. A meno di scambiare due righe della matrice B possiamo supporre che $A^{j_1} = \begin{bmatrix} a_{1j_1} \\ \vdots \\ a_{nj_1} \end{bmatrix}$ con $a_{1j_1} \neq 0$. Consideriamo $\lambda_s = \frac{a_{1j_s}}{a_{1j_1}}$, $s = 2, \ldots, k$. Applicando il Lemma 4.44, i vettori

$$A^{j_1}, A^{j_2} - \lambda_2 A^{j_1}, \ldots, A^{j_k} - \lambda_k A^{j_1}$$

sono linearmente indipendenti. Inoltre

$$B' = (A^{j_1}, A^{j_2} - \lambda_2 A^{j_1}, \ldots, A^{j_k} - \lambda_k A^{j_1}) = \left(\begin{array}{c|ccc} a_{1j_1} & 0 & \cdots & 0 \\ \hline a_{2j_2} & b_{2j_2} & \cdots & b_{2j_k} \\ \vdots & \vdots & \vdots & \vdots \\ a_{nj_1} & b_{nj_2} & \cdots & b_{nj_n} \end{array}\right) \tag{4.1}$$

Sia

$$C = B'_{11} = \begin{pmatrix} b_{2j_2} & \cdots & b_{2j_k} \\ \vdots & \ldots & \vdots \\ b_{nj_2} & \cdots & b_{nj_k} \end{pmatrix} \in \mathrm{M}_{(n-1)\times(k-1)}(\mathbb{K}).$$

Le colonne della matrice C sono linearmente indipendenti. Per ipotesi induttiva esiste un minore $M' = (C)^{\beta_1,\ldots,\beta_{k-1}}_{\alpha_1,\ldots,\alpha_{k-1}}$ di C di ordine $k-1$ tale che $\det M' \neq 0$.

Sia $M = A^{j_1, j_{(\beta_1+1)}, \dots, j_{(\beta_{k-1}+1)}}_{1, (\alpha_1+1), \dots, (\alpha_{k-1}+1)}$. La prima colonna di M è

$$M^1 = \begin{bmatrix} a_{1j_1} \\ \vdots \\ a_{(\alpha_{k-1}+1)j_1} \end{bmatrix}$$

con $a_{1j_1} \neq 0$. Sia $\lambda_s = \frac{a_{(\alpha_{s-1}+1)j_{(\beta_{s-1}+1)}}}{a_{1j_1}}$, $s = 2 \dots, k$. Per le proprietà del determinante, e tenendo in mente 4.1 e la definzione del minore M', si ha

$$\begin{aligned} \det(M^1, \dots, M^k) &= \det(M^1, M^2 - \lambda_2 M^1, \dots, M^k - \lambda_k M^1) \\ &= \det \left(\begin{array}{c|ccc} a_{1j_1} & 0 & \cdots & 0 \\ \hline \vdots & & M' & \end{array} \right) \\ &= a_{1j_1} \det M' \neq 0. \end{aligned}$$

□

4.45 Matrice ridotta a scala e algoritmo di Gauss

Nella sezione anteriore abbiamo introdotto il concetto di rango di una matrice ma non abbiamo ancora visto nessuna tecnica per il calcolo esplicito del rango di una matrice. In questo sezione proveremo come l'algoritmo di Gauss possa essere utilizzato per il calcolo del rango di una matrice.

Una matrice di ordine $m \times n$ siffatta

$$\left[\begin{array}{cccccccccccccc} 0 & \cdots & 0 & s_{1j_1} & * & * & * & \cdots & \cdots & * & \cdots & \cdots & \cdots & * \\ 0 & \cdots & \cdots & 0 & \cdots & 0 & s_{2j_2} & * & * & \cdots & \cdots & \cdots & * & * \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & \cdots & s_{3j_3} & * & \cdots & \cdots & \cdots & * \\ \vdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & 0 & \ddots & \cdots & \cdots & \cdots & * \\ \vdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & \ddots & \cdots & * & * \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & s_{rj_r} & \cdots & * \\ \hline 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & \cdots & \cdots & 0 \end{array} \right]$$

si dice una matrice ridotta a scala. I numeri $s_{1j_1}, \dots, s_{r,j_r}$ sono non nulli e si chiamano *perni, pivot, oppure elementi di testa.* Il numero dei perni coincide con il numero di righe differenti di zero.

Esempio 4.46.

$$S = \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 \\ 0 & 0 & 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \in M_{5\times 6}(\mathbb{R})$$

I perni sono: $s_{12} = 1$, $s_{24} = 1$ *ed* $s_{3,5} = 4$.

Proposizione 4.47. *Sia* $S \in M_{m\times n}(\mathbb{K})$ *una matrice ridotta a scala. Allora* $rg(S)$ *è uguale al numero di righe differenti da zero o equivalentemente al numero di perni. Inoltre le colonne corrispondenti ai perni sono vettori linearmente indipendenti.*

Dimostrazione. Una matrice ridotta a scala ha r righe differenti da zero. Quindi $rg(S) \le r$. Sia M il minore di S di ordine r formato dalle colonne $S^{j_1}, \ldots, S^{j_r}$ e dalle righe $S_1, \ldots, S_r$, ovvero

$$M = \begin{bmatrix} s_{1j_1} & * & \cdots & \cdots & * \\ 0 & s_{2j_2} & * & \cdots & * \\ \vdots & 0 & \ddots & & \vdots \\ \vdots & \vdots & \vdots & \ddots & \\ 0 & \cdots & \cdots & \cdots & s_{rj_r} \end{bmatrix}$$

Poiché $\det M = s_{1j_1} \cdots s_{rj_r} \neq 0$ si ha $r \le rg(S)$. Quindi $rg(S) = r$.

Siano $S^{j_1}, \ldots, S^{j_r}$ le colonne corrispondenti ai perni. Sia $A = (S^{j_1}, \ldots, S^{j_r}) \in \mathrm{M}_{m\times r}(\mathbb{K})$

$$\begin{bmatrix} s_{1j_1} & * & \cdots & \cdots & * \\ 0 & s_{2j_2} & * & \cdots & * \\ \vdots & 0 & \ddots & & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & s_{rj_r} \\ 0 & \cdots & \cdots & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & \cdots & \cdots & 0 \end{bmatrix}.$$

Sia, nuovamente, M il minore formato di formato $r \times r$ ottenunto da A eliminando le ultime $m-r$ righe, ovvero

$$M = \begin{bmatrix} s_{1j_1} & * & \cdots & \cdots & * \\ 0 & s_{2j_2} & * & \cdots & * \\ \vdots & 0 & \ddots & & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & s_{rj_r} \end{bmatrix}.$$

Poiché $s_{rj_r}, \ldots, s_{1j_1}$ sono tutti non nulli si ha che $\det M \neq 0$, ovvero $\mathrm{rg}(A) \geq r$. Poiché $A \in \mathrm{M}_{m \times r}(\mathbb{K})$, si ha che $\mathrm{rg}(A) = r$ e quindi $S^{j_1}, \ldots, S^{j_r}$ sono linearmente indipendenti. $\square$

Teorema 4.48. *Ogni matrice A può essere ridotta in forma a scala mediante operazioni elementari di righe.*

Dimostrazione. Sia $A = (A^1, \ldots, A^n)$.
Passo 1
sia $1 \leq j_1 \leq n$ il più piccolo intero affinché $A^{j_1} \neq 0$.
Passo 2
Se a_{1j_1} è nullo, allora scambiamo due righe in modo che l'elemento $a_{1j_1} \neq 0$. Quindi la matrice A, dopo aver effettuato eventualmente una operazione elementare di riga, ha la seguente forma:

$$\begin{bmatrix} 0 & \cdots & 0 & a_{1j_1} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & a_{mj_k} & \cdots \end{bmatrix}$$

Il nostro obbiettivo è quello di passare, attraverso operazioni elementari di righe, ad una matrice i cui elementi sulla colonna A^{j_1} sotto il primo elemento a_{1j_1} siano tutti nulli. Alla k-riga A_k, $k \geq 2$, sottraggo $-a_{kj_1}/a_{1,j_1}A_1$, ottenedo la matrice:

$$\begin{bmatrix} A_1 \\ A_2 - a_{2j_k}/a_{1j_1}A_1 \\ \vdots \\ \vdots \\ A_m - a_{mj_1}/a_{1j_1}A_1 \end{bmatrix} = \left[\begin{array}{cccc|cccc} 0 & \cdots & 0 & a_{1j_1} & * & \cdots & \cdots & * \\ \hline 0 & \cdots & 0 & 0 & a_{2,j_1+1} & \cdots & \cdots & a_{2n} \\ \vdots & & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 & a_{mj_1+1} & \cdots & \cdots & a_{mn} \end{array}\right]$$

Quindi

$$A = \left[\begin{array}{cccc|ccc} 0 & \cdots & 0 & a_{1j_1} & * & \cdots & * \\ \hline \vdots & & \vdots & 0 & & & \\ \vdots & & \vdots & \vdots & & B & \\ \vdots & \cdots & 0 & 0 & & & \end{array}\right]$$

dove $B \in M_{m-1\times(n-(j_1+1))}(\mathbb{R})$. Se B è la matrice nulla oppure $A \in M_{1\times n}(\mathbb{R})$, allora ho finito. Altrimenti ripeto lo stesso procedimento per la matrice B. Dopo un numero finito di passi, arriviamo ad una matrice le cui ultime righe sono nulle; oppure in cui l'ultimo elemento di testa appartiene all'ultima riga. In entrambi i casi abbiamo ridotto a scala la matrice di partenza A. $\square$

Il metodo anteriore è chiamato *metodo di Gauss* oppure *metodo di eliminazione di Gauss* e permette di ridurre una matrice a scala attraverso operazioni elementari di riga.

Corollario 4.49. *Sia A una matrice ed S una sua riduzione a scala. Allora $rg(A) = rg(S)$ per cui il rango di una matrice è uguale al numero di elementi di testa di una sua riduzione a scala; uguale al numero di righe non nulle di una sua riduzione a scala.*

Dimostrazione. Poiché S è ottenuta da A attraverso operazioni elementari di righe si ha $\mathrm{rg}(A) = \mathrm{rg}(S)$. $\square$

4.50 Sistemi lineari

Studiando i sistemi lineari, i problemi principali che vogliamo risolvere sono: quando un sistema ammette soluzioni; se le ammette, quante sono; come si trovano. In questa sezione con $\mathbb{K}$ intendiamo $\mathbb{R}$ oppure $\mathbb{C}$. La forma generale di un sistema di m equazioni in n-incognite:

$$\left\{\begin{array}{rcl} a_{11}x_1 + \cdots + a_{1n}x_n & = & b_1 \\ \vdots & & \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n & = & b_m \end{array}\right. \tag{4.2}$$

I termini $b_1, \ldots b_m$ sono i termini noti, $(a_{ij})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$ i coefficienti del sistema, $x_1, \ldots x_n$ le incognite, o variabili, del sistema lineare. Se tutti i termini noti sono uguali a zero, il sistema si dice *omogeneo*.

Definizione 4.51. *Una soluzione del sistema (4.2) è una $n-$pla di numeri $v_1, \ldots v_n$ che sostituiti ordinatamente alle incognite $x_1, \ldots, x_n$ soddisfano le equazioni del sistema.*

Scriviamo un sistema lineare in forma matriciale. Sia

$$A = (a_{ij})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}, \qquad b = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}, \qquad X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

Allora un sistema lineare ha la forma

$$AX = b, \tag{4.3}$$

dove $A \in M_{m \times n}(\mathbb{K})$ è chiamata *matrice incompleta* oppure *matrice dei coefficienti*, b vettore dei termini noti ed infine X vettore delle incognite. La matrice $(A|b) \in M_{m \times (n+1)}(\mathbb{K})$ che si ottiene aggiungendo ad A il vettore dei termini noti, si chiama la *matrice completa*. In questo linguaggio $Sol(A|b) = \{Y \in \mathbb{K}^n : \ AY = b\}$ è l'insieme delle soluzione del sistema lineare $AX = b$.

Un sistema lineare $AX = b$ si dice *compatibile* oppure *risolubile* se $Sol(A|b) \neq \emptyset$; incompatibile altrimenti.

Sia $X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ e siano $e_1 = \begin{bmatrix} 1 \\ \vdots \\ 0 \end{bmatrix}, \ldots, e_n = \begin{bmatrix} 0 \\ \vdots \\ 1 \end{bmatrix} \in \mathbb{K}^n$. Poiché

$$X = x_1 e_1 + \cdots + x_n e_n,$$

si ha

$$\begin{aligned} AX &= A(x_1 e_1 + \cdots + x_n e_n) \\ &= x_1 A e_1 + \cdots + x_n A e_n \\ &= x_1 A^1 + \cdots + x_n A^n, \end{aligned}$$

per cui il sistema $AX = b$ è compatibile se e solamente se $b \in \mathcal{L}(A^1, \ldots, A^n)$. Inoltre, $AX = b$ è compatibile se se e solamente se

$$\mathcal{L}(A^1, \ldots, A^n, b) = \mathcal{L}(A^1, \ldots, A^n).$$

In generale, $\mathcal{L}(A^1, \ldots, A^n) \subseteq \mathcal{L}(A^1, \ldots, A^n)$ (verificare!). Quindi è sufficiente dimostrare che la condizione $\mathcal{L}(A^1, \ldots, A^n, b) \subseteq \mathcal{L}(A^1, \ldots, A^n)$ è equivalente alla compatibilità del sistema lineare $AX = b$.

Se $\mathcal{L}(A^1,\dots,A^n,b) \subseteq \mathcal{L}(A^1,\dots,A^n)$, allora $b \in \mathcal{L}(A^1,\dots,A^n)$ e quindi il sistema è compatibile. Viceversa, se il sistema $AX = b$ è compatibile, allora $b \in \mathcal{L}(A^1,\dots,A^n)$. Quindi $b, A_1,\dots,A_n \in \mathcal{L}(A^1,\dots,A^n)$. Poiché $\mathcal{L}(A^1,\dots,A^n)$ è un sottospazio vettoriale di $\mathbb{K}^n$, si ha

$$\mathcal{L}(A^1,\dots,A^n,b) \subseteq \mathcal{L}(A^1,\dots,A^n).$$

Osserviamo, infine, che $\mathcal{L}(A^1,\dots,A^n)$ è l'insieme di tutti termini noti $b \in \mathbb{K}^m$ tali che il sistema $AX = b$ ammette soluzioni. In particolare è un sottospazio vettoriale di $\mathbb{K}^m$.

Un sistema lineare omogeneo è sempre compatibile o risolubile poiché ammette come soluzione il vettore nullo $0_{\mathbb{K}^n} \in \mathbb{K}^n$ ed l'insieme delle soluzioni ammette un sottospazio vettoriale.

Proposizione 4.52. *Sia $AX = 0_{\mathbb{K}^n}$ un sistema lineare omogeneo. Allora $Sol(A|0_{\mathbb{K}^n}) = \{X \in \mathbb{K}^n : AX = 0_{\mathbb{K}^n}\}$ è un sottospazio vettoriale di $\mathbb{K}^n$.*

Dimostrazione. Siano $X, Y \in Sol(A|0_{\mathbb{K}^n})$ e sia $\lambda \in \mathbb{K}$. Allora

$$A(X+Y) = AX + AY = 0_{\mathbb{K}^n},$$

rispettivamente

$$A(\lambda X) = \lambda AX = 0_{\mathbb{K}^n}.$$

Quindi $X+Y, \lambda X \in \mathrm{Sol}(A|0_{\mathbb{K}^n})$ per ogni $X, Y \in \mathrm{Sol}(A|0_{\mathbb{K}^n})$ e per ogni $\lambda \in \mathbb{K}$. □

Osservazione 4.53. *Sia $AX = b$ un sistema lineare di m equazioni in n incognite. Allora $Sol(A|b)$ è un sottospazio vettoriale di $\mathbb{K}^n$ se e solamente se $b = 0_{\mathbb{K}^n}$.*

Sia $AX = b$ un sistema lineare. Diremo che il sistema omogeneo $AX = 0_{\mathbb{K}^n}$ è il sistema lineare omogeneo associato a $AX = b$.

Teorema 4.54 (teorema di struttura)**.** *Sia $AX = b$ un sistema lineare compatibile. Sia X_o una soluzione particolare del sistema $AX = b$. Allora ogni altra soluzione del sistema lineare $AX = b$ è della forma $X_o + W$, dove W è una soluzione del sistema lineare omogeneo associato $AX = 0_{\mathbb{K}^n}$; Quindi*

$$Sol(A|b) = \{X_o + W,\ W \in Sol(A|0_{\mathbb{K}^n})\}.$$

Dimostrazione. Indichiamo con $E = \{X_o + W,\ X \in Sol(A|0_{\mathbb{K}^n})\}$. Vogliamo dimostrare che $E = Sol(A|b)$.

Sia $Y \in Sol(A|b)$. Allora $A(Y - X_o) = AX - AX_o = b - b = 0_{\mathbb{K}^n}$, ovvero $Y - X_o$ è soluzione del sistema lineare omogeneo associato, da cui segue che $Y - X_o \in Sol(A|0_{\mathbb{K}^n})$, cioé $Y = X_o + W$ per un certo $W \in Sol(A|0_{\mathbb{K}^n})$. Quindi $Sol(A|b) \subseteq E$. Viceversa, ogni elemento di E è della forma $X_o + W$, dove W è una soluzione del sistema lineare omogeneo associato. Allora

$$A(X_o + W) = AX_o + AW = b + 0_{\mathbb{K}^n} = b,$$

da cui segue che $E \subseteq Sol(A|b)$. Quindi $Sol(A|b) = \{X_o + X,\ X \in Sol(A|0_{\mathbb{K}^n})\}$. □

Definizione 4.55. *Siano $A \in M_{m\times n}(\mathbb{K})$, $B \in M_{m'\times n}(\mathbb{K})$. Diremo che i due sistemi lineari $AX = b$ e $A'X = c$ si dicono* equivalenti *se* $\mathrm{Sol}(A|b) = \mathrm{Sol}(A'|c)$*, i.e., se hanno le stesse soluzioni.*

Osservazione 4.56. *Sia $AX = b$ un sistema lineare. Se scambio due righe alla matrice $(A|b)$, il sistema lineare associato è equivalente al sistema di partenza $AX = b$.*

Definizione 4.57. *Date due equazioni $a_1x_1 + \cdots + a_nx_n = a$ e $b_1x_1 + \cdots + b_nx_n = b$, si dice combinazione lineare delle due equazioni a coefficienti $h, k \in \mathbb{K}$, l'equazione*

$$h(a_1x_1 + \cdots + a_nx_n) + k(b_1x_1 + \cdots + b_nx_n) = ha + kb.$$

Lemma 4.58. *Sia $AX = d$ un sistema lineare contenente le equazioni*

$$\begin{array}{c} a_1x_1 + \cdots + a_nx_n = a \\ b_1x_1 + \cdots + b_nx_n = b \end{array}.$$

Sia $BX = \tilde{d}$ il sistema lineare ottenuto sostituendo in $AX = d$ l'equazione

$$h(a_1x_1 + \cdots + a_nx_n) + k(b_1x_1 + \cdots + b_nx_n) = ha + kb$$

al posto dell'equazione

$$b_1x_1 + \cdots + b_nx_n = b\ ,$$

dove $h, k \in \mathbb{K}$ con $k \neq 0$. Allora i sistemi $AX = d$ e $BX = \tilde{d}$ sono equivalenti.

Facoltativa. Sia $(A|d) \in \mathrm{M}_{m\times(n+1)}(\mathbb{K})$ la matrice completa. A meno di scambiare righe possiamo supporre che le equazioni

$$\begin{array}{c} a_1x_1 + \cdots + a_nx_n = a \\ b_1x_1 + \cdots + b_nx_n = b \end{array}.$$

siano la 1 e la 2 equazione rispettivamente del sistema lineare $AX = d$, per cui la matrice completa è così siffatta:

$$(A|d) = \left[\begin{array}{ccc|c} a_1 & \cdots & a_n & a \\ b_1 & \cdots & b_n & b \\ \vdots & \vdots & \vdots & \vdots \\ \cdots & \cdots & \cdots & \cdots \end{array}\right].$$

Quindi

$$(B|\tilde{d}) = \left[\begin{array}{ccc|c} a_1 & \cdots & a_n & a \\ ka_1 + kb_1 & \cdots & ka_n + hb_n & ha + kb \\ \vdots & \vdots & \vdots & \vdots \\ \cdots & \cdots & \cdots & \cdots \end{array}\right].$$

Se $\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ è soluzione del sistema $AX = d$, allora soddisfa sicuramente tutte le equazioni del sistema $BX = \tilde{d}$ poiché l'unica equazione differente è una combinazione delle prime due. Viceversa, se $\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ è soluzione del sistema $BX = \tilde{d}$, allora le coordinate del vettore $\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ sono sicuramente soluzione di tutte le equazioni del sistema $AX = d$ tranne la seconda, che andiamo adesso a trattare. Poiché vale

$$a_1x_1 + \cdots + a_nx_n = a$$

allora la seconda equazione di $BX = \tilde{d}$ diventa

$$ha + k(b_1x_1 + \cdots + b_nx_n) = ha + kb.$$

Essendo $k \neq 0$ si ha $b_1x_1 + \cdots + b_nx_n = b$. $\square$

Sia $AX = b$ un sistema lineare. Osserviamo che scambiare due equazioni del sistema $AX = b$ è equivalente a scambiare due righe della matrice $(A|b)$. Analogamente moltiplicare una equazione per uno scalare non nullo, sommare ad una equazione un multiplo di un'altra, è equivalente a moltiplicare una riga della matric $(A|b)$ per uno scalare non nullo, rispettivamente sommare ad una riga di $(A|b)$ un multiplo di un'altra. Quindi, il lemma anteriore prova il seguente risultato.

Corollario 4.59. *Sia $AX = b$ un sistema lineare ed indichiamo con $(A|b)$ la matrice completa. Sia $(C|d)$ la matrice ottenuta attraverso operazioni elementari di riga sulla matrice completa $(A|b)$. I sistemi lineari $AX = b$ e $CX = d$ sono equivalenti ovvero l'insieme delle soluzioni di un sistema lineare non cambia quando si effettuano operazioni elementari di riga sulla matrice completa.*

4.59.1 Sistemi ridotti a scala

Definizione 4.60. *Un sistema $SX = c$ si dice ridotto a scala se la matrice S è ridotta a scala.*

Esempio 4.61.

$$\begin{cases} x_1 + x_2 - x_3 = 2 \\ x_2 - x_4 = 0 \\ x_3 + x_4 = 8 \end{cases}$$

è un sistema ridotto a scala poiché la matrice incompleta ha la forma

$$S = \begin{pmatrix} 1 & 1 & -1 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

Proposizione 4.62. *Sia $SX = c$ un sistema ridotto a scala, dove $S \in M_{m\times n}(\mathbb{K})$, con $rg(S) = r$. Il sistema $SX = c$ è compatibile se e solamente se $m = r$, oppure le ultime $m-r$ coordinante del vettore c sono nulle. Inoltre le soluzioni, se esistono, dipendono da $n - \mathrm{rg}(S)$ parametri.*

Dimostrazione. Sia $SX = c$ un sistema ridotto a scala. La matrice completa ha la forma

$$\left[\begin{array}{cccccccccccccc|c}
0 & \cdots & 0 & s_{1j_1} & * & * & \cdots & \cdots & * & \cdots & \cdots & \cdots & * & * & c_1 \\
0 & \cdots & \cdots & \cdots & 0 & s_{2j_2} & * & * & * & \cdots & \cdots & \cdots & * & * & c_2 \\
0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & s_{3j_3} & * & * & \cdots & \cdots & \cdots & * & c_3 \\
\vdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & \ddots & \cdots & \cdots & \cdots & \cdots & * & \vdots \\
\vdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & \ddots & \cdots & \cdots & * & * & \vdots \\
0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & s_{rj_r} & \cdots & \cdots & * & c_r \\
\hline
0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & 0 & \cdots & 0 & c_{r+1} \\
\vdots & & & & & & & & & & \vdots & \vdots & & \vdots & \vdots \\
\vdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \vdots & \vdots \\
0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & 0 & \cdots & \cdots & 0 & c_m
\end{array}\right]$$

La condizione, $m = r$ oppure le ultime $m - r$ coordinante del vettore c sono nulle è sicuramente necessaria. Dobbiamo dimostrare che questa condizione è anche sufficiente. Supponiamo che le ultime $m - r$ coordinate siano nulle. L'altro caso è analogo. Allora la matrice completa è così siffatta.

$$\left[\begin{array}{cccccccccccccc|c}
0 & \cdots & 0 & s_{1j_1} & * & * & \cdots & \cdots & * & \cdots & \cdots & \cdots & * & * & c_1 \\
0 & \cdots & \cdots & \cdots & 0 & s_{2j_2} & * & * & * & \cdots & \cdots & \cdots & * & * & c_2 \\
0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & s_{3j_3} & * & * & \cdots & \cdots & \cdots & * & c_3 \\
\vdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & \ddots & \cdots & \cdots & \cdots & \cdots & * & \vdots \\
\vdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & \ddots & \cdots & \cdots & * & * & \vdots \\
0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & s_{rj_r} & \cdots & \cdots & * & c_r \\
\hline
0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & 0 & \cdots & 0 & 0 \\
\vdots & & & & & & & & & & \vdots & \vdots & & \vdots & \vdots \\
\vdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \vdots & \vdots \\
0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & 0 & \cdots & \cdots & 0 & 0
\end{array}\right]$$

L'elemento s_{rj_r} è differente di zero. Quindi possiamo scrivere la variabile

$$x_{j_r} = s_{rj_r}^{-1} \left(c_r - s_{rj_r+1} x_{j_r+1} - \cdots - s_{rn} x_n\right)$$

in funzione delle variabili $x_{j_r+1}, \ldots, x_n$. Analogamente,

$$x_{j_{r-1}} = s_{r-1j_{r-1}}^{-1} (c_{r-1} - s_{r-1j_{r-1}+1} x_{j_{r-1}+1} - \cdots - s_{r-1n} x_n)$$

e sostituendo ad x_{j_r} il valore precedente, possiamo scrivere $x_{j_{r-1}}$ in funzione delle variabili $x_{j_{r-1}+1}, \ldots x_{j_r-1}, x_{j_r+1}, \ldots, x_n$; e così via. Quindi le soluzioni esistono. Inoltre abbiamo rappresentato le variabili corrispondenti ai perni, che sono $\mathrm{rg}(S)$, in funzioni delle rimanenti, ovvero le soluzioni dipendono da $n - \mathrm{rg}(S)$ parametri. $\square$

Questo metodo si chiama *metodo della risoluzione all'indietro*. Le variabili corrispondenti agli elementi di testa, $x_{j_1}, \ldots, x_{j_r}$, che chiameremo *variabili basiche* si scrivono in funzione delle altre, che chiameremo *variabili libere*. Questo significa che le soluzioni dipendono da esattamente $n - r$, ovvero una volta che abbiamo assegnato un valore alle variabili libere, le variabili basiche sono univocamente determinate.

Proposizione 4.63. *Sia $SX = c$ un sistema ridotto a scala, dove $S \in M_{m\times n}(\mathbb{K})$, con $rg(S) = r$. Il sistema $SX = c$ è compatibile se e solamente se $\mathrm{rg}(S) = \mathrm{rg}(S|c)$. Inoltre le soluzioni, se esistono, dipendono da $n - \mathrm{rg}(S)$ parametri.*

Dimostrazione. Utilizzando la notazione anteriore, abbiamo visto che il sistema $SX = c$ è compatibile se e solamente se $m = \mathrm{rg}(S)$ oppure le ultime $m - \mathrm{rg}(S)$ coordinate di c sono nulle. Questo è equivalente alla condizione che la matrice completa $(S|c)$ è ridotta a scala e $\mathrm{rg}(S) = \mathrm{rg}(S|c)$. $\square$

Teorema 4.64 (Rouché-Capelli)**.** *Sia $AX = b$, un sistema lineare con $A \in M_{m\times n}(\mathbb{K})$. Il sistema è compatibile se e solamente se $rg(A) = rg(A|b)$. In tal caso, le soluzioni dipendono da $n - rg(A)$ parametri.*

Dimostrazione. Sia $(A|b)$ la matrice completa. Indichiamo con $(S|c)$ una sua riduzione a scala ottenuta attraverso il metodo di Gauss. La matrice S è una riduzione a scala di A ed il sistema lineare $SX = c$ è equivalente al sistema $AX = b$. Quindi il sistema lineare $AX = b$ è compatibile se e solamente se $SX = c$ è compatibile per cui se solamente se se $m = r$, oppure le ultime $m-r$ coordinante del vettore c sono nulle, dove $r = \mathrm{rg}(S)$. Questa condizione è equivalente a $rg(S) = rg(S|c)$, ovvero il sistema $AX = b$ è compatibile se e solamente se $rg(A) = r(S) = rg(S|b) = rg(A|b)$. La secondo parte del Teorema segue dal procedimento della risoluzione all'indietro. $\square$

Corollario 4.65. *Sia $A \in M_{m\times n}(\mathbb{K})$ e sia $AX = b$ un sistema lineare. Se il sistema $AX = b$ è compatibile, allora Il sistema $AX = b$ ammette una ed una sola soluzione se e solamente se $\mathrm{rg}(A) = n$.*

Dimostrazione. Poiché il sistema è compatibile, applicando il Teorema di Rouché-Capelli, le soluzioni dipendono da $n - \mathrm{rg}(A)$ parametri. Quindi $AX = b$ ammette una ed una sola soluzione se e solamente se il numero dei parametri è zero e quindi se e solamente se $\mathrm{rg}(A) = n$. $\square$

Osservazione 4.66. $\mathrm{rg}(A) \leq \mathrm{rg}(A|b)$. *Inoltre*

$$\mathrm{rg}(A|b) = \begin{cases} \mathrm{rg}(A) \\ \text{oppure} \\ \mathrm{rg}(A)+1 \end{cases}$$

Corollario 4.67. *Un sistema lineare omogeneo $AX = 0$ di m equazioni in n incognite ammette soluzioni non banali, se e solamente se $rg(A) < n$. In particolare se $m < n$, allora il sistema lineare $AX = 0$ ammette sempre soluzioni non banali.*

Dimostrazione. Applicando il Teorema di Rouché-Capelli, le soluzione del sistema lineare omogeneo dipendono da $n - \mathrm{rg}(A)$ parametri. Quindi il sistema omogeneo $AX = 0$ ammette soluzioni non banali se e solamente se $\mathrm{rg}(A) < n$. Se $A \in M_{m\times n}(\mathbb{R})$ con $m < n$, tenendo in mente che $\mathrm{rg}(A) \leq m$, si ha

$$n - \mathrm{rg}(A) \geq n - m > 0,$$

ovvero ammette sempre soluzioni non banali. □

Corollario 4.68 (Teorema di Cramer)**.** *Sia $A = (A^1, \ldots, A^n) \in M_{n\times n}(\mathbb{K})$. Il sistema $AX = b$ ammette una ed una soluzione se e solamente se A è invertibile. Se $X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ è l'unica soluzione, allora*

$$x_i = \frac{\det(A^1, \ldots, A^{i-1}, b, A^{i+1}, \ldots, A^n)}{\det(A)}$$

per $i = 1, \ldots, n$.

Dimostrazione. Se il sistema ammette una ed una soluzione, per il Teorema di Rouché-Capelli, $\mathrm{rg}(A) = n$, ovvero A è invertibile.

Se A è invertibile allora $\mathrm{rg}(A) = n$. Poiché $(A|b) \in \mathrm{M}_{n\times(n+1)}(\mathbb{K})$, si ha

$$n = \mathrm{rg}(A) \leq \mathrm{rg}(A|b) \leq n,$$

ovvero $\mathrm{rg}(A) = \mathrm{rg}(A|b) = n$. Applicando il Teorema di Rouché-Capelli, il sistema è compatibile ed ammette una ed una sola soluzione.

Sia $X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ l'unica soluzione del sistema $AX = b$. Poiché

$$x_1A^1 + \ldots + x_nA^n = b,$$

si ha

$$\begin{aligned}\det(A^1,\dots,A^{i-1},b,A^{i+1},\dots,A^n) &= \det(A^1,\dots,A^{i-1},\sum_{m=1}^{n} x_m A^m, A^{i+1},\dots,A^n)\\ &= \sum_{m=1}^{n} x_m \det\underbrace{(A^1,\dots,A^{i-1},A^m,A^{i+1},\dots,A^n)}_{i}\\ &= x_i \det A.\end{aligned}$$

Poiché $\det A \neq 0$ si ha la tesi. $\square$

4.68.1 Metodi di Calcolo

Siano $X_1,\dots,X_s \in \mathbb{K}^n$. Sia $A = (X_1,\dots,X_s) \in \mathrm{M}_{n\times s}(\mathbb{K})$. Poiché $\mathrm{rg}(A)$ è il massimo numero dei vettori $X_1,\dots,X_s$ linearmente indipendenti, si ha che $X_1,\dots,X_s$ sono linearmente dipendenti $\iff rg(A) < s$, rispettivamente $X_1,\dots,X_s$ sono linearmente indipendenti $\iff rg(A) = s$. Se $s = n$, allora $X_1,\dots X_n$ sono linearmente indipendenti se e solamente se $\mathrm{rg}(A) = n$ ovvero se e solamente se $\det(A) \neq 0$. Se $s > n$, tenendo in mente che $A \in M_{n\times s}(\mathbb{K})$ e quindi $\mathrm{rg}(A) \leq n < s$, si ha che i vettori $X_1,\dots,X_s$ sono linearmente dipendenti.

Una combinazione lineare di $X_1,\dots,X_s$ è un vettore Z per il quale esistono $\alpha_1,\dots,\alpha_s \in \mathbb{K}$ tali che

$$\alpha_1 X_1 + \cdots + \alpha_s X_s = Z.$$

Se indichiamo con $A = (X_1,\dots,X_s) \in \mathrm{M}_{n\times s}(\mathbb{K})$, si ha

$$A \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_s \end{bmatrix} = Z.$$

Quindi il vettore Z è combinazione lineare di $X_1,\dots X_s$ se e solamente se il sistema lineare $AX = Z$ è compatibile, ovvero se e solamente se

$$\mathrm{rg}(A) = \mathrm{rg}(A|Z).$$

4.68.2 Mutua posizione di rette e piani

Siano date $r_1 : X = P_1 + tA_1$ e $r_2 : X = P_2 + tA_2$ rette nello spazio. Vogliamo studiare la mutua posizione di r_1 e r_2 nello spazio.

Le due rette hanno punti in comune se e solo se esistono $s_1, t_1 \in \mathbb{R}$ tali che

$$P_1 + t_1 A_1 = P_2 + s_1 A_2,$$

ovvero se e solamente se

$$P_1 - P_2 = (-t_1)A_1 + s_1 A_2 = (A_1, A_2) \begin{bmatrix} -t_1 \\ s_1 \end{bmatrix}.$$

Quindi r_1 e r_2 hanno punti in comune se e solamente se il sistema lineare $AX = P_1 - P_2$, dove $A = (A_1, A_2)$, è compatibile. Poiché i vettori direttori di una retta sono non nulli, la matrice A può avere rango 1 oppure 2. Applicando il Teorema di Rouché-Capelli, si ha:

a) $rg(A) = 1$, allora se $rg(A|P_1 - P_2) = 1$ le due rette sono coincidenti; se $rg(A|P_1 - P_2) = 2$ le due rette sono parallele;

b) $rg(A) = 2$, allora se $rg(A|P_1 - P_2) = 2$, allora le due rette sono incidenti; se $rg(A|P_1 - P_2) = 3$, allora le due rette sono sghembe.

Corollario 4.69. *Le due rette r ed s sono sghembe se e solamente se*

$$\det(A_1, A_2, P_1 - P_2) \neq 0.$$

Consideriamo le due rette in forma cartesiana:

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

$$s = \begin{cases} a''x + b''y + c''z = d'' \\ a'''x + b'''y + c'''z + = d''' \end{cases}$$

Se indichiamo con

$$A = \begin{pmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \\ a''' & b''' & c''' \end{pmatrix} \in M_{4\times 3}(\mathbb{R})$$

e con $h = \begin{bmatrix} d \\ d' \\ d'' \\ d''' \end{bmatrix}$, allora i punti di $r \cap s$ soddisfano il seguente sistema lineare

$$AX = h,$$

dove $X = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$. Poiché una retta è intersezioni di due piano non paralleli, il rango della matrice A può essere 2 oppure 3. Applicando il Teorema di Rouché-Capelli, otteniamo

a) $rg(A) = 2$. Se $rg(A|h) = 2$ il sistema è compatibile e le due rette sono coincidenti; se $rg(A|h) = 3$, allora il sistema è incompatibile e le due rette sono parallele;

b) $rg(A) = 3$. Se $rg(A|h) = 3$ il sistema è compatibile e le due rette sono incidenti; se $rg(A|h) = 4$, allora il sistema è incompatibile e le due rette sono sghembe;

Corollario 4.70. *Le due rette r ed s sono sghembe se e solamente se*

$$\det(A|h) \neq 0.$$

Siano $\pi : ax + bx + cz = d$ e $\pi' : a'x + b'y + c'z = d'$ piani nello spazio. Sia

$$A = \begin{pmatrix} a & b & c \\ a' & b' & c' \end{pmatrix} \in M_{2\times 3}(\mathbb{R})$$

e sia

$$(A|d) = \left(\begin{array}{ccc|c} a & b & c & d \\ a' & b' & c' & d' \end{array}\right) \in M_{2\times 4}(\mathbb{R})$$

Quindi $\pi \cap \pi' = \mathrm{Sol}(A|d)$. Poiché $A \in M_{2\times 3}(\mathbb{R})$, la matrice A può avere rango 1 oppure 2. Applicando il teorema di Roucyhé-Capelli si ha i seguenti casi:

a) $rg(A) = 1$. Se anche $rg(A|d) = 1$ allora i due piani sono coincidenti. Invece se $rg(A|d) = 2$ i due piani sono paralleli;

b) $rg(A) = 2$, allora anche $rg(A|d) = 2$. Applicando il Teorema di Rouché-Capelli le soluzioni dipendendono da un parametro. Quindi due piani non paralleli e non coincidenti si intersecano lungo una retta.

Siano $\pi : ax + by + cz = d$ un piano e

$$r = \begin{cases} a'x + b'y + c'z = d' \\ a''x + b''y + c''z = d'' \end{cases},$$

una retta nello spazio. Un vettore $P \in r \cap \pi$ se e solamente se il sistema

$$AX = h$$

dove

$$A = \begin{pmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \end{pmatrix} \in M_{3\times 3}(\mathbb{R})$$

e

$$h = \begin{bmatrix} d \\ d' \\ d'' \end{bmatrix},$$

è compatibile. Il rango della matrice A può essere 2 oppure 3. Applicando Rouché-Capelli, si ha:

a) $rg(A) = 2$. Se $rg(A|h) = 2$, allora il sistema è compatibile e la retta è contenuta nel piano; se $rg(A|h) = 3$, allora le retta è parallela al piano;

b) $rg(A) = 3$, allora anche $rg(A|h) = 3$ e quindi il piano π e la retta r sono incidenti.

Utilizzando la notazione anteriore, proviamo i seguenti risultati.

Corollario 4.71. *$r \subset \pi$ se e solamente se $\pi : \alpha(a'x + b'y + c'z - d') + \beta(a''x + b''y + c''z - d'') = 0$ con $\alpha, \beta \in \mathbb{R}$ non entrambi nulli.*

Dimostrazione. La retta r è contenuta nel piano π se e solamente se il sistema ammette soluzione e le soluzioni dipendono da un parametro, quindi se e solamente se $rg(A) = rg(A|h) = 2$. Poiché le ultime due righe della matrice $(A|h)$ sono linearmente indipendenti, si ha che la prima riga è combinazione lineare della seconda e della terza, da cui segue la tesi. □

Corollario 4.72. *La retta r ed il piano π sono incidenti se e solamente se* $\det(A) \neq 0$

Dimostrazione. La retta r ed il piano π sono incidenti se e solamente se $\mathrm{rg}(A|h) = 3$. Poiché $(A|h) \in \mathrm{M}_{3\times 3}(\mathbb{R})$ si ha che r ed π sono incidenti se e solamente se $\det(A) \neq 0$. □

Capitolo 5

Basi e dimensione di uno spazio vettoriale

5.1 Basi di uno spazio vettoriale

Sia V uno spazio vettoriale su $\mathbb{K}$.

Definizione 5.2. *Un insieme $\mathcal{B} = \{v_1, \ldots, v_n\}$ si dice una base di V se:*

- $v_1, \ldots, v_n$ *sono linearmente indipendenti:*
- $v_1, \ldots, v_n$ *formano un sistema di generatori, i.e., $\mathcal{L}(v_1, \ldots, v_n) = V$.*

a) i vettori $e_1 = \begin{bmatrix} 1 \\ \vdots \\ 0 \end{bmatrix}, \ldots, e_n = \begin{bmatrix} 0 \\ \vdots \\ 1 \end{bmatrix}$ formano una base di $\mathbb{K}^n$. La base $\mathcal{C} = \{e_1, \ldots, e_n\}$ è chiamata la *base canonica*;

b) siano $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \in \mathrm{M}_{2\times 2}(\mathbb{R})$. Poiché
$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = a\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + b\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + c\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + d\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$
si ha che i vettori $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ formano una base di $M_{2\times 2}(\mathbb{R})$;

c) sia $E_{ij} \in M_{m\times n}(\mathbb{K})$ la matrice i cui elementi sono tutti nulli tranne l'elemento $a_{ij} = 1$. È facile provare che $\mathcal{B} = \{E_{ij}, 1 \leq i \leq m, 1 \leq j \leq n\}$ è una base di $M_{m\times n}(\mathbb{K})$;

d) i polinomi $\{1, x, \ldots, x^n\}$ formano una base di $\mathbb{K}_n[x]$;

e) i vettori $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, formano una base di $\mathbb{R}^3$. Infatti, sia $\begin{bmatrix} x \\ y \\ z \end{bmatrix} \in \mathbb{R}^3$. Allora

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \alpha \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \beta \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + \gamma \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix},$$

se il sistema

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & x \\ 0 & 1 & 1 & y \\ 0 & 0 & 1 & z \end{array}\right],$$

è compatibile. Poiché la matrice $\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$ è inveritible, applicando il Teorema di Cramer si ha che il sistema è compatibile ed ammette una ed una sola soluzione. Quindi i vettori $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ sono linearmente indipendenti e formano un sistema di generatori.

f) i vettori $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ formano una base di $\mathrm{M}_{2\times 2}(\mathbb{R})$. Infatti,

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \alpha \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \beta \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + \gamma \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} + \delta \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

se e solamente se il sistema

$$\begin{cases} \alpha & = & a \\ \beta + \gamma & = & b \\ \beta - \gamma & = & c \\ \delta & = & d \end{cases},$$

è compatibile. La matrice dei coefficienti è

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

quindi invertible. Applicando Il Teorema di Cramer si ha che il sistema ammette una ed una soluzione. Quindi i vettori $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$, $\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ formano una base di $\mathrm{M}_{2\times 2}(\mathbb{R})$.

Proposizione 5.3. *Sia V uno spazio vettoriale su $\mathbb{K}$ e sia $\mathcal{B} = \{v_1, \dots, v_n\}$ una base di V. Allora ogni vettore $v \in V$ si può scrivere in maniera unica come combinazione lineare dei vettori $v_1, \dots v_n$. Viceversa, se ogni elemento si può scrive in maniera unica come combinazione lineare di $v_1, \dots v_n$, allora $\mathcal{B} = \{v_1, \dots, v_n\}$ è una base di V.*

Dimostrazione. Sia $v \in V$. Poiché $\mathcal{B} = \{v_1, \dots, v_n\}$ è un sistema di generatori si ha $V = \mathcal{L}(v_1, \dots, v_n)$. Quindi esistono $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ tale che

$$v = \lambda_1 v_1 + \cdots + \lambda_n v_n.$$

Supponiamo che

$$v = \lambda_1 v_1 + \cdots + \lambda_n v_n = \alpha_1 v_1 + \cdots + \alpha_n v_n.$$

Allora

$$0 = (\lambda_1 - \alpha_1) v_1 + \cdots + (\lambda_n - \alpha_n) v_n.$$

Essendo $\mathcal{B} = \{v_1, \dots, v_n\}$ un insieme di vettori linearmente indipendenti, ne segue che

$$\alpha_1 = \lambda_1, \dots, \alpha_k = \lambda_k.$$

Viceversa, supponiamo che ogni elemento si scrive in maniera unica come combinazione lineare dei vettori $v_1, \dots, v_n$. Vogliamo provare che $\mathcal{B} = \{v_1, \dots, v_n\}$ è una base di V, ovvero i vettori $v_1, \dots, v_n$ formano un sistema di generatori e sono vettori linearmente indipendenti.

Sia $v \in V$. Poiché ogni elemento si scrive come combinazione lineare dei vettori $v_1, \dots, v_n$, esistono $\alpha_1, \dots, \alpha_n \in \mathbb{K}$ tali che $v = \alpha_1 v_1 + \cdots + \alpha_n v_n$. Quindi $v \in \mathcal{L}(v_1, \dots, v_n)$ da cui segue che

$$\mathcal{L}(v_1, \dots, v_n) = V.$$

ovvero $v_1, \dots, v_n$ formano un sistema di generatori. Adesso proviamo che sono linearmente indipendenti.

Siano $\alpha_1, \dots, \alpha_n \in \mathbb{K}$ tali che $\alpha_1 v_1 + \cdots + \alpha_n v_n = 0$. Dall'unicità, segue che $\alpha_1 = \cdots = \alpha_n = 0$, poiché il vettore nullo si può scrivere come combinazione lineare dei vettori $v_1, \dots, v_n$ con tutti i coefficienti nulli. Quindi i vettori $v_1, \dots v_n$ sono linearmente indipendenti. $\square$

Definizione 5.4. *Sia* $v \in V$ *e sia* $\mathcal{B} = \{v_1, \dots, v_n\}$ *una base di* V*. Le coordinate di* v *rispetto a* $\mathcal{B}$ *sono gli unici scalari* $x_1, \dots, x_n \in \mathbb{K}$ *tali che*

$$v = x_1 v_1 + \cdots + x_n v_n.$$

Indicheremo con $[v]_{\mathcal{B}} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{K}^n$ *le coordinate di* v *rispetto a* $\mathcal{B}$*.*

Esempio 5.5. *I vettori* $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$*, formano una base di* $\mathbb{R}^3$*. Vogliamo calcolare le coordinate di un vettore rispetto alla base* $\mathcal{B}$*. Quindi dobbiamo calcolare* $\alpha, \beta, \gamma \in \mathbb{R}$ *tali che*

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \alpha \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \beta \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + \gamma \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix},$$

ovvero risolvere il sistema

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & x \\ 0 & 1 & 1 & y \\ 0 & 0 & 1 & z \end{array}\right].$$

Il sistema è già ridotto a scala. Applicando il metodo della risoluzione all'indietro si ha

$$\begin{cases} \gamma = z \\ \beta = y - z \\ \alpha = x - y \end{cases},$$

ovvero

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = (x-y) \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + (y-z) \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + z \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

Quindi $\begin{bmatrix} x \\ y \\ z \end{bmatrix}_{\mathcal{B}} = \begin{bmatrix} x - y \\ y - z \\ z \end{bmatrix}$.

Sia $\mathcal{B} = \{v_1, \dots, v_n\}$ una base di V. Indicheremo con

$$\mathcal{F}_{\mathcal{B}} : V \longrightarrow \mathbb{K}^n \qquad v \mapsto [v]_{\mathcal{B}}$$

l'applicazione che associa ad ogni vettore le sue coordinate rispetto alla base $\mathcal{B}$. Tale applicazione è iniettiva e suriettiva, ovvero una trasformazione biunivoca che soddisfa alle seguenti proprietà:

- se $v, w \in V$, allora $v = w$ se e solamente se $[v]_{\mathcal{B}} = [w]_{\mathcal{B}}$;
- $[0_V]_{\mathcal{B}} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$;
- $[v + w]_{\mathcal{B}} = [v]_{\mathcal{B}} + [w]_{\mathcal{B}}$;
- $[\lambda v]_{\mathcal{B}} = \lambda [v]_{\mathcal{B}}$;
- Sia $\{e_1, \ldots, e_n\}$ la base canonica di $\mathbb{K}^n$. Allora $[v_j]_{\mathcal{B}} = e_j$, per $j = 1, \ldots, n$.

La prima proprietà è una diretta conseguenza dell'iniettività dell'applicazione $\mathcal{F}_{\mathcal{B}}$. Infatti, $v = w$ se e solamente se $\mathcal{F}_{\mathcal{B}}(v) = \mathcal{F}_{\mathcal{B}}(w)$ quindi se e solamente se $[v]_{\mathcal{B}} = [w]_{\mathcal{B}}$.

Poiché $0_V = 0v_1 + \cdots + 0v_n$, si ha $[0_V]_{\mathcal{B}} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$.

Siano $v, w \in V$. Allora

$$v = x_1 v_1 + \cdots + x_n v_n,$$

rispettivamente

$$w = y_1 v_1 + \cdots + y_n v_n.$$

Quindi

$$v + w = (x_1 + y_1) v_1 + \cdots + (x_n + y_n) v_n,$$

ovvero

$$[v + w]_{\mathcal{B}} = \begin{bmatrix} (x_1 + y_1) \\ \vdots \\ (x_n + y_n) \end{bmatrix} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = [v]_{\mathcal{B}} + [w]_{\mathcal{B}}.$$

Se

$$v = x_1 v_1 + \cdots + x_n v_n,$$

allora

$$\lambda v = \lambda x_1 v_1 + \cdots + \lambda x_n v_n.$$

Quindi

$$[\lambda v]_{\mathcal{B}} = \begin{bmatrix} \lambda x_1 \\ \vdots \\ \lambda x_n \end{bmatrix} = \lambda \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \lambda [v]_{\mathcal{B}}.$$

Sia $v_1 \in V$. Allora
$$v_1 = 1v_1 + 0v_2 + \cdots + 0v_n.$$
Quindi $[v_1]_{\mathcal{B}} = \begin{bmatrix} 1 \\ \vdots \\ 0 \end{bmatrix} = e_1$. In maniera analoga,
$$v_j = 0v_1 + \cdots + 0v_{j-1} + 1v_j + 0v_{j+1} + \cdots + 0v_n,$$
ovvero $[v_j]_{\mathcal{B}} = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} = e_j$.

Osserviamo inoltre che l'applicazione $\mathcal{F}_{\mathcal{B}}$ trasforma combinazioni lineari in combinazioni lineari.

Siano $w_1, \ldots, w_k \in V$ e sia $\alpha_1 w_1 + \cdots + \alpha_k w_k$ una combinazione lineare dei vettori $w_1, \ldots, w_k$ a coefficienti $\alpha_1, \ldots, \alpha_k \in \mathbb{K}$. Quindi
$$\begin{aligned}\mathcal{F}_{\mathcal{B}}(\alpha_1 w_1 + \cdots + \alpha_k w_k) &= \mathcal{F}_{\mathcal{B}}(\alpha_1 w_1) + \mathcal{F}_{\mathcal{B}}(\alpha_2 w_2 + \cdots + \alpha_k w_k) \\ &= \mathcal{F}_{\mathcal{B}}(\alpha_1 w_1) + \cdots + \mathcal{F}_{\mathcal{B}}(\alpha_k w_k) \\ &= \alpha_1 \mathcal{F}_{\mathcal{B}}(w_1) + \cdots + \alpha_k \mathcal{F}_{\mathcal{B}}(w_k),\end{aligned}$$
ovvero
$$[\alpha_1 w_1 + \cdots + \alpha_k w_k]_{\mathcal{B}} = \alpha_1 [w_1]_{\mathcal{B}} + \cdots + \alpha_k [w_k]_{\mathcal{B}}.$$
Quindi, $\mathcal{F}_{\mathcal{B}}$ trasforma una combinazione lineare dei vettori $w_1, \ldots, w_k$ a coefficienti $\alpha_1, \ldots, \alpha_k \in K$ in una combinazione lineare dei vettori $[w_1]_{\mathcal{B}}, \ldots, [w_k]_{\mathcal{B}}$ sempre a coefficienti $\alpha_1, \ldots, \alpha_k \in \mathbb{K}$.

Proposizione 5.6. *Siano $w_1, \ldots, w_k \in V$ e sia $\mathcal{B} = \{v_1, \ldots, v_n\}$ una base di V. Allora:*

a) $w_1, \ldots, w_k$ sono linearmente dipendenti, rispettivamente linearmente indipendenti, se e solamente se $[w_1]_{\mathcal{B}}, \ldots, [w_k]_{\mathcal{B}} \in \mathbb{K}^n$ sono linearmente dipendenti, rispettivamente linearmente indipendenti;

b) $v \in V$ è combinazione lineare dei vettori $w_1, \ldots, w_k$ se e solamente se $[v]_{\mathcal{B}} \in \mathbb{K}^n$ è combinazione lineare dei vettori $[w_1]_{\mathcal{B}}, \ldots, [w_k]_{\mathcal{B}} \in \mathbb{K}^n$;

c) $\mathcal{F}_{\mathcal{B}}(\mathcal{L}(w_1, \ldots, w_k)) = \mathcal{L}([w_1]_{\mathcal{B}}, \ldots, [w_k]_{\mathcal{B}})$.

Dimostrazione. Siano $w_1, \dots, w_k$ linearmente dipendenti. Allora esistono $\alpha_1, \dots, \alpha_k \in \mathbb{K}$ non tutti nulli tali che $\alpha_1 w_1 + \cdots + \alpha_k w_k = 0_V$. Quindi

$$\begin{aligned} 0_{\mathbb{K}^n} &= \mathcal{F}_{\mathcal{B}}(0_V) \\ &= \mathcal{F}_{\mathcal{B}}(\alpha_1 w_1 + \cdots + \alpha_k w_k) \\ &= \alpha_1 \mathcal{F}_{\mathcal{B}}(w_1) + \cdots + \alpha_k \mathcal{F}_{\mathcal{B}}(w_k) \\ &= \alpha_1 [w_1]_{\mathcal{B}} + \cdots + \alpha_k [w_k]_{\mathcal{B}}, \end{aligned}$$

ovvero i vettori $[w_1]_{\mathcal{B}}, \dots, [w_k]_{\mathcal{B}}$ sono linearmente dipendenti. Viceversa, se i vettori $[w_1]_{\mathcal{B}}, \dots, [w_k]_{\mathcal{B}}$ sono linearmente dipendenti, allora esistono $\alpha_1, \dots, \alpha_k \in \mathbb{K}$ non tutti nulli tali che

$$\alpha_1 [w_1]_{\mathcal{B}} + \cdots + \alpha_k [w_k]_{\mathcal{B}} = 0_{\mathbb{K}^n}.$$

Quindi

$$\mathcal{F}_{\mathcal{B}}(\alpha_1 w_1 + \cdots + \alpha_k w_k) = \alpha_1 [w_1]_{\mathcal{B}} + \cdots + \alpha_k [w_k]_{\mathcal{B}} = 0_{\mathbb{K}^n}.$$

Poiché $\mathcal{F}_{\mathcal{B}}$ è biunivoca, si ha

$$\alpha_1 w_1 + \cdots + \alpha_k w_k = 0_V.$$

Quindi i vettori $w_1, \dots, w_k \in V$ sono linearmente dipendenti. In maniera analoga è possibile dimostrare lo stesso risultato per la lineare indipendenza.

Sia $v \in V$. Il vettore v è combinazione lineare dei vettori $w_1, \dots, w_k \in V$ se esistono $\beta_1, \dots, \beta_k \in \mathbb{K}$ tale che

$$v = \beta_1 w_1 + \cdots + \beta_k w_k.$$

Poiché $\mathcal{F}_{\mathcal{B}}$ è biunivoca, $v = \beta_1 w_1 + \cdots + \beta_k w_k$ se e solamente se $\mathcal{F}_{\mathcal{B}}(v) = \mathcal{F}_{\mathcal{B}}(\beta_1 w_1 + \cdots + \beta_k w_k)$, quindi se e solamente se

$$[v]_{\mathcal{B}} = \mathcal{F}_{\mathcal{B}}(v) = \mathcal{F}_{\mathcal{B}}(\beta_1 w_1 + \cdots + \beta_k w_k) = \beta_1 [w_1]_{\mathcal{B}} + \cdots + \beta_k [w_k]_{\mathcal{B}}.$$

Quindi $v \in V$ è combinazione lineare dei vettori $w_1, \dots, w_k$ se e solamente se $[v]_{\mathcal{B}}$ è combinazione lineare dei vettori $[w_1]_{\mathcal{B}}, \dots, [w_k]_{\mathcal{B}}$.

Sia $w \in \mathcal{L}(w_1, \dots, w_k)$. Esistono $\alpha_1, \dots, \alpha_k \in \mathbb{K}$ tali che $w = \alpha_1 w_1 + \cdots + \alpha_k w_k$. Applicando $\mathcal{F}_{\mathcal{B}}$ si ha

$$\begin{aligned} \mathcal{F}_{\mathcal{B}}(w) &= \mathcal{F}_{\mathcal{B}}(\alpha_1 w_1 + \cdots + \alpha_k w_k) \\ &= \alpha_1 \mathcal{F}_{\mathcal{B}}(w_1) + \cdots + \alpha_k \mathcal{F}_{\mathcal{B}}(w_k) \\ &= \alpha_1 [w_1]_{\mathcal{B}} + \cdots + \alpha_k [w_k]_{\mathcal{B}} \end{aligned}$$

e viceversa. Quindi, $\mathcal{F}_{\mathcal{B}}(\mathcal{L}(w_1, \dots, w_k)) = \mathcal{L}([w_1]_{\mathcal{B}}, \dots, [w_k]_{\mathcal{B}})$. □

Esempio 5.7. *Vogliamo stabilire se i vettori* $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 0 & -3 \\ 1 & 2 \end{bmatrix} \in$ $\mathrm{M}_{2\times 2}(\mathbb{R})$ *sono linearmente indipendenti. Abbiamo dimostrato che*

$$\mathcal{B} = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$$

è una base di $\mathrm{M}_{2\times 2}(\mathbb{R})$ *e*

$$\begin{bmatrix} x & y \\ z & t \end{bmatrix}_{\mathcal{B}} = \begin{bmatrix} x \\ y \\ z \\ t \end{bmatrix}.$$

Quindi stibilire se $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 0 & -3 \\ 1 & 2 \end{bmatrix}$ *sono linearmente indipendent è equivalente a stabilire se i vettori* $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}_{\mathcal{B}} = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix}_{\mathcal{B}} = \begin{bmatrix} 1 \\ 2 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 & -3 \\ 1 & 2 \end{bmatrix}_{\mathcal{B}} = \begin{bmatrix} 0 \\ -3 \\ 1 \\ 2 \end{bmatrix} \in \mathbb{R}^4$ *sono linearmente indipendenti. Poiché la matrice*

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

ha rango 3*, verificare!, si ha che i vettori* $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 0 & -3 \\ 1 & 2 \end{bmatrix}$ *sono linearmente indipendent*

5.7.1 Dimensione di uno spazio vettoriale

Sia V uno spazio finitamente generato. L'obiettivo di questa sezione è dimostrare che due basi hanno lo stesso numero di elementi. Cominciamo con il seguente Lemma.

Lemma 5.8. *Siano* $v_1, \ldots, v_n$ *vettori di* V*. I vettori* $v_1, \ldots, v_n$ *sono linearmente dipendenti se e solamente se uno di essi si scrive come combinazione* *lineare degli altri, ovvero, esiste* $1 \leq j \leq n$ *tale che*

$$v_j \in \mathcal{L}(v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_n).$$

Inoltre

$$\mathcal{L}(v_1, \ldots, v_n) = \mathcal{L}(v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_n).$$

Dimostrazione. Supponiamo che $v_1, \ldots, v_n$ siano linearmente dipendenti. Esistono $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$ non tutti nulli, tali che $\lambda_1 v_1 + \cdots + \lambda_n v_n = 0$. Supponiamo che $\lambda_j \neq 0$. Allora dall'equazione

$$\lambda_1 v_1 + \cdots \underbrace{\lambda_j v_j} + \cdots + \lambda_n v_n = 0$$

posso portare al secondo membro tutti i termini, tranne $\lambda_j v_j$, e poi dividire per λ_j, ottenedo

$$v_j = (-\lambda_j^{-1}\lambda_1)v_1 + \cdots + (-\lambda_j^{-1}\lambda_{j-1})v_{j-1} + (-\lambda_j^{-1}\lambda_{j+1})v_{j+1} + \cdots + (-\lambda_j^{-1}\lambda_n)v_n.$$

Quindi v_j è combinazione lineare dei rimanenti, ovvero

$$v_j \in \mathcal{L}(v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_n).$$

Viceversa supponiamo che esista $j \in \{1, \ldots, n\}$, tale che v_j è combinazione lineare dei rimanenti, ovvero

$$v_j = \lambda_1 v_1 + \cdots + \lambda_{j-1} v_{j-1} + \lambda_{j+1} v_{j+1} + \cdots + \lambda_n v_n.$$

Allora

$$\lambda_1 v_1 + \cdots + \lambda_{j-1} v_{j-1} - v_j + \lambda_{j+1} v_{j+1} + \cdots + \lambda_n v_n = 0,$$

e quindi i vettori sono linearmente dipendenti poiché il coefficiente che moltiplica v_j è -1.

Adesso, proviamo che $\mathcal{L}(v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_n) = \mathcal{L}(v_1, \ldots, v_n)$ se v_j è combinazione lineare dei vettori $v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_n$.

Poiché $\mathcal{L}(v_1, \ldots, v_n)$ è un sottospazio vettoriale di V e i vettori

$$v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_n \in \mathcal{L}(v_1, \ldots, v_n),$$

si ha $\mathcal{L}(v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_n) \subseteq \mathcal{L}(v_1, \ldots, v_n)$.

Viceversa, poiché v_j è combinazione lineare di $v_1, \ldots, v_{j-1}, v_{j+1}, \ldots v_n$, si ha

$$v_j = \lambda_1 v_1 + \cdots + \lambda_{j-1} v_{j-1} + \lambda_{j+1} v_{j+1} + \cdots + \lambda_n v_n$$

ovvero

$$v_j \in \mathcal{L}(v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_n).$$

Quindi $v_1, \ldots, v_n \in \mathcal{L}(v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_n)$ da cui segue $\mathcal{L}(v_1, \ldots, v_n) \subseteq \mathcal{L}(v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_n)$ concludendo la dimostrazione. $\square$

Proposizione 5.9. *Sia V uno spazio vettoriale finitamente generato. Allora V ammette una base.*

Facoltativa. Siano $v_1, \ldots, v_n$ un sistema di generatori dello spazio vettoriale V, i.e., $V = \mathcal{L}(v_1, \ldots, v_n)$. Se $v_1, \ldots, v_n$ sono linearmente indipendenti allora formano una base. Altrimenti i vettori $v_1, \ldots, v_n$ sono linearmente dipendenti. Applicando il Lemma 5.8, esiste $1 \leq j_1 \leq n$ tale che v_{j_1} è combinazione lineare dei rimanenti. Inoltre

$$V = \mathcal{L}(v_1, \ldots, v_n) = \mathcal{L}(v_1, \ldots, v_{j_1-1}, v_{j_1+1}, \ldots, v_n),$$

ovvero $v_1, \ldots, v_{j_1-1}, v_{j_1+1}, \ldots, v_n$ formano un sistema di generatori. Se i vettori $v_1, \ldots, v_{j_1-1}, v_{j_1+1}, \ldots, v_n$ non fossero linearmente indipendenti allora potrei iterare questo procedimento, al massimo un numero finito di volte, fino ad ottenere una base di V. $\square$

Vogliamo dimostrare che due basi hanno lo stesso numero di vettori. Cominciamo con il seguente risultato.

Lemma 5.10 (Steinitz)**.** *Sia V uno spazio vettoriale su $\mathbb{K}$ generato da n vettori $v_1, \ldots, v_n$. Siano $w_1, \ldots, w_m \in V$ con $m > n$. Allora $w_1, \ldots, w_m$ sono vettori linearmente dipendenti.*

Dimostrazione. Vogliamo dimostrare che esistono $\alpha_1, \ldots, \alpha_m$ non tutti nulli tali che $\alpha_1 w_1 + \cdots + \alpha_m w_m = 0_V$. Poiché $V = \mathcal{L}(v_1, \ldots, v_n)$, ogni vettore w_j, per $j = 1, \ldots, m$, si scrive come combinazione lineare dei vettori $v_1, \ldots, v_n$:

$$w_j = a_{1j}v_1 + \cdots + a_{nj}v_n = \sum_{k=1}^{n} a_{kj}v_k.$$

Quindi

$$\begin{aligned}
0 &= \alpha_1 w_1 + \cdots + \alpha_m w_m \\
&= \sum_{j=1}^{m} \alpha_j w_j \\
&= \sum_{j=1}^{m} \sum_{k=1}^{n} \alpha_j a_{kj} v_k \\
&= \sum_{k=1}^{n} \left(\sum_{j=1}^{m} a_{kj}\alpha_j \right) v_k.
\end{aligned}$$

Consideriamo il seguente sistema lineare:

$$\begin{cases} \sum_{j=1}^m a_{1j}\alpha_j = 0 \\ \vdots \\ \sum_{j=1}^m a_{nj}\alpha_j = 0 \end{cases}$$

È un sistema lineare omogeneo di n equazioni in m incognite, con $n < m$. Per il corollario 4.67 (del Teorema di Rouché-Capelli) il sistema ammette soluzioni non banali. Quindi esistono $\alpha_1, \dots \alpha_m \in \mathbb{K}$ non tutti nulli tali che

$$\begin{cases} \sum_{j=1}^m a_{1j}\alpha_j = 0 \\ \vdots \\ \sum_{j=1}^m a_{nj}\alpha_j = 0 \end{cases}$$

Adesso,

$$\begin{aligned} \sum_{j=1}^m \alpha_j w_j &= \sum_{j=1}^m \alpha_j \left(\sum_{k=1}^n a_{kj} v_k \right) \\ &= \sum_{k=1}^n \left(\sum_{j=1}^m a_{kj}\alpha_j \right) v_k \\ &= 0_V, \end{aligned}$$

ovvero i vettori $\mathcal{C} = \{w_1, \dots, w_m\}$ sono linearmente dipendenti. $\square$

Corollario 5.11. *Sia V uno spazio vettoriale su $\mathbb{K}$ generato da n vettori. Siano $w_1, \dots, w_m \in V$ linearmente indipendenti. Allora $m \leq n$.*

Dimostrazione. Se $m > n$, allora per il Lemma di Steinitz sarebbero linearmente dipendenti. Assurdo. Quindi $m \leq n$. $\square$

Corollario 5.12. *Siano $\mathcal{B} = \{v_1, \dots, v_n\}$, $\mathcal{C} = \{w_1, \dots, w_m\}$ due basi di V. Allora $m = n$.*

Dimostrazione. Poiché $V = \mathcal{L}(v_1, \dots, v_n)$ e $w_1, \dots, w_m$ sono linearmente indipendenti, applicando il corollario anteriore ottengo che $m \leq n$. Scambiando il ruolo delle due basi, i.e., $V = \mathcal{L}(w_1, \dots, w_m)$ e $v_1, \dots, v_n$ linearmente indipendenti, sempre per il corollario anteriore si ha $n \leq m$. Quindi $n = m$. $\square$

Definizione 5.13. *Sia V uno spazio vettoriale generato da un numero finito di elementi. Il numero dei vettori di una qualsiasi base di V si dice dimensione di V.*

Osservazione 5.14. *Sia V uno spazio vettoriale su $\mathbb{K}$ di dimensione n. Allora:*

- *n è il massimo numero di vettori di V linearmente indipendenti;*
- *se $v_1, \ldots, v_m$ formano un sistema di generatori di V, allora $n \leq m$.*

Infine, se $V = \{0_V\}$, ovvero uno spazio vettoriale formato solamente dal vettore nullo, la sua dimensione è per definizione 0.

Esempio 5.15.

a) *i vettori* $e_1 = \begin{bmatrix} 1 \\ \vdots \\ 0 \end{bmatrix}, \ldots, e_n = \begin{bmatrix} 0 \\ \vdots \\ 1 \end{bmatrix}$ *formano una base di* $\mathbb{R}^n$. *Quindi* $\dim \mathbb{R}^n = n$*;*

b) *i vettori* $e_1 = \begin{bmatrix} 1 \\ \vdots \\ 0 \end{bmatrix}, \ldots, e_n = \begin{bmatrix} 0 \\ \vdots \\ 1 \end{bmatrix}$ *formano una base di* $\mathbb{C}^n$. *Quindi* $\dim \mathbb{C}^n = n$*;*

c) *Sia $E_{ij} \in M_{m\times n}(\mathbb{K})$ la matrice i cui elementi sono tutti nulli tranne l'elemento $a_{ij} = 1$. È facile provare che $\mathcal{B} = \{E_{ij}, 1 \leq i \leq m, 1 \leq j \leq n\}$ è una base di $M_{m\times n}(\mathbb{K})$. Quindi* $\dim M_{m\times n}(\mathbb{K}) = mn$.

d) *i polinomi $\{1, x, \ldots, x^n\}$ formano una base di $\mathbb{K}_n[x]$. Quindi* $\dim \mathbb{K}_n[x] = n + 1$*;*

e) *Sia $V = \left\{A \in M_{2\times 2}(\mathbb{R}) :\ A = A^T\right\}$, i.e., l'insieme delle matrici simmetriche. Si può dimostrare che i vettori*

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix},$$

formano una base di V. Quindi $\dim V = 3$.

f) *Sia $V = \left\{A \in M_{3\times 3}(\mathbb{R}) :\ A = -A^T\right\}$, i.e., l'insieme delle matrici antisimmetriche. Si può dimostrare che le matrici*

$$\begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix},$$

formano una base di V. Quindi $\dim V = 3$.

Proposizione 5.16. *Sia V uno spazoio vettoriale su $\mathbb{K}$. Siano $w_1, \ldots, w_m \in V$ e sia $W = \mathcal{L}(w_1, \ldots, w_m)$. Indichiamo con k il massimo numero dei vettori $\{w_1, \ldots, w_m\}$ linearmente indipendenti. Allora* $\dim W = k$

Dimostrazione. Poiché k è il massimo numero dei vettori $\{w_1, \ldots, w_m\}$ linearmente indipendenti, esistono $1 \leq j_1 < \cdots < j_k \leq m$ tali che i vettori i vettori $w_{j_1}, \ldots, w_{j_k}$ sono linearmente indipendenti. Vogliamo dimostrare che $W = \mathcal{L}(w_{j_1}, \ldots, w_{j_k})$. Una inclusione è immediate, ovvero $\mathcal{L}(w_{j_1}, \ldots, w_{j_k}) \subseteq W$ (perché?). Per dimostrare che $W \subseteq \mathcal{L}(w_{j_1}, \ldots, w_{j_k})$ è sufficiente dimostrare che $w_j \in \mathcal{L}(w_{j_1}, \ldots, w_{j_k})$ per $j = 1, \ldots, n$.

Se $j = j_i$ per un certo $1 \leq i \leq k$, allora $w_j \in \mathcal{L}(w_{j_1}, \ldots, w_{j_k})$. Supponiamo che $j \neq j_i$. I vettori $w_j, w_{j_1}, \ldots, w_{j_k}$ sono linearmente dipendenti, poiché sono $k+1$ e k è il massimo numero dei vettori $\{w_1, \ldots, w_m\}$ linearmente indipendenti. Quindi esistono $\alpha_j, \alpha_{j_1}, \ldots, \alpha_{j_k} \in \mathbb{K}$ non tutti nulli tali che

$$\alpha_j w_j + \alpha_{j_1} w_{j_1} + \cdots + \alpha_{j_k} w_{j_k} = 0_V.$$

Se $\alpha_j = 0$, allora

$$\alpha_{j_1} w_{j_1} + \cdots + \alpha_{j_k} w_{j_k} = 0_V,$$

ovvero esisterebbe una combinazione lineare non banale dei vettori $w_{j_1}, \ldots, w_{j_k}$ uguale al vettore nullo. Assurdo perché i vettori $w_{j_1}, \ldots, w_{j_k}$ sono linearmente indipendenti. Quindi $\alpha_j \neq 0$ da cui segue che

$$w_j = (-\alpha_{j_1}/\alpha_j) w_{j_1} + \cdots + (-\alpha_{j_k}/\alpha_j) w_{j_k} \in \mathcal{L}(w_{j_1}, \ldots, w_{j_k}).$$

Quindi abbiamo dimostrato che

$$w_1, \ldots, w_n \in \mathcal{L}(w_{j_1}, \ldots, w_{j_k}),$$

ovvero

$$W = \mathcal{L}(w_1, \ldots, w_n) \subseteq \mathcal{L}(w_{j_1}, \ldots, w_{j_k}).$$

□

Corollario 5.17. *Siano $v_1, \ldots, v_n \in \mathbb{K}^m$ e sia $A = (v_1, \ldots, v_n) \in \mathrm{M}_{m \times n}(\mathbb{K})$. Allora* $\dim \mathcal{L}(v_1, \ldots, v_n) = \mathrm{rg}(A)$. *In particolare se $A = (A^1, \ldots, A^n) = \begin{bmatrix} A_1 \\ \vdots \\ A_m \end{bmatrix} \in \mathrm{M}_{m \times n}(\mathbb{K})$, allora $rg(A) = \dim (A^1, \ldots, A^n) = \dim(A_1^T, \ldots, A_m^T)$. Infine il rango di una matrice non cambia se effettuo operazioni elementari di colonna, rispettivamente di righe, sulla matrice A.*

Facoltativa. Nella proposizione anteriore abbiamo dimostrato che la dimensione del sottospazio $\mathcal{L}(v_1, \ldots, v_n)$ è il massimo numero dei vettori $\{v_1, \ldots, v_n\}$ linearmente indipendenti.

Sia $A = (v_1, \ldots, v_n) \in \mathrm{M}_{m \times n}(\mathbb{K})$. Per definizione di rango di una matrice si ha

$$\dim \mathcal{L}(v_1, \ldots, v_n) = \mathrm{rg}(A).$$

Adesso dimostreremo che il rango di una matrice non cambia se effettuo operazioni elementari di colonna. La dimostrazione che il rango di una matrice non cambia se effettuo operazioni elementari di riga è analogo.

Poiché $rg(A) = \dim(A^1, \ldots, A^n)$, il rango di una matrice non cambia se scambio due colonne oppure se moltiplico una colonna per un multiplo non nullo. Proviamo che il rango rimane invariato se sommiamo ad una colonna un multiplo di una altro. Dobbiamo dimostrare che per ogni $1 \leq i \neq j \leq n$, e per ogni $\lambda \in \mathbb{K}$, si ha $\mathcal{L}(A^1, \ldots, A^n) = \mathcal{L}(A_1, \ldots, A^i + \lambda A_j, \ldots, A_n)$.

Poiché

$$A^1, \ldots, A^i + \lambda A^j, \ldots, A^n \in \mathcal{L}(A^1, \ldots, A^n),$$

ne segue che

$$\mathcal{L}(A^1, \ldots, A_i + \lambda A_j, \ldots, A_n) \subseteq \mathcal{L}(A^1, \ldots, A^n).$$

Viceversa, $A^1, \ldots, \widehat{A^i}, \ldots, A^n \in \mathcal{L}(A^1, \ldots, A^i + \lambda A^j, \ldots, A^n)$. Inoltre $A^i = (A^i + \lambda A^j) - \lambda A^j \in \mathcal{L}(A^1, \ldots, A^i + \lambda A^j, \ldots, A^n)$ da cui segue che

$$\mathcal{L}(A^1, \ldots, A^n) \subseteq \mathcal{L}(A^1, \ldots, A^i + \lambda A^j, \ldots, A^n).$$

□

Proposizione 5.18. *Sia V uno spazio vettoriale di dimensione n. Sia $\mathcal{B} = \{v_1, \ldots, v_n\}$ un insieme formata da n vettori di V. Le seguenti condizioni sono equivalenti:*

a) $\mathcal{B} = \{v_1, \ldots, v_n\}$ è costituito da vettori linearmente indipendenti;

b) $\mathcal{B} = \{v_1, \ldots, v_n\}$ formano un sistema di generatori;

c) $\mathcal{B} = \{v_1, \ldots, v_n\}$ è una base di V.

Dimostrazione. $(a) \Rightarrow (b)$. Sia $v \in V$. Devo dimostrare che $v \in \mathcal{L}(v_1, \ldots, v_n)$, ovvero che v è combinazione lineare di $v_1, \ldots, v_n$.

Poiché n è il massimo numero di vettori linearmente indipendenti di V, allora $v, v_1, \ldots, v_n$ sono linearmente dipendenti. Quindi esistono $\alpha, \lambda_1, \ldots, \lambda_n$ non tutti nulli tali che

$$\alpha v + \lambda_1 v_1 + \cdots + \lambda_n v_n = 0.$$

Se $\alpha = 0$, allora i vettori $v_1, \ldots, v_n$ sarebbero linearmente dipendenti. Quindi, $\alpha \neq 0$ e $v = -\frac{\lambda_1}{\alpha} v_1 - \cdots - \frac{\lambda_n}{\alpha} v_n$, ovvero $v \in \mathcal{L}(v_1, \ldots, v_n)$.

$(b) \Rightarrow (c)$. Poiché $v_1, \ldots, v_n$ formano un sistema di generatori, allora $V = \mathcal{L}(v_1, \ldots, v_n)$. Se $v_1, \ldots, v_n$ fossero linearmente dipendenti, esisterebbe $1 \leq j \leq n$ tale che v_j è combinazione linare dei rimanenti. Inoltre $V = \mathcal{L}(v_1, \ldots, v_n) = \mathcal{L}(v_1, \ldots v_{j-1}, v_{j+1}, \ldots, v_n)$. Applicando il Lemma di Steinitz si ha che una base di V sarebbe formata da al massimo $n-1$ elementi. Assurdo. $(c) \Rightarrow (a)$ è immediata dalla definizione di base. $\square$

Una base è un insieme formato da vettori linearmente indipendenti e generatori. Se $\dim V = n$ e $\mathcal{C} = \{w_1, \ldots, w_m\}$ è un insieme di vettori linearmente indipendenti, quindi $m \leq n$, è possibile completare $\mathcal{C}$ a base di V?La risposta è si. Cominciamo con il seguente Lemma.

Lemma 5.19. *Siano $v_1, \ldots v_n$ vettori linearmente indipedenti di V e sia $v \in V$. I vettori $v, v_1, \ldots v_n$ sono linearmente indipendenti se e solamente se $v \notin \mathcal{L}(v_1, \ldots, v_n)$.*

Dimostrazione. Supponiamo che $v, v_1, \ldots v_n$ sono linearmente indipendenti. Per il Lemma 5.8 v non appartiene a $\mathcal{L}(v_1, \ldots, v_n)$ poiché altrimenti i vettori $v, v_1, \ldots, v_n$ sarebbero linearmente dipendenti. Viceversa, supponiamo che v non appartenga a $\mathcal{L}(v_1, \ldots, v_n)$ e sia $\alpha v + \beta_1 v_1 + \cdots + \beta_n v_n = 0$ una combinazione lineare uguale al vettore nullo. Se α fosse differente da zero, allora

$$v = -(\beta_1/\alpha) v_1 - \cdots - (\beta_n/\alpha) v_n,$$

ovvero $v \in \mathcal{L}(v_1, \ldots, v_n)$ che non è possibile per ipotesi. Quindi necessariamente $\alpha = 0$. Se $\alpha = 0$, allora necessariamente $\beta_1 = \cdots = \beta_n = 0$ poiché i vettori $v_1, \ldots, v_n$ sono linearmente indipendenti. $\square$

Teorema 5.20 (Teorema di completamento a base)**.** *Sia V uno spazio vettoriale su $\mathbb{K}$ di dimensione n. Sia $\mathcal{C} = \{w_1, \ldots, w_m\}$ un insieme formato da vettori linearmente indipendenti e sia $\mathcal{B} = \{v_1, \ldots, v_n\}$ una base di V. Allora esistono $n-m$ vettori di $\mathcal{B}$ che insieme a $\mathcal{C} = \{w_1, \ldots, w_m\}$ formano una base di V.*

Facoltativa. Poiché $n = \dim V$ è il massimo numero di vettori linearmente indipendenti allora $m \leq n$. Poiché $\mathcal{B} = \{v_1, \ldots, v_n\}$ è una base di V, si ha $\mathcal{L}(v_1, \ldots, v_n) = V$ da cui segue che

$$\mathcal{L}(w_1, \ldots, w_m) \subseteq \mathcal{L}(v_1, \ldots, v_n).$$

Se $\mathcal{L}(w_1, \ldots, w_m) = \mathcal{L}(v_1, \ldots v_n)$, allora $\{w_1, \ldots, w_m\}$ è una base di V. Quindi $m = n$ e la dimostrazione è finita. Altrimenti $\mathcal{L}(w_1, \ldots, w_m) \subsetneq \mathcal{L}(v_1, \ldots v_n)$. Affermiamo che esiste $1 \leq j_1 \leq n$ tale che $v_{j_1} \notin \mathcal{L}(w_1, \ldots, w_m)$. Infatti, se $v_1, \ldots, v_n \in \mathcal{L}(w_1, \ldots, w_m)$, allora $\mathcal{L}(v_1, \ldots, v_n) \subseteq \mathcal{L}(w_1, \ldots, w_m)$ e quindi

$$\mathcal{L}(w_1, \ldots, w_m) = \mathcal{L}(v_1, \ldots, v_n).$$

Assurdo perché per ipotesi $\mathcal{L}(w_1, \ldots, w_m) \subsetneq \mathcal{L}(v_1, \ldots v_n)$. Applicando il Lemma 5.19, i vettori $v_{j_1}, w_1, \ldots, w_m$ sono linearmente indipendenti. Se

$$\mathcal{L}(v_{j_1}, w_1, \ldots, w_m) = \mathcal{L}(v_1, \ldots, v_n),$$

allora $m + 1 = n$ ed il Teorema sarebbe dimostrato. Altrimenti esisterebbe $1 \leq j_2 \leq n$, $j_2 \neq j_1$, tale che v_{j_2} non apparterrebbe al sottospazio $\mathcal{L}(v_{j_1}, w_1 \ldots, w_m)$ e quindi i vettori $v_{j_2}, v_{j_1}, w_1, \ldots w_m$ sarebbero linearmente indipendenti. Se

$$\mathcal{L}(v_{j_1}, v_{j_2}, w_1, \ldots, w_m) = \mathcal{L}(v_1, \ldots, v_n),$$

allora $m+2 = n$ ed il Teorema sarebbe dimostrato. Altrimenti posso iterare questo procedimento per un numero finito di volte: esattamente $n-m$ volte, concludendo la dimostrazione del Teorema. $\square$

Corollario 5.21. *Sia V uno spazio vettoriale su $\mathbb{K}$ e sia W un sottospazio di V. Allora* $\dim W \leq \dim V$. *Inoltre* $\dim W = \dim V$ *se e solamente se* $V = W$.

Dimostrazione. Sia $\mathcal{C} = \{w_1, \ldots, w_m\}$ una base di W e sia $\mathcal{B} = \{v_1, \ldots, v_n\}$ una base di V. Poiché i vettori $w_1, \ldots, w_m$ sono linearmente indipendenti, applicando il Lemma di Steinitz si ha $m \leq n$, ovvero $\dim W \leq \dim V$.

Se $\dim W = \dim V$, $m = n$ e $\mathcal{C} = \{w_1, \ldots, w_m\}$ è anche una base di V, da cui segue che $V = \mathcal{L}(w_1, \ldots, w_m) = W$. $\square$

Diamo una prova alternativa della I parte del Teorema di Rouché-Capelli utilizzando i risultati anteriori.

Corollario 5.22. *Sia $A \in \mathrm{M}_{m \times n}(\mathbb{K})$. Il sistema $AX = b$ è compatibile se e solamente se* $\mathrm{rg}(A) = \mathrm{rg}(A|b)$.

Facoltativa. Abbiamo visto che Il sistema lineare $AX = b$ è compatibile se e solamente se $b \in \mathcal{L}(A^1, \ldots, A^n)$. Poiché $\mathcal{L}(A^1, \ldots, A^n) \subseteq \mathcal{L}(A^1, \ldots, A^n, b)$, si ha $b \in \mathcal{L}(A^1, \ldots, A^n)$ se e solamente se $\mathcal{L}(A^1, \ldots, A^n) = \mathcal{L}(A^1, \ldots, A^n, b)$. Riassumendo, abbiamo dimostrato che il sistema $AX = b$ è compatibile se e solamente se $\mathcal{L}(A^1, \ldots, A^n) = \mathcal{L}(A^1, \ldots, A^n, b)$ Applicando il corollario anteriore, si ha che il sistema lineare $AX = b$ è combatibile se e solamente se $\dim \mathcal{L}(A^1, \ldots, A^n) = \dim \mathcal{L}(A^1, \ldots, A^n, b)$. Applicando il corollario 5.17 si ha che il sistema $AX = b$ è compatibile se e solamente se $\mathrm{rg}(A) = \mathrm{rg}(A|b)$. $\square$

5.22.1 Equazioni cartesiane di un sottospazio di $\mathbb{K}^n$

Sia $W \subset \mathbb{K}^n$ un sottospazio vettoriale generato dai vettori $w_1, \ldots, w_k$, ovvero $W = \mathcal{L}(w_1, \ldots, w_k)$. Il vetore $w \in W$ se e solamente se esistono degli scalari $\alpha_1, \ldots, \alpha_k \in \mathbb{K}$ tali che

$$\alpha_1 w_1 + \cdots + \alpha_k w_k = w,$$

ovvero se e solamente se il sistema lineare

$$AX = w,$$

è compatibile, dove $A = (w_1, \ldots, w_k) \in M_{n\times k}(\mathbb{K})$. Applicando il metodo di Gauss alla matrice $(A|w)$ ottengo un sistema lineare $SX = c$ ridotto a scala e equivalente al sistema lineare $AX = w$. Il sistema lineare $SX = c$ ha la forma:

$$\left[\begin{array}{ccccccccccc|c}
\cdots & s_{1j_1} & * & * & \cdots & \cdots & * & \cdots & \cdots & \cdot & * & c_1 \\
\cdots & \cdots & 0 & s_{2j_2} & * & * & * & \cdots & \cdots & \cdot & * & c_2 \\
\cdots & \cdots & \cdots & \cdots & 0 & s_{3j_3} & * & * & \cdots & \cdots & * & c_3 \\
\vdots & \cdots & \cdots & \cdots & \cdots & 0 & \ddots & \cdots & \cdots & \cdots & * & \vdots \\
0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & s_{(r-1)j_{r-1}} & * & * & * & c_{r-1} \\
0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & s_{rj_r} & * & * & c_r \\
\hline
0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & \cdots & 0 & c_{r+1} \\
\vdots & & & & & & & & \vdots & \vdots & & \vdots \\
0 & \cdots & & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & \cdots & 0 & c_n
\end{array}\right]$$

Quindi il sistema $SX = c$ è compatibile se e solamente se le ultime $n - rg(S)$ coordinate di c sono nulle, ovvero se e solamente se

$$c_{r+1} = 0, \ldots, c_n = 0,$$

le quali definiscono equazioni cartesiane. Quindi,

$$W = \mathcal{L}(w_1, \ldots, w_k) = \mathrm{Sol}(A|w) = \mathrm{Sol}(S|c) = \{c_{r+1} = 0, \ldots, c_n = 0\}\,.$$

Queste equazioni sono chiamate *equazioni cartesiane* di W.

Esempio 5.23. *Sia* $W = \mathcal{L}\left(\begin{bmatrix}1\\0\\1\\1\end{bmatrix}, \begin{bmatrix}0\\1\\1\\-1\end{bmatrix}, \begin{bmatrix}1\\1\\2\\0\end{bmatrix}\right)$.

$$\left[\begin{array}{ccc|c} 1 & 0 & 1 & x_1 \\ 0 & 1 & 1 & x_2 \\ 1 & 1 & 2 & x_3 \\ 1 & -1 & 0 & x_4 \end{array}\right] \begin{array}{c} A_3 - A_1 \\ \longrightarrow \\ A_4 - A_1 \\ \longrightarrow \end{array} \left[\begin{array}{ccc|c} 1 & 0 & 1 & x_1 \\ 0 & 1 & 1 & x_2 \\ 0 & 1 & 1 & x_3 - x_1 \\ 0 & -1 & -1 & x_4 - x_1 \end{array}\right]$$

$$\left[\begin{array}{c|cc|c} 1 & 0 & 1 & x_1 \\ \hline 0 & 1 & 1 & x_2 \\ 0 & 1 & 1 & x_3 - x_1 \\ 0 & -1 & -1 & x_4 - x_1 \end{array}\right] \begin{array}{c} A_3 - A_2 \\ \longrightarrow \\ A_4 + A_2 \\ \longrightarrow \end{array} \left[\begin{array}{ccc|c} 1 & 0 & 1 & x_1 \\ 0 & 1 & 1 & x_2 \\ \hline 0 & 0 & 0 & x_3 - x_1 - x_2 \\ 0 & 0 & 0 & x_4 - x_1 + x_2 \end{array}\right]$$

Il sistema lineare è compatibile se e solamente se $x_3 - x_1 - x_2 = 0$ *e* $x_4 - x_1 + x_2 = 0$, *ovvero*

$$W = \left\{ \begin{bmatrix}x_1\\x_2\\x_3\\x_4\end{bmatrix} \in \mathbb{R}^4 :\ x_3 - x_1 - x_2 = 0, x_4 - x_1 + x_2 = 0 \right\}.$$

5.23.1 Metodi di Calcolo

Siano $X_1, \ldots, X_k$ vettori di $\mathbb{K}^n$, dove $\mathbb{K} = \mathbb{R}$ oppure $\mathbb{K} = \mathbb{C}$.

a) Affinché i vettori $X_1, \ldots, X_k$ formino una base è necessario che $n = k$ essendo $n = \dim \mathbb{K}^n$. Sia $(X_1, \ldots, X_n) \in \mathrm{M}_{n\times n}(\mathbb{K})$, ovvero la matrice le cui colonne sono i vettori $X_1, \ldots, X_n$. Applicando la Proposizione 5.18 si ha che i vettori $X_1, \ldots, X_n \in \mathbb{K}^n$ formano una base se e solamente se $X_1, \ldots, X_n$ sono linearmente indipendenti, quindi se e solamente se $\det(X_1, \ldots, X_n) \neq 0$ ($\mathrm{rg}(X_1, \ldots, X_n) = n$);

b) Sia $W = \mathcal{L}(X_1, \ldots, X_k)$ dove $X_1, \ldots X_k \in \mathbb{K}^n$. Abbiamo visto che $\dim W = rg((X_1, \ldots, X_k))$. Per determinare una base possiamo seguire il seguente procedimento.

Applicando l'algoritmo di Gauss alla matrice $A = (X_1, \ldots, X_k) \in M_{m\times k}(\mathbb{K})$ otteniamo una matrice $S \in M_{m\times k}(\mathbb{K})$ ridotta a scala. Siano $1 \leq j_1 < \ldots < j_k \leq n$ tali che le colonne $S^{j_1}, \ldots, S^{j_k}$ corrispondono ai perni di S. Si può dimostrare che $X_{j_1}, \ldots, X_{j_k}$ sono linearmente indipendenti. Poiché $k = \mathrm{rg}(A)$, allora $(X_{j_1}, \ldots, X_{j_k})$ formano una base di W.

c) Siano $X_1, \ldots, X_k \in \mathbb{K}^n$ linearmente indipendenti. Abbiamo dimostrato che possiamo completarli a base di $\mathbb{K}^n$. Una possibilità è di trovare $n-k$ vettori $Y_{k+1}, \ldots, Y_n \in \mathbb{K}^n$ tali che

$$\det(X_1, \ldots, X_k, Y_{k+1}, \ldots, Y_n) \neq 0.$$

Non abbiamo un algoritmo che ci guida nella scelta dei vettori da aggiungere. Tuttavia, il Teorema di completamento a base afferma che possiamo completare a base i vettori indipendenti $X_1, \ldots, X_k$ aggiungendo, per esempio, $n-k$ vettori della base canonica.

Il metodo di Gauss, ci fornisce un algoritmo per completare a base i vettori $X_1, \ldots, X_k$ linermente indipendenti.

Sia $\mathcal{C} = \{e_1, \ldots e_n\}$ la base canonica di $\mathbb{K}^n$ e sia

$$A = (X_1, \ldots X_k, e_1, \ldots, e_n) \in M_{n\times(n+k)}(\mathbb{K}).$$

Si osservi che $rg(A) = n$. Infatti $rg(A) \leq n$ poiché la matrice A ha n righe. Dall'altro lato, i vettori $e_1, \ldots, e_n$ sono linearmente indipendenti e quindi esistono almento n colonne linearmente indipendenti, ovvero $rg(A) \geq n$. Quindi $rg(A) = n$. Applicando l'algoritmo di Gauss alla matrice A otteniamo una matrice ridotta a scala che indichiamo con $S = (S^1, \ldots, S^k, B^1, \ldots, B^n)$. Poiché i vettori $X_1, \ldots, X_k$ sono linearmente indipendenti, si può dimostrare che le prime k colonne della matrice S contengono perni (perché?). Inoltre, se i perni della matrice S sono contenuti nelle colonne $S^1, \ldots, S^k, B^{j_1}, \ldots, B^{j_{n-k}}$, allora i vettori $X_1, \ldots X_k, e_{j_1}, \ldots e_{j_{n-k}}$ sono linearmente indipenti e quindi formano un base di $\mathbb{K}^n$. Questo procedimento può essere lungo poiché se dovessi completare a base $X_1, X_2 \in \mathbb{R}^4$, dovrei ridurre a scala una matrice di formato 4×6.

d) sia $\mathcal{B} = \{v_1, \ldots, v_n\}$ una base di $\mathbb{K}^n$. Vogliamo calcolare le coordinate di un vettore v rispetto alla base $\mathcal{B}$. Le coordinate di un vettore v rispetto alla base $\mathcal{B}$ sono gli unici scalari $x_1, \ldots, x_n \in \mathbb{K}$ tale che

$$v = x_1 v_1 + \cdots + x_n v_n.$$

Sia $A = (v_1, \ldots, v_n) \in M_{n\times n}(\mathbb{K})$ e sia $X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$. Poiché

$$v = x_1 v_1 + \cdots + x_n v_n \iff AX = v,$$

le coordinare di v rispetto alla base $\mathcal{B}$ è l'unica soluzione del sistema lineare $AX = v$. Quindi calcolare le coordinate di un vettore rispetto ad una base è equivalente a risolvere un sistema lineare di n equazioni in n incognite.

Sia V uno spazio vettoriale su $\mathbb{K}$ di dimensione n e sia $\mathcal{B} = \{v_1, \ldots, v_n\}$ una base di V. Sia $\mathcal{F}_\mathcal{B} : V \longrightarrow \mathbb{K}^n$ l'applicazione biunivoca che associa a v le sue coordinate rispetto a $\mathcal{B}$. Ricordiamo che dalla definizione di coordinate si ha che $\mathcal{F}_\mathcal{B}(v_i) = e_i$, per $i = 1, \ldots, n$, dove $\{e_1, \ldots, e_n\}$ è la base canonica di $\mathbb{K}^n$. Inoltre

$$\mathcal{F}_\mathcal{B}^{-1}\left(\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}\right) = x_1 v_1 + \cdots + x_n v_n.$$

Siano $w_1, \ldots, w_k \in V$. Allora:

a) $w_1, \ldots, w_k \in V$ sono linearmente indipendenti se e solamente se i vettori $[w_1]_\mathcal{B}, \ldots, [w_k]_\mathcal{B} \in \mathbb{K}^n$ sono linearmente indipendenti se e solamente se $\mathrm{rg}([w_1]_\mathcal{B}, \ldots, [w_k]_\mathcal{B}) = k$.

b) $w_1, \ldots, w_k \in V$ sono linearmente dipendenti se e solamente se i vettori $[w_1]_\mathcal{B}, \ldots, [w_k]_\mathcal{B} \in \mathbb{K}^n$ sono linearmente dipendenti se e solamente se $\mathrm{rg}([w_1]_\mathcal{B}, \ldots, [w_k]_\mathcal{B}) < k$.

c) $w_1, \ldots, w_k \in V$ formano una base di V se e solamente se $k = n$ ed i vettori $[w_1]_\mathcal{B}, \ldots, [w_n]_\mathcal{B}$ formano una base di $\mathbb{K}^n$ se e solamente se $\det([w_1]_\mathcal{B}, \ldots, [w_n]_\mathcal{B}) \neq 0$;

d) siano $v, w_1, \ldots, w_k \in V$. v è combinazione lineare di $w_1, \ldots, w_k$ se e solamente se $[v]_\mathcal{B}$ è combinazione lineare dei vettori $[w_1]_\mathcal{B}, \ldots, [w_k]_\mathcal{B} \in \mathbb{K}^n$ se e solamente se $\mathrm{rg}([w_1]_\mathcal{B}, \ldots, [w_k]_\mathcal{B}) = \mathrm{rg}([w_1]_\mathcal{B}, \ldots, [w_k]_\mathcal{B}, [v]_\mathcal{B})$;

Siano $w_1, \ldots, w_k$ vettori linearmente indipendenti di V. Possiamo completarli a base di V. Una maniera può essere la seguente.

Siano $[w_1]_\mathcal{B}, \ldots, [w_k]_\mathcal{B} \in \mathbb{K}^n$. Poiché i vettori $[w_1]_\mathcal{B}, \ldots, [w_k]_\mathcal{B}$ sono linearmente indipendenti applicando il Teorema del completamento a base, li posso completare a base aggiungendo $n-k$ vettori della base canonica $\{e_1,\ldots,e_n\}$. Quindi esistono $\leq j_1 < \ldots < j_{n-k} \leq n$ tale che

$$[w_1],\ldots,[w_k]_{\mathcal{B}}, e_{j_1},\ldots,e_{j_{n-k}},$$

formano una base di $\mathbb{K}^n$. Tenendo in mente che $\mathcal{F}_{\mathcal{B}}(v_j) = e_j$ per $j=1,\ldots,n$, e che $\mathcal{F}_{\mathcal{B}}$ è biunivoca, non è difficile dimostrare che i vettori

$$w_1,\ldots,w_k, v_{j_1},\ldots,v_{j_{n-k}},$$

formano una base di V.

Sia $W = \mathcal{L}(w_1,\ldots,w_k) \subseteq V$. Per calcolare una base possiamo seguire il seguente procedimento.

Sia $W' = \mathcal{L}([w_1]_{\mathcal{B}},\ldots,[w_k]_{\mathcal{B}})$. Noi sappiamo che $\dim W = \dim W' = \mathrm{rg}([w_1]_{\mathcal{B}},\ldots,[w_k]_{\mathcal{B}})$. Siano $1 \leq j_1 < \cdots < j_k \leq n$ tale che i vettori $[w_{j_1}]_{\mathcal{B}},\ldots,[w_{j_k}]_{\mathcal{B}})$ formano una base di W'. Non è difficile dimostrare che i vettori $w_{j_1},\ldots,w_{j_k}$ formano una base di W.

Sia $\mathcal{C} = \{w_1,\ldots,w_n\}$ una base di V. Le coordinate di un vettore v rispetto alla base $\mathcal{C}$ sono gli unici scalari $y_1,\ldots,y_n \in \mathbb{K}$ tali che

$$v = y_1 w_1 + \cdots + y_n w_n.$$

Poiché $\mathcal{F}_{\mathcal{B}}$ è biunivoca e trasforma combinazioni lineari in combinazioni lineari si ha

$$[v]_{\mathcal{B}} = y_1[w_1]_{\mathcal{B}} + \cdots + y_n[w_n]_{\mathcal{B}}.$$

Tenendo in mente che $\mathcal{C}' = \{[w_1]_{\mathcal{B}},\ldots,[w_n]_{\mathcal{B}}\}$ è una base di $\mathbb{K}^n$, si ha che le coordinate di v rispetto alla base $\mathcal{C}$ sono le coordinate del vettore $[v]_{\mathcal{B}}$ rispetto alla base $\mathcal{C}'$.

5.24 Formula di Grassman

Sia V uno spazio vettoriale su $\mathbb{K}$ e siano $U, W \subseteq V$ sottospazi vettoriali di V. Non è difficile verificare che $U \cap W$ è un sottospazio vettoriale di V (esercizio). Invece, in generale, $U \cup W$ non è un sottospazio vettoriale. Per esempio, sia $V = \mathbb{R}^2$ e sia $U = \mathcal{L}(e_1)$ and $W = \mathcal{L}(e_2)$. Allora $U \cup W$,

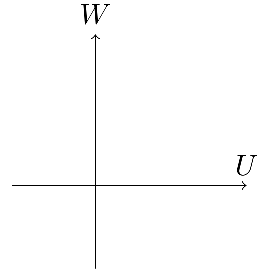


Figura 5.24.1.

è l'unione degli assi, il quale non è un sottospazio vettoriale poiché non è chiuso rispetto alla somma (ma è chiuso rispetto alla moltiplicazione per scalare).

Definiamo

$$U + W := \{u + w : u \in U \; w \in W\}.$$

Proposizione 5.25. *$U + W$ è un sottospazio vettoriale di V.*

Dimostrazione. Dobbiamo dimostrare che $U+W$ è chiuso rispetto alla somma e la moltiplicazione per scalare. Ovvero, se $z, h \in U+W$ e $\lambda \in \mathbb{K}$, allora $z + h \in U + W$ e $\lambda z \in U + W$.

Siano $z, h \in U + W$. Per definizione di $U + W$, esitono $u_1, u_2 \in U$ e $w_1, w_2 \in W$ tale che $z = u_1 + w_1$ e $h = u_2 + w_2$ rispettivamente. La tesi è che anche $z + h$ e λz li posso scrivere come una somma di un elemento di U e di un elemento di W. Infatti,

$$z + h = u_1 + w_1 + u_2 + w_2 = \underbrace{u_1 + u_2}_{U} + \underbrace{w_1 + w_2}_{W} \in U + W.$$

Analogamente

$$\lambda z = \lambda u_1 + \lambda w_1 \in U + W,$$

poiché $\lambda u_1 \in U$ e $\lambda w_1 \in W$ □

È facile verificare che $U + W$ contiene sia U e sia W. Se volessimo essere rigorosi dovremmo affermare che $U + W$ è il più piccolo sottospazio vettoriale di V che contiene $U \cup W$. Il prossimo risultato fornisce un criterio per determinare un sistema di generatori di $U + W$.

Lemma 5.26. *Sia $\mathcal{D} = \{u_1, \ldots, u_n\}$ un sistema di generatori di U e $\mathcal{C} = \{w_1, \ldots, w_m\}$ un sistema di generatori di W. Allora $\mathcal{D} \cup \mathcal{C}$ è un sistema di generatori di $U + W$.*

Dimostrazione. I vettori $u_1, \ldots, u_n, w_1, \ldots w_m \in U + W$. Quindi, tenendo in mente che $U + W$ è un sottospazio di V, si ha

$$\mathcal{L}(u_1, \ldots, u_n, w_1, \ldots, w_m) \subseteq U + W.$$

Dimostriamo l'inclusione opposta.

Sia $z \in U + W$. Il vettore $z = u + w$ per un certo $u \in U$ e $w \in W$. Poiché $\mathcal{D} = \{u_1, \ldots, u_n\}$, rispettivamente $\mathcal{C} = \{w_1, \ldots, w_m\}$, è un sistema di generatori di U, rispettivamente W, esistono $\alpha_1, \ldots, \alpha_n$, rispettivamente $\beta_1, \ldots, \beta_m$, tale che

$$u = \alpha_1 u_1 + \cdots + \alpha_n u_n,$$

rispettivamente
$$w = \beta_1 w_1 + \cdots + \beta_m w_m.$$
Quindi
$$u+w = \alpha_1 u_1 + \cdots + \alpha_n u_n + \beta_1 w_1 + \cdots + \beta_m w_m \in \mathcal{L}(u_1, \ldots, u_n, w_1, \ldots, w_m),$$
ovvero $U + W \subseteq \mathcal{L}(u_1, \ldots, u_n, w_1, \ldots, w_m)$. $\square$

Esempio 5.27. *Si considerino i sottospazi*
$$U = \mathcal{L}\left(\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}\right), \qquad W = \mathcal{L}\left(\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\right)$$
di $\mathbb{R}^3$*. Allora*
$$U + W = \mathcal{L}\left(\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\right),$$
e
$$\dim(U + W) = \operatorname{rg}\begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 2 & 1 & 1 \end{pmatrix} = 3 \quad \text{verificare!}.$$
Quindi $U + W = \mathbb{R}^3$*.*

Teorema 5.28 (Formula di Grassman)**.** *Sia V uno spazio vettoriale su $\mathbb{K}$. Siano U e W sottospazi vettoriali di V. Allora*
$$\dim(U + W) + \dim(U \cap W) = \dim U + \dim W$$

Facoltativa. Sia $\{s_1, \ldots, s_k\}$ una base di $U \cap W$. Per il Teorema di completamento a base, possiamo completarla a base di U, $\{s_1, \ldots, s_k, u_1, \ldots, u_p\}$, rispettivamente a una base di W, $\{s_1, \ldots, s_k, w_1, \ldots, w_q\}$. Quindi $\dim(U \cap W) = k$, $\dim U = k + p$ ed infine $\dim W = k + q$. Per il Lemma anteriore, i vettori
$$\{s_1, \ldots, s_k, u_1, \ldots, u_p, w_1, \ldots, w_q\},$$
formano un sistema di generatori di $U + W$. Vogliamo dimostrare che sono linearmente indipendenti e quindi formano una base di $U+W$. Consideriamo una loro combinazione lineare uguale al vettore nullo:
$$\underbrace{\alpha_1 s_1 + \cdots + \alpha_k s_k}_{s} + \underbrace{\beta_1 u_1 + \cdots + \beta_p u_p}_{u} + \underbrace{\gamma_1 w_1 + \cdots + \gamma_q w_q}_{w} = 0,$$

ovvero

$$s + u + w = 0,$$

dove $s \in U \cap W$, $u \in U$ e $w \in W$. Quindi $u = -s - w \in W$, rispettivamente $w = -s - u \in U$, da cui segue che $u, w \in U \cap W$. In particolare esistono $\lambda_1, \ldots, \lambda_k$, rispettivamente $\mu_1, \ldots, \mu_k$, tali che

$$u = \lambda_1 s_1 + \cdots + \lambda_k s_k = \beta_1 u_1 + \cdots + \beta_p u_p,$$

rispettivamente

$$w = \mu_1 s_1 + \cdots + \mu_k s_k = \gamma_1 w_1 + \cdots + \gamma_q w_q,$$

ovvero

$$\lambda_1 s_1 + \cdots + \lambda_k s_k - \beta_1 u_1 - \cdots - \beta_p u_p = 0,$$

rispettivamente

$$\mu_1 s_1 + \cdots + \mu_k s_k - \gamma_1 w_1 - \cdots - \gamma_q w_q = 0.$$

Poiché i vettori $\{s_1, \ldots, s_k, u_1, \ldots, u_p\}$ formano una base di U, rispettivamente $\{s_1, \ldots, s_k, w_1, \ldots, w_q\}$ formano una base di W, ne segue che

$$\beta_1 = \cdots = \beta_p = 0, \text{ rispettivamente } \gamma_1 = \cdots = \gamma_q = 0.$$

Quindi $u = w = 0$ ed $\alpha_1 s_1 + \cdots + \alpha_k s_k = 0$, Poiché i vettori $s_1, \ldots, s_k$ sono linearmente indipendenti si ha $\alpha_1 = \cdots = \alpha_k = 0$, ovvero i vettori $s_1, \ldots, s_k, u_1, \ldots, u_p, w_1, \ldots, w_q$ sono linearmente indipendenti. Quindi,

$$\begin{aligned} \dim(U + W) &= k + p + q \\ &= (k + p) + (k + q) - k \\ &= \dim U + \dim W - \dim(U \cap W). \end{aligned}$$

$\square$

Definizione 5.29. *Sia V uno spazio vettoriale su $\mathbb{K}$. Siano U e W sottospazi vettoriali di V. Diremo U e W sono in somma diretta se $U \cap W = \{0\}$. Diremo inoltre che V è in somma diretta di U e W se:*

- $U \cap W = \{0\}$;
- $U + W = V$.

Se V è in soma diretta di U e W scriveremo $V = U \oplus W$.

Esempio 5.30. *Si considerino i sottospazi* $U = \mathcal{L}(e_1)$ *e* $W = \mathcal{L}(e_2)$ *di* $\mathbb{R}^2$. *È facile verificare che* $U \cap W = \{0\}$.

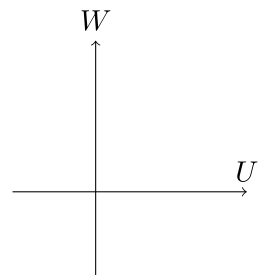


Figura 5.30.1.

Quindi U *e* W *sono in somma diretta. Inoltre,* $U + W = \mathbb{R}^2$, *poiché* $U + W = \mathcal{L}(e_1, e_2) = \mathbb{R}^2$, *ovvero* $\mathbb{R}^2 = U \oplus W$.

Applicando la formula di Grassman si ha il seguente risultato.

Corollario 5.31. *Siano* U *e* W *due sottospazi vettoriali di* V. *Allora:*

a) U *e* W *sono in somma diretta se e solamente se* $\dim(U+W) = \dim U + \dim W$*;*

b) V *è somma diretta di* U *e* W *se e solamente se* $\dim V = \dim(U + W)$ *e* $\dim(U + W) = \dim U + \dim W$.

Dimostrazione. Tenendo in mente che $U \cap W = \{0\}$ è equivalente a $\dim(U \cap W) = 0$, applicando la formula di Grassmann

$$\dim(U + W) + \dim(U \cap W) = \dim U + \dim W,$$

si ha che $U \cap W = \{0\}$ se e solamente se $\dim(U + W) = \dim U + \dim W$.

Se V è somma diretta di U e W, allora $U \cap W = \{0\}$ ed $V = U + W$. Quindi $\dim V = \dim(U + W)$. Applicando la formula di Grassmann si ha $\dim(U + W) = \dim U + \dim W$. Viceversa, supponiamo che $\dim V = \dim(U+W) = \dim U+\dim W$. Poiché $U+W \subseteq V$, applicando il Corollario 5.21 si ha $V = U + W$. Applicando nuovamente la formula di Grassman, se $\dim(U + W) = \dim U + \dim W$, allora $U \cap W = \{0\}$ concludendo la dimostrazione. $\square$

Il seguente risultato descrive geometricamente cosa significa che V è somma diretta di U e W.

Corollario 5.32. *Sia* V *uno spazio vettoriale su* $\mathbb{K}$ *e siano* U *e* W *sottospazi di* V. *Allora* V *è somma diretta di* U *e* W *se e solamente se ogni elemento di* V *si scrive in maniera unica come somma di un vettore di* U *ed un vettore di* W.

Facoltattiva. Se V è somma diretta di U e W, allora $U \cap W = \{0\}$ e $V = U + W$. La seconda proprietà implica che ogni vettore $v \in V$ si scrive come somma $v = u + w$, dove $u \in U$ e $w \in W$. Se $v = u + w = u' + w'$, allora

$$u - u' = w - w' \in U \cap W.$$

Quindi $u = u'$ e $w = w'$. Viceversa, supponiamo che ogni elemento di V si scriva, in maniera unica, come somma di un elemento di U e di un elemento di W. Allora $V = U + W$ (Perché?). Se $z \in U \cap W$, allora $z = \frac{3}{4}z + \frac{1}{4}z = \frac{1}{2}z + \frac{1}{2}z$. Quindi se $z \neq 0$, allora potrei scrivere z come somma di un elemento di U e di un elemento di W in almeno due modi differenti. Assurdo. Quindi $U \cap W = \{0\}$. $\square$

Esempio 5.33. *Sia $V = \mathrm{M}_{n\times n}(\mathbb{R})$ e siano $U = \{A \in V : A = A^T\}$ e $W = \{A \in V : A = -A^T\}$. Ogni matrice $A \in V$ si scrive in maniera unica come combinazione lineare di una matrice simmetrica ed una antisimmetrica:*

$$A = \frac{1}{2}(A + A^T) + \frac{1}{2}(A - A^T).$$

Quindi V è somma diretta di U e W.

Sia V uno spazio vettoriale su $\mathbb{K}$ e sia W un sottospazio vettoriale di V. Una domanda naturale è se esiste W' sottospazio di V tale che $V = U \oplus W$. La risposta è affermativa ed una diretta conseguenza del Teorema di completamento a base.

Corollario 5.34. *Sia V uno spazio vettoriale su $\mathbb{K}$ e sia W un sottospazio vettoriale di V. Esistono (infiniti) W' tali che $V = W \oplus W'$.*

Facoltativa. Sia $\mathcal{C} = \{w_1, \ldots, w_k\}$ una base di W. Possiamo completarla a basa di V. Siano $v_1, \ldots, v_s$ tale che i vettori $w_1, \ldots, w_k, v_1, \ldots, v_s$ formano una base di V. Affermiamo che V è in somma diretta di W e $W' = \mathcal{L}(v_1, \ldots, v_s)$.

Poiché $W + W' = \mathcal{L}(w_1, \ldots, w_k, v_1, \ldots, v_s)$ ed $w_1, \ldots, w_s, v_1, \ldots, v_s$ formano una base di V si ha $W + W' = V$.

Sia $z \in W \cap W'$. Allora

$$z = \alpha_1 w_1 + \cdots + \alpha_k w_k = \beta_1 v_1 + \cdots + \beta_s v_s,$$

da cui segue

$$\alpha_1 w_1 + \cdots + \alpha_k w_k + (-\beta_1)v_1 + \cdots + (-\beta_s)v_s = 0.$$

Poiché i vettori $w_1, \ldots, w_k, v_1, \ldots, v_s$ sono linearmente indipendenti i coefficienti

$$\alpha_1 = \cdots = \alpha_k = \beta_1 = \cdots = \beta_s = 0,$$

ovvero $z = 0$. $\square$

5.34.1 Metodi di Calcolo

Siano $U = \mathcal{L}(X_1, \ldots, X_k)$ e $W = \mathcal{L}(Y_1, \ldots, Y_s)$ sottospazi vettoriale di $\mathbb{K}^n$. Siano $A = (X_1, \ldots, X_k) \in M_{n\times k}(\mathbb{K})$, $B = (Y_1, \ldots, Y_s) \in M_{n\times s}(\mathbb{K})$ e $C = (X_1, \ldots, X_k, Y_1, \ldots, Y_s) \in M_{n\times(k+s)}(\mathbb{K})$. Allora

- $\dim U = \mathrm{rg}(A)$;
- $\dim W = \mathrm{rg}(B)$;
- $\dim(U + W) = \mathrm{rg}(C)$;
- $\dim(U \cap W)$ possiamo calcolarla applicando la formula di Grassman;
- U e W sono in somma diretta se e solamente se $\mathrm{rg}(C) = \mathrm{rg}(A)+\mathrm{rg}(B)$;
- $\mathbb{K}^n$ è in somma ditretta di U e W se e solamente se $n = \mathrm{rg}(C) = \mathrm{rg}(A) + \mathrm{rg}(B)$.

Per quanto rigurda l'intersezione possiamo così procedere.

Siano U e W. Calcoliamo equzioni cartesiane di U e W rispettivamente. Allora l'unione di queste equazioni sono equazioni cartesiane per l'intersezione $U \cap W$.

Sia V uno spazio vettoriale su $\mathbb{K}$. Siano $U = \mathcal{L}(u_1, \ldots, u_p)$, $W = \mathcal{L}(w_1, \ldots, w_q)$ sottospazi vettoriali di V. Sia infine $\mathcal{B}$ una base di V. Allora:

- $\dim U = rg(([u_1]_{\mathcal{B}}, \ldots, [u_p]_{\mathcal{B}})$;
- $\dim W = rg(([w_1]_{\mathcal{B}}, \ldots, [w_q]_{\mathcal{B}})$;
- $\dim U + W = rg(([u_1]_{\mathcal{B}}, \ldots, [u_p]_{\mathcal{B}}, [w_1]_{\mathcal{B}}, \ldots, [w_q]_{\mathcal{B}}))$.
- la $\dim(U \cap W)$ possiamo ricavarla utilizzando la formula di Grassmann.

Capitolo 6

Applicazioni lineari e matrici

6.1 Applicazioni lineari

Definizione 6.2. *Siano V, W spazi vettoriali su $\mathbb{K}$. Una applicazione $T : V \longrightarrow W$ si dice lineare se:*

a) $\forall v_1, v_2 \in V$, $T(v_1 + v_2) = T(v_1) + T(v_2)$ *(additiva);*

b) $\forall v \in V$ *e* $\lambda \in \mathbb{K}$, $T(\lambda v) = \lambda T(v)$ *(omogenea).*

Proposizione 6.3. *Sia $T : V \longrightarrow W$ una applicazione lineare. Allora*

- $T(0_V) = 0_W$*;*
- $T(-v) = -T(v)$*;*
- *siano $v_1, \dots, v_n \in V$ e siano $\alpha_1, \dots, \alpha_n \in \mathbb{K}$. Allora*

$$T(\alpha_1 v_1 + \cdots + \alpha_n v_n) = \alpha_1 T(v_1) + \cdots + \alpha_n T(v_n).$$

Dimostrazione. Dimostreremo solamente l'ultima proprietà lasciando le altre due per esercizo.

$$\begin{aligned} T(\alpha_1 v_1 + \cdots + \alpha_n v_n) &= T(\alpha_1 v_1) + T(\alpha_2 v_2 + \cdots + \alpha_n v_n) \\ &= T(\alpha_1 v_1) + T(\alpha_2 v_1) + \cdots + T(\alpha_n v_n) \\ &= \alpha_1 T(v_1) + \alpha_2 T(v_2) + \cdots + \alpha_n T(v_n) \\ &= \alpha_1 T(v_1) + \cdots + \alpha_n T(v_n). \end{aligned}$$

□

Vediamo alcuni esempi di applicazioni lineari.

Esempio 6.4.

a) $T : V \longrightarrow W$, $v \mapsto 0_W$ *è lineare;*

b) $\mathrm{Id}_V : V \longrightarrow V$, $v \mapsto v$ *è una applicazione lineare;*

c) L'applicazione

$$M_{m\times n}(\mathbb{R}) \longrightarrow M_{n\times m}(\mathbb{R}), \qquad A \mapsto A^T$$

è una applicazione lineare;

d) sia V uno spazio vettoriale su $\mathbb{K}$ e sia $\mathcal{B}$ una base di V. Allora

$$\mathcal{F}_{\mathcal{B}} : V \longrightarrow \mathbb{K}^n, \qquad v \mapsto [v]_{\mathcal{B}},$$

è una applicazione lineare;

e) sia $A \in M_{m\times n}(\mathbb{K})$. Le applicazioni

$$M_{p\times m}(\mathbb{K}) \longrightarrow M_{p\times n}(\mathbb{K}), \qquad X \mapsto XA,$$

e

$$M_{n\times q}(\mathbb{K}) \longrightarrow M_{m\times q}(\mathbb{K}), \qquad X \mapsto AX,$$

sono applicazioni lineari.

f) Sia $A \in M_{m\times n}(\mathbb{K})$. Allora

$$L_A : \mathbb{K}^n \longrightarrow \mathbb{K}^m, \qquad X \mapsto AX,$$

è una applicazione lineare.

g) L'applicazione

$$M_{n\times n}(\mathbb{K}) \longrightarrow \mathbb{K} \qquad A \mapsto \mathrm{Tr}(A)$$

è una applicazione lineare.

Osservazione 6.5. *L'applicazione $T : \mathbb{R} \longrightarrow \mathbb{R}$, $T(x) = x + 1$ non è una applicazione lineare poiché $T(0) = 1 \neq 0$. Lo stesso vale per l'applicazione $T : \mathbb{R} \longrightarrow \mathbb{R}$, $T(x) = x^2$ poiché $T(-x) = T(x)$ per ogni $x \in \mathbb{R}$. Un altro esempio di applicazione che non è lineare è l'applicazione*

$$M_{m\times n}(\mathbb{C}) \longrightarrow M_{n\times m}(\mathbb{C}), \qquad A \in M_{m\times n}(\mathbb{C}) \mapsto A^* \in M_{n\times m}(\mathbb{C}).$$

Infatti $(\lambda A)^* = \overline{\lambda} A$, ovvero non è omogenea (ma è additiva).

Siano V, W spazi vettoriali su $\mathbb{K}$. Indichiamo con $Lin(V, W)$ oppure $Hom(V, W)$ l'insieme di tutte le applicazioni lineari fra V e W. Se $T, L \in Lin(V, W)$, diremo che $T = L$ se per ogni $v \in V$ si ha $T(v) = L(v)$. $Lin(V, W)$ ammette una struttura di spazio vettoriale su $\mathbb{K}$.

Siano $T, L \in Lin(V, W)$ e $\lambda \in \mathbb{K}$. Definiamo

$$(T + L)(v) := T(v) + L(v), \qquad (\lambda T)(v) = \lambda T(v).$$

Proposizione 6.6. *Se V, W sono spazi vettoriali su $\mathbb{K}$, allora $Lin(V, W)$ con le operazioni di somma e moltiplicazione per scalare definite come sopra è uno spazio vettoriale su $\mathbb{K}$.*

Proposizione 6.7. *Siano $f \in Lin(V, W)$ e $g \in Lin(W, Z)$. Allora $g \circ f \in Lin(V, Z)$ è ancora una applicazione lineare, ovvero la composizione di applicazioni lineari è ancora una applicazione lineare. L'inversa, quando esiste, di una applicazione lineare è ancora lineare. Inoltre: siano $L : V \longrightarrow W$, $T, H : W \longrightarrow Z$ e $S, Q : D \longrightarrow W$. Allora*

a) $(T + H) \circ L = T \circ L + H \circ L$;

b) $T \circ (S + Q) = T \circ S + T \circ Q$;

c) $\lambda(H \circ L) = (\lambda H) \circ L = H \circ (\lambda L)$.

Facoltativa. Siano $v_1, v_2 \in V$. Allora

$$\begin{aligned}(g \circ f)(v_1 + v_2) &= g(f(v_1 + v_2)) = g(f(v_1) + f(v_2)) \\ &= g(f(v_1)) + g(f(v_2)) = (g \circ f)(v_1) + (g \circ f)(v_2).\end{aligned}$$

Analogamente se $\lambda \in \mathbb{K}$ e $v \in V$ si ha

$$\begin{aligned}(g \circ f)(\lambda v) &= g(f(\lambda v)) = g(\lambda f(v)) \\ &= \lambda g(f(v)) = \lambda(g \circ f)(v).\end{aligned}$$

Sia $T : V \longrightarrow W$ lineare e biunivoca. Dalla teoria degli insieme sappiamo che esiste $T^{-1} : W \longrightarrow V$ tale che $T \circ T^{-1} = \mathrm{Id}_W$ e $T^{-1} \circ T = \mathrm{Id}_V$. L'applicazione $T^{-1} : W \longrightarrow V$ è così definta: $T^{-1}(w) = v$, dove v è l'unico elemento di V tale che $T(v) = w$. Vogliamo dimostrare che T^{-1} è lineare.

Siano $w_1, w_2 \in W$ e siano v_1, rispettivamente $v_2 \in V$, tali che $T(v_1) = w_1$, rispettivamente $T(v_2) = w_2$. Poiché T è lineare si ha

$$T(v_1 + v_2) = T(v_1) + T(v_2) = w_1 + w_2.$$

Per definizione di T^{-1} si ha

$$T^{-1}(w_1 + w_2) = v_1 + v_2 = T^{-1}(w_1) + T^{-1}(w_2).$$

Siano $\lambda \in \mathbb{K}$ e $w \in W$. Sia $v \in V$ tale che $T(v) = w$. Poiché T è lineare si ha $T(\lambda v) = \lambda T(v) = \lambda w$, ovvero

$$T^{-1}(\lambda w) = \lambda T^{-1}(w).$$

Le rimanenti proprietà sono lasciate per esercizio. □

Il prossimo risultato garantisce che una applicazione lineare è completamente determinate dai valori che assume su una base di V.

Proposizione 6.8. *Siano V e W spazi vettoriali su $\mathbb{K}$. Sia $\mathcal{B} = \{v_1, \ldots, v_n\}$ una base di V e siano $w_1, \ldots, w_n \in W$. Esiste una unica applicazione lineare $T : V \longrightarrow W$, tale che $T(v_1) = w_1, \ldots, T(v_n) = w_n$.*

Dimostrazione. Supponiamo che esistono $T, L : V \longrightarrow W$ lineari tali che $T(v_1) = L(v_1), \ldots, T(v_n) = L(v_n)$ e dimostriamo che $T = L$.

Sia $v \in V$. Poiché $\mathcal{B} = \{v_1, \ldots, v_n\}$ è una base di V esistono, e sono unici, $x_1, \ldots, x_n \in \mathbb{K}$, tale che $v = x_1 v_1 + \cdots + x_n v_n$. Quindi

$$\begin{aligned} T(v) &= T(x_1 v_1 + \cdots + x_n v_n) \\ &= x_1 T(v_1) + \cdots + x_n T(v_n) \\ &= x_1 L(v_1) + \cdots + x_n L(v_n) \\ &= L(x_1 v_1 + \cdots + x_n v_n) \\ &= L(v). \end{aligned}$$

Quindi se esite una tale applicazione è unica. Adesso dimostriamo che esiste una applicazione lineare $T : V \longrightarrow W$ tale che $T(v_1) = w_1, \ldots, T(v_n) = w_n$.

Sia $v \in V$. Allora esistono, e sono unici, $x_1, \ldots, x_n \in \mathbb{K}$ tale che

$$v = x_1 v_1 + \cdots + x_n v_n.$$

Definiamo

$$T(v) = x_1 w_1 + \cdots + x_n w_n.$$

La definizione è ben posta poiché le coordinate sono univocamente determinate. Inoltre, tenendo in mente che $[v_1]_{\mathcal{B}} = e_1, \ldots, [v_n]_{\mathcal{B}} = e_n$, si ha $T(v_1) = w_1, \ldots, T(v_n) = w_n$. Rimane quindi da dimostrare che T è lineare.

Siano $v, w \in V$. Allora

$$\begin{array}{l} v = x_1 v_1 + \cdots + x_n v_n \\ w = y_1 v_1 + \cdots + y_n v_n \\ v + w = (x_1 + y_1) v_1 + \cdots + (x_n + y_n) v_n \end{array}.$$

Quindi

$$\begin{aligned} T(v + w) &= (x_1 + y_1) w_1 + \cdots + (x_n + y_n) w_n \\ &= \underbrace{x_1 w_1 + \cdots + x_n w_n} + \underbrace{y_1 w_1 + \cdots + y_n w_n} \\ &= T(v) + T(w). \end{aligned}$$

Sia $v \in V$ e sia $\lambda \in \mathbb{K}$. Allora

$$v = x_1 v_1 + \cdots + x_n v_n \qquad \lambda v = \lambda x_1 v_1 + \cdots + \lambda x_n v_n.$$

Quindi

$$\begin{aligned} T(\lambda v) &= (\lambda x_1) w_1 + \cdots + (\lambda x_n) w_n \\ &= \lambda (x_1 w_1 + \cdots + x_n w_n) \\ &= \lambda T(v). \end{aligned}$$

$\square$

Esempio 6.9. *Sia $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^4$ l'applicazione lineare definita da:*

$$T\left(\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 1 \\ 2 \\ -1 \end{bmatrix}, T\left(\begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 0 \\ -3 \\ 1 \\ 1 \end{bmatrix}, T\left(\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 2 \\ -1 \\ 5 \\ -1 \end{bmatrix}.$$

Vogliamo calcolare $T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right)$.

Sia $\mathcal{B} = \left\{\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\right\}$, la base di $\mathbb{R}^3$ sulla quale conosciamo l'applicazione lineare T. Se

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix}_{\mathcal{B}} = \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix}$$

sono le coordinate del vettore $\begin{bmatrix} x \\ y \\ z \end{bmatrix}$ *rispetto alla base* $\mathcal{B}$, *allora*

$$T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \alpha \begin{bmatrix} 1 \\ 1 \\ 2 \\ -1 \end{bmatrix} + \beta \begin{bmatrix} 0 \\ -3 \\ 1 \\ 1 \end{bmatrix} + \gamma \begin{bmatrix} 2 \\ -1 \\ 5 \\ -1 \end{bmatrix}.$$

Le coordinate del vettore $\begin{bmatrix} x \\ y \\ z \end{bmatrix}$ *rispetto alla base* $\mathcal{B}$ *è l'unica soluzione del sistema lineare*

$$\begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix},$$

ovvero

$$\begin{cases} \alpha = -x - y + 2z \\ \beta = -x + z \\ \gamma = 2x + y - 2z \end{cases}$$

Infatti

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = (-x - y + 2z) \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + (-x + z) \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix} + (2x + y - 2z) \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

Quindi

$$\begin{aligned} T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) &= (-x - y + 2z) \begin{bmatrix} 1 \\ 1 \\ 2 \\ -1 \end{bmatrix} + (-x + z) \begin{bmatrix} 0 \\ -3 \\ 1 \\ 1 \end{bmatrix} \\ &+ (2x + y - 2z) \begin{bmatrix} 2 \\ -1 \\ 5 \\ -1 \end{bmatrix} \\ &= \begin{bmatrix} 3x + y - 2z \\ -2y + z \\ 7x + 3y - 5z \\ -2x + z \end{bmatrix}. \end{aligned}$$

Corollario 6.10. *Siano $T, L : V \longrightarrow W$ due applicazioni lineari e sia $\mathcal{B} = \{v_1, \ldots, v_n\}$ una base di V. Allora $T = L$, se e solamente se $T(v_1) = L(v_1), \ldots, T(v_n) = L(v_n)$.*

Dimostrazione. Se $T = L$, allora $T(v) = L(v)$ per ogni $v \in V$. In particolare vale per ogni elemento della base $\mathcal{B} = \{v_1, \ldots, v_n\}$. Viceversa, supponiamo che $T(v_i) = L(v_i)$ per $i = 1, \ldots, n$. Applicando la proposizione anteriore si ha $T = L$. $\square$

Definizione 6.11. *Sia $T : V \longrightarrow W$ una applicazione lineare. Si pone*

- $Ker\, T = \{v \in V : T(v) = 0_W\}$ *(nucleo di T).*
- $Im\, T = \{T(v) : v \in V\} = \{w \in W : \exists v \in V \text{ per cui } T(v) = w\}$ *(Immagine di T).*

Proposizione 6.12. *Sia $T : V \longrightarrow W$ una applicazione lineare. Allora*

- $Ker\, T$ *è un sottospazio vettoriale di V.*
- $Im\, T$ *è un sottospazio vettoriale di W.*
- *Se $\mathcal{B} = \{v_1, \ldots, v_n\}$ è una base di V, allora*

  $$Im\, T = \mathcal{L}(T(v_1), \ldots, T(v_n)).$$

  Quindi $\dim Im\, T \leq \dim V$.

Dimostrazione.

- Dobbiamo dimostrare che $\mathrm{Ker}\, T$ è chiuso rispetto alla somma e la moltiplicazione per scalare.

  Siano $v, z \in Ker\, T$. Allora $T(v) = T(z) = 0_W$. Dobbiamo dimostrare che $v + z \in Ker\, T$, ovvero

  $$T(v + z) = 0_W.$$

  Poiché T è lineare, si ha

  $$T(v + z) = T(v) + T(z) = 0_W,$$

  quindi $v + z \in \mathrm{Ker}\, T$.

  Siano $v \in \mathrm{Ker}\, T$ e sia $\lambda \in \mathbb{K}$. Dobbiamo dimostrare che $\lambda v \in \mathrm{Ker}\, T$, cioè $T(\lambda v) = 0_W$. Poiché T è lineare e $v \in \mathrm{Ker}\, T$ si ha

  $$T(\lambda v) = \lambda T(v) = 0_W.$$

- Siano $w_1, w_2 \in Im\,T$ e $\lambda \in \mathbb{K}$. Dobbiamo dimostrare che $w_1 + w_2 \in \operatorname{Im} T$ e $\lambda w_1 \in \operatorname{Im} T$, ovvero che esistono $v, z \in V$ tali che $T(v) = w_1 + w_2$ e $T(z) = \lambda w_1$, rispettivamente. Per ipotesi esistono $v_1, v_2 \in V$ tali che $T(v_1) = w_1$ e $T(v_2) = w_2$. Quindi

$$T(v_1 + v_2) = T(v_1) + T(v_2) = w_1 + w_2 \in Im\,T,$$

rispettivamente

$$T(\lambda v_1) = \lambda T(v_1) = \lambda w_1 \in Im\,T.$$

- Sia $\mathcal{B} = \{v_1, \dots, v_n\}$ una base di V e sia $w \in Im\,T$. Allora $w = T(v)$ per un certo $v \in V$. Poiché $\mathcal{B} = \{v_1, \dots, v_n\}$ è una base di V si ha $v = x_1 v_1 + \cdots + x_n v_n$. Per la linearità di T, si ha

$$\begin{aligned} w = T(v) &= T(x_1 v_1 + \cdots + x_n v_n) \\ &= x_1 T(v_1) + \cdots + x_n T(v_n) \in \mathcal{L}(T(v_1), \dots, T(v_n)). \end{aligned}$$

Quindi $Im\,T \subseteq \mathcal{L}(T(v_1), \dots, T(v_n))$. Poiché $T(v_1), \dots, T(v_n) \in Im\,T$, tenendo in mente che $Im\,T$ è un sottospazio vettoriale di W, si ha

$$\mathcal{L}(T(v_1), \dots, T(v_n)) \subseteq Im\,T.$$

Quindi

$$Im\,T = \mathcal{L}(T(v_1), \dots, T(v_n)).$$

Infine, applicando il Lemma di Steintz si ha $\dim Im\,T \leq \dim V$.

$\square$

Esempio 6.13. *Si consideri l'applicazione lineare* $T : \mathbb{R}^4 \longrightarrow \mathbb{R}^3$ *definita da*

$$T\left(\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix}\right) = \begin{bmatrix} x - 2z + 2w \\ x + 2y + 4w \\ y + z + w \end{bmatrix},$$

Vogliamo calcolare $\operatorname{Ker} T$, *ovvero i vettori* $\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} \in \mathbb{R}^4$ *tali che*

$$T\left(\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix}\right) = \begin{bmatrix} x - 2z + 2w \\ x + 2y + 4w \\ y + z + w \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Quindi $\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} \in \operatorname{Ker} T$ *se e solamente se*

$$\begin{cases} x - 2z + 2w = 0 \\ x + 2y + 4w = 0 \\ y + z + w = 0. \end{cases}$$

Applicando l'algoirtmo di Gauss e il metodo della risoluzione all'indietro, si ha

$$\operatorname{Ker} T = \left\{ \begin{bmatrix} 2z - 2w \\ -z - w \\ z \\ w \end{bmatrix} : z, w \in \mathbb{R} \right\} = \left\{ z \begin{bmatrix} 2 \\ -1 \\ 1 \\ 0 \end{bmatrix} + w \begin{bmatrix} -2 \\ -1 \\ 0 \\ 1 \end{bmatrix} : z, w \in \mathbb{R} \right\}.$$

Quindi $\operatorname{Ker} T$ *ha dimensione* 2 *ed una base è formata dai vettori* $\left\{ \begin{bmatrix} 2 \\ -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ -1 \\ 0 \\ 1 \end{bmatrix} \right\}$.

Adesso vogliamo calcolare l'immagine di T.

Sia $\mathcal{C} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right\} = \{e_1, e_2, e_3, e_4\}$, *la base canonica di* $\mathbb{R}^4$. *Per la Proposizione 6.12, si ha*

$$\operatorname{Im} T = \mathcal{L}\left(T(e_1), T(e_2), T(e_3), T(e_4)\right) = \mathcal{L}\left(\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 4 \\ 1 \end{bmatrix} \right).$$

Poiché la matrice (verificare)

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

ha rango 2, *si ha che* $\dim \operatorname{Im} T = 2$ *ed una base è formata dai vettori*

$$\left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix} \right\}.$$

Esempio 6.14. *Si consideri l'applicazione lineare* $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ *definita da:*

$$T\left(\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}, T\left(\begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 3 \\ -2 \\ 1 \end{bmatrix}, T\left(\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 2 \\ 3 \\ 5 \end{bmatrix}.$$

Per la Proposizione 6.12, si ha

$$\operatorname{Im} T = \mathcal{L}\left(\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 3 \\ -2 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \\ 5 \end{bmatrix}\right).$$

Poiché la matrice (verificare)

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

ha rango 2 *si ha che* $\dim \operatorname{Im} T = 2$ *ed una base è formata dai vettori* $\left\{\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 3 \\ -2 \\ 1 \end{bmatrix}\right\}$*. Per calcolare il nucleo, dobbiamo calcolare* $T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right)$.
Sia $\mathcal{B} = \left\{\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\right\}$*, la base di* $\mathbb{R}^3$ *sulla quale conosciamo l'applicazione lineare* T*. Le coordinate del vettore* $\begin{bmatrix} x \\ y \\ z \end{bmatrix}$ *rispetto alla base* $\mathcal{B}$ *è l'unica soluzione del sistema lineare*

$$\begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix},$$

ovvero

$$\begin{cases} \alpha = -x - y + 2z \\ \beta = -x + z \\ \gamma = 2x + y - 2z \end{cases}$$

Infatti

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = (-x - y + 2z) \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + (-x + z) \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix} + (2x + y - 2z) \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

Quindi

$$T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = (-x-y+2z)\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} + (-x+z)\begin{bmatrix} 3 \\ -2 \\ 1 \end{bmatrix}$$
$$+ (2x+y-2z)\begin{bmatrix} 2 \\ 3 \\ 5 \end{bmatrix}$$
$$= \begin{bmatrix} y+z \\ 7x+2y-6z \\ 7x+3y-5z \end{bmatrix}.$$

Quindi $\begin{bmatrix} x \\ y \\ z \end{bmatrix} \in \operatorname{Ker} T$ *se e solamente se*

$$T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} y+z \\ 7x+2y-6z \\ 7x+3y-5z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

ovvero $\begin{bmatrix} x \\ y \\ z \end{bmatrix} \in \operatorname{Ker} T$ *se e solamente se*

$$\begin{cases} y+z=0 \\ 7x+2y-6z=0 \\ 7x+3y-5z=0. \end{cases}$$

Applicando l'algoirtmo di Gauss e il metodo della risoluzione all'indietro, si ha

$$\operatorname{Ker} T = \left\{ \begin{bmatrix} 8/7z \\ -z \\ z \end{bmatrix} : z \in \mathbb{R} \right\} = \left\{ z \begin{bmatrix} 8/7 \\ -1 \\ 1 \end{bmatrix} : z \in \mathbb{R} \right\}.$$

Quindi $\operatorname{Ker} T$ *ha dimensione* 1 *ed una base è formata dal vettore* $\left\{ \begin{bmatrix} 8/7 \\ -1 \\ 1 \end{bmatrix} \right\}$.

Osservazione 6.15. *Sia* $T : V \longrightarrow W$ *una applicazione lineare. Ricordiamo che se* $Z \subseteq W$, *allora la controimmagine di* Z *rispetto a* T, *che indicheremo con* $T^{-l}(Z)$, *è sottoinsieme di* V *così definito:*

$$T^{-1}(Z) = \{v \in V : T(v) \in Z\}.$$

Se Z è un sottospazio vettoriale, allora si può dimostrare che $T^{-1}(Z)$ è un sottospazio vettoriale di V. Osserviamo che $\operatorname{Ker} T = T^{-1}(0_W)$.

Se $L \subseteq V$, allora l'immagine di L rispetto a T è il sottoinsieme di W così definito:

$$T(L) = \{T(s) : s \in L\}$$

Se L è un sottospazio vettoriale di V, allora si può dimostrare che $T(L)$ è un sottospazio vettoriale di W. Inoltre, se $L = \mathcal{L}(v_1, \ldots, v_k)$, allora

$$T(L) = \mathcal{L}(T(v_1), \ldots, T(v_k)).$$

Proposizione 6.16. *Sia $T : V \longrightarrow W$ una applicazione lineare. Allora:*

- *T è iniettiva se e solamente se $Ker\, T = \{0_V\}$;*
- *T è suriettiva se e solamente se* $\dim \operatorname{Im} T = \dim W$.

Dimostrazione. Una applicazione T è iniettiva se $T(v) \neq T(w)$ ogni volta che $v \neq w$. Poiché T è lineare, si ha $T(0_V) = 0_W$. Quindi se $v \neq 0_V$, allora $T(v) \neq T(0_V) = 0_W$, ovvero $\operatorname{Ker} T = \{0_V\}$.

Viceversa, supponiamo che $Ker\, T = \{0_V\}$. Vogliamo dimostrare che se $T(v) = T(w)$, allora $v = w$.

Se $T(v) = T(w)$, allora per la linearità di T si ha $T(v - w) = 0$. Quindi $v - w \in Ker\, T = \{0_V\}$ da cui segue $v = w$.

Un'applicazione T è suriettiva se $\operatorname{Im} T = W$. Poiché $\operatorname{Im} T \subseteq W$, per il Corollario 5.21, la condizione $\operatorname{Im} T = W$ è equivalente a $\dim W = \dim \operatorname{Im} T$. □

Teorema 6.17 (Teorema della dimensione)**.** *Sia $T : V \longrightarrow W$ una applicazione lineare. Allora*

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

Facoltativa. Sia $\{v_1, \ldots, v_k\}$ una base di $Ker\, T$. Per il Teorema del complemento a base, esistono $v_{k+1}, \ldots, v_n$ tale che $\{v_1, \ldots, v_n\}$ è una base di V. Poiché $T(v_1) = \cdots = T(v_k) = 0$, applicando la Proposizione 6.12, si ha

$$Im\, T = \mathcal{L}(T(v_1), \ldots, T(v_n)) = \mathcal{L}(T(v_{k+1}), \ldots, T(v_n)).$$

Quindi $\dim V = n$, $\dim \operatorname{Ker} T = k$ e $\operatorname{Im} T$ è generata da $n - k$ vettori. Il Teorema sarà dimostrato se proveremo che i vettori $T(v_{k+1}), \ldots, T(v_n)$ sono linearmente indipendenti e quindi una base di $Im\, T$. In tal caso avrei

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

Siano $\alpha_{k+1}, \dots, \alpha_n \in \mathbb{K}$ tali che

$$\alpha_{k+1}T(v_{k+1}) + \cdots + \alpha_n T(v_n) = 0_W.$$

Per la linearità di T si ha

$$T(\alpha_{k+1}v_{k+1} + \cdots + \alpha_n v_n) = 0_W,$$

ovvero $\alpha_{k+1}v_{k+1} + \cdots + \alpha_n v_n \in Ker\,T$. Poiché $\{v_1, \dots, v_k\}$ formano una base di $\operatorname{Ker} T$ esistono $\alpha_1, \dots, \alpha_k \in \mathbb{K}$ tali che

$$\alpha_1 v_1 + \cdots + \alpha_k v_k = \alpha_{k+1}v_{k+1} + \cdots + \alpha_n v_n,$$

ovvero

$$\alpha_1 v_1 + \cdots + \alpha_k v_k - \alpha_{k+1}v_{k+1} - \cdots - \alpha_n v_n = 0_V.$$

Poiché i vettori $v_1, \dots, v_n$ formano una base di V, si ha $\alpha_1 = \cdots = \alpha_n = 0$, ed in particolare $\alpha_{k+1} = \cdots = \alpha_n = 0$, per cui i vettori $T(v_{r+1}), \dots, T(v_n)$ sono linearmente indipendenti. $\square$

Vediamo alcune conseguenze del Teorema della dimensione.

Corollario 6.18. *Sia $T : V \longrightarrow W$ una applicazione lineare. Allora T è iniettiva se e solamente se* $\dim Im\,T = \dim V$.

Dimostrazione. Per la Proposizione 6.16, T è iniettiva se e solamente se $\operatorname{Ker} T = \{0\}$. Applicando il Teorema della dimensione,

$$\dim V = \dim \operatorname{Ker} T + \operatorname{Im} T,$$

si ha che T è iniettiva se e solamente se $\dim Im\,T = \dim V$. $\square$

Esempio 6.19. *Sia $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^4$ l'applicazione lineare definita da:*

$$T\left(\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 1 \\ 1 \\ -1 \end{bmatrix}, T\left(\begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 2 \\ 0 \\ -1 \\ 3 \end{bmatrix}, T\left(\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 0 \\ 2 \\ 3 \\ -5 \end{bmatrix}.$$

Quindi

$$\operatorname{Im} T = \mathcal{L}\left(\begin{bmatrix} 1 \\ 1 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ -1 \\ 3 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ 3 \\ -5 \end{bmatrix}\right)$$

e

$$\dim \operatorname{Im} T = \operatorname{rg}\begin{pmatrix} 1 & 2 & 0 \\ 1 & 0 & 2 \\ 1 & -1 & 3 \\ -1 & 3 & -5 \end{pmatrix} = 2 \quad \text{verificare!.}$$

Quindi T non è iniettiva.

Corollario 6.20. *Sia $T : V \longrightarrow W$ una applicazione lineare. Allora:*

- *se T è iniettiva, allora* $\dim V \leq \dim W$*;*
- *se T è suriettiva, allora* $\dim V \geq \dim W$*;*
- *se T è biunivoca, allora* $\dim V = \dim W$*.*

Dimostrazione. Dimostriamo solamente la seconda affermazione poiché le altre sono analoghe. Se T è suriettiva, allora per la Proposizione 6.16, si ha $\dim W = \dim \operatorname{Im} T$. Per il Teorema della dimensione si ha

$$\dim V = \dim Ker\, T + \dim Im\, T = \dim Ker\, T + \dim W \geq \dim W.$$

□

Corollario 6.21. *Sia $T : V \longrightarrow W$ lineare. Supponiamo che* $\dim V = \dim W$*. Le seguenti condizioni sono equivalenti:*

a) T è iniettiva;

b) T è suriettiva;

c) T è biunivoca.

Dimostrazione. Proveremo solamente $(b) \Rightarrow (c)$.

Se T è suriettiva, allora $\dim Im\, T = \dim W$, applicando il Teorema della dimensione si ha

$$\dim V = \dim Ker\, T + \dim Im\, T = \dim Ker\, T + \dim W.$$

Poiché $\dim V = \dim W$ si ha $\dim Ker\, T = 0$, ovvero T è iniettiva. □

Definizione 6.22. *Due spazi vettoriali V, W sono* isomorfi *se esiste una applicazione lineare $T : V \longrightarrow W$ iniettiva e suriettiva, ovvero biunivoca. Un'applicazione lineare e biunivoca si chiama* isomorfismo*.*

Proposizione 6.23. *Due spazi vettoriali sono isomorfi se e solamente se* $\dim V = \dim W$.

Facoltativa. Applicando il Corollario 6.20 si ha che due spazi vettoriali isomorfi hanno la stessa dimensione . Viceversa, supponiamo che $\dim V = \dim W = n$. Sia $\mathcal{B} = \{v_1, \ldots, v_n\}$, rispettivamente $\mathcal{C} = \{w_1, \ldots, w_n\}$, una base di V, rispettivamente W. Sia $T : V \longrightarrow W$ l'unica applicazione lineare tale che $T(v_1) = w_1, \ldots, T(v_n) = w_n$ (Proposizione 6.8). Poiché l'immagine contiene una base di W è suriettiva. Per il corollario anteriore T è biunivoca e quindi V e W sono isomorfi. $\square$

Capitolo 7

Matrici e applicazioni lineari

7.1 Applicazioni lineari da $\mathbb{K}^n$ a $\mathbb{K}^m$

Sia $A \in M_{m\times n}(\mathbb{K})$ e sia $L_A : \mathbb{K}^n \longrightarrow \mathbb{K}^m$ l'applicazione lineare così definita: $L_A(X) = AX$. Per esempio, se $A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 3 \end{bmatrix} \in M_{2\times 3}(\mathbb{R})$, allora

$$L_A : \mathbb{R}^3 \longrightarrow \mathbb{R}^2,\ L_A\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 3 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x+2y+3z \\ y+3z \end{bmatrix}.$$

Allora:

a) $\operatorname{Ker} L_A = \{X \in \mathbb{K}^n : L_A(X) = 0_{\mathbb{K}^m}\} = \{X \in \mathbb{K}^n : AX = 0_{\mathbb{K}^m}\} = Sol(A|0_{\mathbb{K}^m})$;

b) $\operatorname{Im} L_A = \{b \in \mathbb{K}^m : \exists X \in \mathbb{K}^n \text{ per cui } AX = b\} = \{b \in \mathbb{K}^m : Sol(A|b) \neq \emptyset\}$ ovvero l'immagine di L_A è l'insieme dei vettori di $\mathbb{K}^m$ per i quali il sistema $AX = b$ è compatibile;

c) sia $e_1, \ldots, e_n$ la base canonica di $\mathbb{K}^n$. Poiché

$$L_A(e_i) = Ae_i = A^i,$$

per $i = 1, \ldots, n$, si ha

$$\operatorname{Im} L_A = \mathcal{L}\left(L_A(e_1), \ldots, L_A(e_n)\right) = \mathcal{L}(A^1, \ldots, A^n),$$

da cui segue che $\dim \operatorname{Im} L_A = \operatorname{rg}(A)$;

d) sia $b \in \mathbb{K}^m$. Allora $L_A^{-1}(b) = \{X \in \mathbb{K}^n : AX = b\} = \mathrm{Sol}(A|b)$. Quindi $b \in \mathrm{Im}\, L_A$ se e solamente se il sistema lineare $AX = b$ è compatibile e quindi se e solamente se $\mathrm{rg}(A) = \mathrm{rg}(A|b)$.

Applicando il Teorema della dimensione, otteniamo il seguente risultato.

Teorema 7.2 (Teorema di nullità più rango)**.** *Sia $A \in M_{m\times n}(\mathbb{K})$ e sia $L_A : \mathbb{K}^n \longrightarrow \mathbb{K}^m$ l'applicazione lineare associata ad A. Allora*

$$n = \dim Ker\, L_A + rg(A)$$

.

Il risultato anteriore ci permette di calcolare la dimensione dell'insieme delle soluzioni di un sistema lineare omogeneo.

Corollario 7.3. *Sia $A \in \mathrm{M}_{m\times n}(\mathbb{K})$. Allora* $\dim \mathrm{Sol}(A|0_{K^m}) = n - \mathrm{rg}(A)$.

Dimostrazione. Sia $L_A : \mathbb{K}^n \longrightarrow \mathbb{K}^m$, l'applicazione lineare $L_A(X) = AX$. Per Il Teorema anteriore si ha

$$\dim \mathrm{Sol}(A|0_{\mathbb{K}^m}) = \dim \mathrm{Ker}\, L_A = n - \mathrm{rg}(A).$$

□

Applicando il Teorema della dimensione, ed i suoi corollari, si ha il seguente risultato.

Corollario 7.4. *Sia $A \in M_{m\times n}(\mathbb{K})$ e sia $L_A : \mathbb{K}^n \longrightarrow \mathbb{K}^m$. Allora*

- *L_A è suriettiva se e solamente se $rg(A) = m$;*
- *L_A è iniettiva se e solamente se $rg(A) = n$;*
- *L_A è biunivoca se e solamente se $A \in M_{n\times n}(\mathbb{K})$ e $rg(A) = n$ se e solamente se $A \in M_{n\times n}(\mathbb{K})$ e $\det(A) \neq 0$.*

Infine, $L_{\mathrm{Id}_n} = \mathrm{Id}_{\mathbb{K}^n}$.

Dimostrazione. L_A è suriettiva se e solamente se $\dim \mathrm{Im}\, L_A = \dim \mathbb{K}^m$ se e solamente se $rg(A) = m$.

L_A è iniettiva se e solamente $\dim \mathrm{Im}\, L_A = \dim \mathbb{K}^n$ se e solamente se $rg(A) = n$.

L_A è biunivoca se e solamente se A è una matrice quadrata e $\dim \mathrm{Im}\, L_A = \dim \mathbb{K}^n = n$ quindi se e solamente se A è una matrice quadrata di formato $n \times n$ e $\mathrm{rg}(A) = n$, ovvero se e solamente se A è una matrice invertbile.

Infine, $L_{\mathrm{Id_n}} : \mathbb{K}^n \longrightarrow \mathbb{K}^n$ è l'applicazione lineare,

$$L_{\mathrm{Id_n}}(X) = \mathrm{Id}_n X = X = \mathrm{Id}_{\mathbb{K}^n}(X).$$

$\square$

Esempio 7.5. *Sia* $A = \begin{bmatrix} 1 & 1 & 2 & 1 \\ -1 & 2 & 3 & 2 \\ 4 & 1 & 3 & 1 \end{bmatrix}$. *Allora* $L_A : \mathbb{R}^4 \longrightarrow \mathbb{R}^3$,

$$L_A\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}\right) = \begin{bmatrix} x_1 + x_2 + 2x_3 + x_4 \\ -x_1 + 2x_2 + 3x_3 + 2x_2 \\ 4x_1 + x_2 + 3x_3 + x_4 \end{bmatrix}$$

La matrice matrice A ha rango 2 (verificare!). Quindi $\dim \mathrm{Im}\, L_A = 2$ *e* $\dim \mathrm{Ker}\, L_A = 2$. *Una base per l'immagine di L_A è formata dai vettori*

$$\left\{ \begin{bmatrix} 1 \\ -1 \\ 4 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \right\}.$$

Il nucleo di L_A è l'insieme delle soluzione del sistema lineare omogeneo

$$AX = 0_{\mathbb{R}^3}.$$

Una base per il nucleo di L_A è formata dai vettori

$$\left\{ \begin{bmatrix} 0 \\ -1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 5 \\ -3 \\ 0 \end{bmatrix} \right\}$$

Sia $A \in \mathrm{M}_{m\times n}(\mathbb{K})$ e sia $B \in \mathrm{M}_{n\times p}(\mathbb{K})$. Allora

$$L_A : \mathbb{K}^n \longrightarrow \mathbb{K}^m,$$

rispettivamente

$$L_B : \mathbb{K}^p \longrightarrow \mathbb{K}^n.$$

L'applicazione

$$L_A \circ L_B : \mathbb{K}^p \longrightarrow \mathbb{K}^m,$$

è lineare.

Proposizione 7.6. *$L_A \circ L_B = L_{AB}$. In particolare, se $A \in M_{n\times n}(\mathbb{K})$ è una matrice invertibile, allora L_A è biunivoca e $L_A^{-1} = L_{A^{-1}}$*

Dimostrazione.

$$(L_A \circ L_B)(X) = L_A(L_B(X)) = L_A(BX) = ABX = (AB)(X) = L_{AB}(X).$$

Sia $A \in \mathrm{M}_{n\times n}(\mathbb{K})$ una matrice invertibile. Abbiamo già visto che L_A è biunivoca e viceversa. Inoltre, tenendo in mente che $L_{\mathrm{Id}_n} = \mathrm{Id}_{\mathbb{K}^n}$, si ha $L_A \circ L_{A^{-1}} = L_{AA^{-1}} = L_{\mathrm{Id}_n} = \mathrm{Id}_{\mathbb{K}^n}$ e $L_{A^{-1}} \circ L_A = L_{A^{-1}A} = L_{\mathrm{Id}} = \mathrm{Id}_{\mathbb{K}^n}$, ovvero

$$(L_A)^{-1} = L_{A^{-1}}$$

□

Esempio 7.7. *Sia $A = \begin{bmatrix} 1 & 1 & 2 \\ 1 & -1 & 1 \end{bmatrix} \in M_{2\times 3}(\mathbb{R})$ e sia $B = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & -1 \end{bmatrix} \in M_{3\times 2}(\mathbb{R})$. Allora:*

$$L_A : \mathbb{R}^3 \longrightarrow \mathbb{R}^2, \qquad L_A\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} x+y+2z \\ x-y+z \end{bmatrix},$$

$$L_B : \mathbb{R}^2 \longrightarrow \mathbb{R}^3, \qquad L_B\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x+y \\ y \\ x-y \end{bmatrix},$$

$L_A \circ L_B : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$,

$$L_A \circ L_B\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = L_{AB}\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} 3 & 0 \\ 2 & -1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 3x \\ 2x-y \end{bmatrix},$$

$L_B \circ L_A : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$,

$$L_B \circ L_A\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = L_{BA}\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} 2 & 0 & 3 \\ 1 & -1 & 1 \\ 0 & 2 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2x+3z \\ x-y+z \\ 2y+z \end{bmatrix},$$

Sia

$$\mathcal{K} : \mathrm{M}_{m\times n}(\mathbb{K}) \longrightarrow \mathrm{Lin}(\mathbb{K}^n, \mathbb{K}^m) \qquad A \mapsto L_A.$$

L'applicazione $\mathcal{K}$ è lineare. Infatti

$$L_{A+B}(v) = (A+B)(v) = Av + Bv = L_A(v) + L_B(v),$$

per ogni $v \in \mathbb{K}^n$. Quindi $\mathcal{K}(A+B) = L_{A+B} = L_A + L_B = \mathcal{K}(A) + \mathcal{K}(B)$.

Se $\lambda \in \mathbb{K}$ e $A \in \mathrm{M}_{m\times n}(\mathbb{K})$, allora per ogni $v \in \mathbb{K}^n$ si ha

$$L_{\lambda A}(v) = (\lambda A)v = \lambda(Av) = \lambda(L_A)(v),$$

ovvero $\mathcal{K}(\lambda A) = \lambda \mathcal{K}(A)$. Tale applicazione è anche iniettiva, poiché $\mathcal{K}(A) = L_A$ è l'applicazione lineare nulla se e solamente se $\operatorname{Im} L_A = \{0_{K^m}\}$ se e solamente se

$$\mathcal{L}(A^1, \ldots, A^n) = \{0\},$$

quindi se e solamente se $A = 0$.

Vogliamo dimostrare che ogni applicazione lineare è della forma L_A, ovvero che $\mathcal{K}$ è suriettiva. Quindi $\mathcal{K}$ è un isomorfismo fra l'insieme delle matrici $M_{m\times n}(\mathbb{K})$ e $Lin(\mathbb{K}^n, \mathbb{K}^m)$.

Sia $T : \mathbb{K}^n \longrightarrow \mathbb{K}^m$ una applicazione lineare. Sia $(e_1, \ldots, e_n)$ la base canonica di $\mathbb{K}^n$. Se $X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{K}^n$, allora $X = x_1e_1 + \cdots + x_ne_n$, e

$$\begin{aligned} T(X) &= T(x_1e_1 + \cdots + x_ne_n) \\ &= x_1T(e_1) + \cdots + x_nT(e_n) \\ &= M_T \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}. \end{aligned}$$

dove $M_T = (T(e_1), \ldots, T(e_n)) \in M_{m\times n}(\mathbb{K})$. Quindi

$$T(X) = M_TX = L_{M_T}(X).$$

In particolare $\operatorname{Ker} T = \operatorname{Sol}(M_T|0)$ e $\dim \operatorname{Im} T = \operatorname{rg}(M_T)$ Riassumento abbiamo dimostrato il seguente risultato.

Proposizione 7.8. *Sia $T : \mathbb{K}^n \longrightarrow \mathbb{K}^m$ una applicazione lineare e sia $\{e_1, \ldots, e_n\}$ la base canonica di $\mathbb{K}^n$. Allora $T = L_{M_T}$ dove*

$$M_T = (T(e_1), \ldots, T(e_n)) \in M_{m\times n}(\mathbb{K}).$$

Inoltre, valgono le seguenti proprietà:

- $KerT = Sol(M_T|0)$;
- $ImT = \mathcal{L}(M_T^1, \ldots, M_T^n)$ *e* $\dim Im\, T = rg(M_T)$.

Corollario 7.9. *Sia $T : \mathbb{K}^n \longrightarrow \mathbb{K}^m$ una applicazione lineare e sia $M_T = (T(e_1), \ldots, T(e_n)) \in M_{m\times n}(\mathbb{K})$, dove $\{e_1, \ldots, e_n\}$ è la base canonica di $\mathbb{K}^n$. Allora*

a) T è iniettiva se e solamente se $rg(M_T) = n$;

b) T è suriettiva se e solamente se $rg(M_T) = m$;

c) T è biunivoca se e solamente se M_T è invertibile;

d) $b \in \operatorname{Im} T$ se e solamente se il sistema lineare $M_T X = b$ è compatibile.

Dimostrazione. Poiché $T = L_{M_T}$, applicando il Corolario 7.4, le prima tre proprietà sono di verifica immediata.

Sia $b \in \mathbb{K}^m$. $b \in \operatorname{Im} T$ se esiste $Y \in \mathbb{K}^n$ tale che $T(Y) = b$. Poiché $T(Y) = L_{M_T}(Y) = M_T Y = b$, si ha $b \in \operatorname{Im} T$ se e solamente se il sistema lineare $M_T X = b$ è compatibile. $\square$

Esempio 7.10. *Si consideri l'applicazione lineare $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ definita da*

$$T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} x+y+2z \\ 2x+y+3z \\ x+2y+3z \end{bmatrix},$$

Sia $\left\{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\right\}$ *la base canonica di $\mathbb{R}^3$. Allora*

$$T\left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, T\left(\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}, T\left(\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 2 \\ 3 \\ 3 \end{bmatrix},$$

e quindi

$$M_T = \begin{bmatrix} 1 & 1 & 2 \\ 2 & 1 & 3 \\ 1 & 2 & 3 \end{bmatrix}.$$

Poiché $\operatorname{rg}(M_T) = 2$, (verificare!) T non è iniettiva, rispettivamente suriettiva, rispettivamente biunivoca. Inoltre,

$$\operatorname{Ker} T = \operatorname{Sol}(M_T|0) = \mathcal{L}\left(\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}\right).$$

Sia $b = \begin{bmatrix} 2 \\ -1 \\ 7 \end{bmatrix} \in \mathbb{R}^3$. *Stabilire se* $b \in \operatorname{Im} T$ *è equivalente a stabilire se il sistema*

$$M_T X = b,$$

è compatibile. Poiché $\operatorname{rg}(M_T) = \operatorname{rg}(M_T|b)$ *(verificare!),* $b \in \operatorname{Im} T$.

Il prossimo risultato riguarda la corrispondenza $T \mapsto M_T$.

Proposizione 7.11. *L'applicazione*

$$\mathcal{H} : \operatorname{Lin}(\mathbb{K}^n, \mathbb{K}^m) \mapsto \mathrm{M}_{m\times n}(\mathbb{K}) \qquad T \mapsto M_T,$$

è una applicazione lineare, i.e.,

a) $M_{T+L} = M_T + M_L$;

b) $M_{\lambda T} = \lambda M_T$;

biunivoca la cui inversa è $\mathcal{K}$. *Inoltre:*

- *se* $T : \mathbb{K}^n \longrightarrow \mathbb{K}^p$ *e* $G : \mathbb{K}^p \longrightarrow \mathbb{K}^m$, *allora* $M_{G\circ T} = M_G M_T$;
- *se* $M_{\mathrm{Id}_{\mathbb{K}^n}} = \mathrm{Id}_n$;
- *se* $T : \mathbb{K}^n \longrightarrow \mathbb{K}^n$ *è invertibile, allora* $M_{T^{-1}} = M_T^{-1}$.

In particolare $\dim \operatorname{Lin}(\mathbb{K}^n, \mathbb{K}^m) = mn$.

Facoltativa. Abbiamo dimostrato che

$$\mathcal{K} : \mathrm{M}_{m\times n}(\mathbb{K}) \longrightarrow \operatorname{Lin}(\mathbb{K}^n, \mathbb{K}^m) \qquad A \mapsto L_A.$$

è un isomorfismo. È facile verifivare che l'applicazione

$$\mathcal{H} : \operatorname{Lin}(\mathbb{K}^n, \mathbb{K}^m) \longrightarrow \mathrm{M}_{m\times n}(\mathbb{K}) \qquad T \mapsto M_T,$$

è l'inversa di $\mathcal{K}$ e quindi è lineare.

Siano $T : \mathbb{K}^n \longrightarrow \mathbb{K}^p$ e $G : \mathbb{K}^p \longrightarrow \mathbb{K}^m$ lineari. Allora $G \circ T : \mathbb{K}^n \longrightarrow \mathbb{K}^m$ è lineare. Vogliamo dimostrare che $M_{G\circ T} = M_G M_T$. Poiché $T = L_{M_T}$ e $G = L_{M_G}$, allora

$$G \circ T = L_{M_G} \circ L_{M_T} = L_{M_G M_T}.$$

Quindi, tenendo in mente che l'applicazione $T \mapsto M_T$ è biunivoca, si ha

$$M_{G\circ T} = M_G M_T.$$

Infine, sia $T : \mathbb{K}^n \longrightarrow \mathbb{K}^n$ invertibile. Quindi M_T è invertibile. Inoltre

$$L_{M_T} \circ L_{M_T^{-1}} = L_{M_T(M_T)^{-1}} = \mathrm{Id}_{\mathbb{K}^n},$$

ovvero $T^{-1} = L_{M_T^{-1}}$. Infine, poiché $\mathcal{H}$ è un isomorfismo,

$$\dim \mathrm{Lin}(\mathbb{K}^n, \mathbb{K}^m) = \mathrm{M}_{m\times n}(\mathbb{K}) = mn.$$

□

7.12 Matrice associata ad una applicazione lineare

In questa sezione vogliamo associare ad una applicazione lineare una matrice generalizzando la costruzione che abbiamo introdotto nelle sezioni anteriori.

Sia $T : V \longrightarrow W$ una applicazione lineare. Siano $\mathcal{B} = \{v_1, \ldots, v_n\}$ e $\mathcal{C} = \{w_1, \ldots, w_m\}$ basi di V e W rispettivamente. Se $v \in V$, allora $v = x_1v_1 + \cdots + x_nv_n$ e $[v]_{\mathcal{B}} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ sono le coordinate di v rispetto a $\mathcal{B}$.

Analogamente se $w \in W$, allora $w = y_1w_1 + \ldots + y_mw_m$ e $[w]_{\mathcal{C}} = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}$ sono le coordinate di w rispetto alla base $\mathcal{C}$. Vogliamo calcolare le coordinate di $T(v)$ rispetto alla base $\mathcal{C}$. Poiché $\mathcal{F}_{\mathcal{C}} : W \longrightarrow \mathbb{K}^n$ è lineare, si ha

$$\begin{aligned} [T(v)]_{\mathcal{C}} &= [T(x_1v_1 + \cdots + x_nv_n)]_{\mathcal{C}} \\ &= [x_1T(v_1) + \cdots + x_nT(v_n)]_{\mathcal{C}} \\ &= x_1[T(v_1)]_{\mathcal{C}} + \cdots + x_n[T(v_n)]_{\mathcal{C}} \\ &= \mathcal{M}_{\mathcal{C},\mathcal{B}}(T)[v]_{\mathcal{B}}, \end{aligned}$$

dove $\mathcal{M}_{\mathcal{C},\mathcal{B}}(T) = ([T(v_1)]_{\mathcal{C}}, \ldots, [T(v_n)]_{\mathcal{C}}) \in M_{m\times n}(\mathbb{K})$. $\mathcal{M}_{\mathcal{C},\mathcal{B}}(T)$ è chiamata la *matrice associata a T rispetto alla basi $\mathcal{B}$ in partenza e $\mathcal{C}$ in arrivo.* La matrice $\mathcal{M}_{\mathcal{C},\mathcal{B}}(T)$ è l'unica matrice di ordine $m \times n$, dove $m = \dim W$ e $n = \dim V$, a coefficienti in $\mathbb{K}$ che verifica

$$[T(v)]_{\mathcal{C}} = \mathcal{M}_{\mathcal{C},\mathcal{B}}(T)[v]_{\mathcal{B}},$$

per ogni $v \in V$. In maniera equivalente, il seguente diagramma

$$\begin{tikzcd} V \arrow[r, "T"] \arrow[d, "\mathcal{F}_{\mathcal{B}}"] & W \arrow[d, "\mathcal{F}_{\mathcal{C}}"] \\ \mathbb{K}^n \arrow[r, "L_{\mathcal{M}_{\mathcal{C},\mathcal{B}}(T)}"'] & \mathbb{K}^m \end{tikzcd}$$

è commutativo, i.e., $\mathcal{F}_{\mathcal{C}} \circ T = L_{\mathcal{M}_{\mathcal{C},\mathcal{B}}(T)} \circ \mathcal{F}_{\mathcal{B}}$.

Osservazione 7.13. *Sia $T : \mathbb{K}^n \longrightarrow \mathbb{K}^m$. Sia $\mathcal{C}$ la base canonica di $\mathbb{K}^n$ e $\mathcal{C}'$ la base canonica di $\mathbb{K}^m$. Allora $M_T = (T(e_1), \ldots, T(e_n)) = \mathcal{M}_{\mathcal{C}',\mathcal{C}}(T)$*

Proposizione 7.14. *Sia $T : V \longrightarrow W$ e siano $\mathcal{B}$ e $\mathcal{C}$ basi di V e W rispettivamente. Allora $\dim ImT = rg(\mathcal{M}_{\mathcal{C},\mathcal{B}}(T))$ e quindi $\dim KerT = \dim V - rg(\mathcal{M}_{\mathcal{C},\mathcal{B}}(T))$. Inoltre*

- $\mathcal{F}_{\mathcal{C}}(Im\,T) = Im\,L_{\mathcal{M}_{\mathcal{C},\mathcal{B}}(T)}$;
- $\mathcal{F}_{\mathcal{B}}(Ker\,T) = Ker\,L_{\mathcal{M}_{\mathcal{C},\mathcal{B}}(T)}$;

Facoltativa. Sia $v \in V$. $T(v) = 0$ se e solamente se $[T(v)]_{\mathcal{C}} = 0$. Poiché

$$[T(v)]_{\mathcal{C}} = \mathcal{M}_{\mathcal{C},\mathcal{B}}(T)[v]_{\mathcal{B}},$$

si ha $T(v) = 0$ se e solamente se $\mathcal{M}_{\mathcal{C},\mathcal{B}}(T)[v]_{\mathcal{B}} = 0$. Quindi abbiamo dimostrato che

$$\operatorname{Ker} T = \{v \in V : T(v) = 0\} = \{v \in V : [v]_{\mathcal{B}} \in \operatorname{Sol}(\mathcal{M}_{\mathcal{C},\mathcal{B}}(T)|0)\}.$$

In particolare $\mathcal{F}_{\mathcal{B}}(\operatorname{Ker} T) = \operatorname{Sol}(\mathcal{M}_{\mathcal{C},\mathcal{B}}(T)|0)$. Poiché $\mathcal{F}_{\mathcal{B}}$ è un isomorfismo si ha $\dim \operatorname{Ker} T = \dim \operatorname{Sol}(\mathcal{M}_{\mathcal{C},\mathcal{B}}(T)|0) = \dim V - \operatorname{rg}(\mathcal{M}_{\mathcal{C},\mathcal{B}}(T))$. Applicando il Teorema della dimensione si ha $\dim ImT = rg(\mathcal{M}_{\mathcal{C},\mathcal{B}}(T))$.

Per la seconda parte osserviamo che

$$\operatorname{Ker} L_{\mathcal{M}_{\mathcal{C},\mathcal{B}}(T)} = \operatorname{Sol}(\mathcal{M}_{\mathcal{C},\mathcal{B}}(T)|0),$$

quindi $\mathcal{F}_{\mathcal{B}}(Ker\,T) = Ker\,L_{\mathcal{M}_{\mathcal{C},\mathcal{B}}(T)}$.

Sia $w \in W$. $w \in Im\,T$ se e solamente se esiste $v \in V$ tale che $T(v) = w$. Quindi se e solamente se $[T(v)]_{\mathcal{C}} = [w]_{\mathcal{C}}$. Poiché

$$[T(v)]_{\mathcal{C}} = \mathcal{M}_{\mathcal{C},\mathcal{B}}(T)[v]_{\mathcal{B}},$$

si ha $w \in Im\,T$ se e solamente se $[w]_{\mathcal{C}} \in Im\,L_{\mathcal{M}_{\mathcal{C},\mathcal{B}}(T)}$, ovvero $\mathcal{F}_{\mathcal{C}}(Im\,T) = Im\,L_{\mathcal{M}_{\mathcal{C},\mathcal{B}}(T)}$. □

Applicando il risultato anteriore ed il Corollario 6.16 si ha il seguente risultato.

Corollario 7.15. *Sia $T : V \longrightarrow W$ e siano $\mathcal{B}$ e $\mathcal{C}$ basi di V e W rispettivamente. Allora:*

- *T è iniettiva se e solamente se $rg(\mathcal{M}_{\mathcal{C},\mathcal{B}}(T)) = \dim V$;*
- *T è suriettiva se e solamente se $rg(\mathcal{M}_{\mathcal{C},\mathcal{B}}(T)) = \dim W$;*
- *T è biunivoca se e solamente se $\mathcal{M}_{\mathcal{C},\mathcal{B}}(T)$ è invertibile.*

Esempio 7.16. *Sia $A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$ e sia Sia $T : M_{2\times 2}(\mathbb{R}) \longrightarrow M_{2\times 2}(\mathbb{R})$ così definita:*

$$T(X) = X - \mathrm{Tr}(X)A.$$

L'applicazione T è lineare (verificare). Sia

$$\mathcal{B} = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\},$$

una base di $M_{2\times 2}(\mathbb{R})$. Allora

$$T\left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \right) = \begin{bmatrix} 1 & -1 \\ -1 & -1 \end{bmatrix}, T\left(\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \right) = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

$$T\left(\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \right) = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, T\left(\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right) = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}.$$

Poiché

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = a \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + b \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + c \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + d \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix},$$

si ha

$$\mathcal{F}_{\mathcal{B}} : M_{2\times 2}(\mathbb{R}) \longrightarrow \mathbb{R}^4 \qquad \begin{bmatrix} a & b \\ c & d \end{bmatrix} \mapsto \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix},$$

e quindi

$$\mathcal{M}_{\mathcal{B},\mathcal{B}}(T) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & -1 \\ -1 & 0 & 1 & -1 \\ -1 & 0 & 0 & 0 \end{bmatrix}.$$

Poiché $\mathrm{rg}(\mathcal{M}_{\mathcal{B},\mathcal{B}}(T)) = 3$, *allora* $\dim \mathrm{Im}\, T = 3$ *e* $\dim \mathrm{Ker}\, T = 1$

Esempio 7.17. *Sia $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ l'applicazione lineare definita da:*

$$T\left(\begin{bmatrix}1\\0\\1\end{bmatrix}\right)=\begin{bmatrix}1\\0\\2\end{bmatrix}, T\left(\begin{bmatrix}0\\2\\1\end{bmatrix}\right)=\begin{bmatrix}1\\2\\4\end{bmatrix}, T\left(\begin{bmatrix}1\\1\\1\end{bmatrix}\right)=\begin{bmatrix}1\\1\\3\end{bmatrix}.$$

Allora

$$\mathcal{M}_{\mathcal{C},\mathcal{B}}(T)=\begin{bmatrix}1&1&1\\0&2&1\\2&4&3\end{bmatrix},$$

è la matrice associata a T rispetto alle basi $\mathcal{B}=\left\{\begin{bmatrix}1\\0\\1\end{bmatrix},\begin{bmatrix}0\\2\\1\end{bmatrix},\begin{bmatrix}1\\1\\1\end{bmatrix}\right\}$ in partenza e la base canonica $\mathcal{C}$ in arrivo. Poiché $\mathcal{M}_{\mathcal{C},\mathcal{B}}(T)$ ha rango 2, allora T non è iniettiva. Per calcolare il nucleo possiamo procedere come segue. Poiché

$$T(X)=\mathcal{M}_{\mathcal{C},\mathcal{B}}(T)[X]_{\mathcal{B}},$$

si ha $X \in \operatorname{Ker} T$ se e solamente se $T(X)=\mathcal{M}_{\mathcal{C},\mathcal{B}}(T)[X]_{\mathcal{B}}=0$. Quindi $\mathrm{Sol}(\mathcal{M}_{\mathcal{C},\mathcal{B}}(T)|0)$ descrive le coordinate dei vettori $X \in \mathbb{R}^3$ rispetto alla base $\mathcal{B}$ tali che $T(X)=0$. Applicando il metodo di Gauss e il metodo della risoluzione all'indietro, si ha

$$\mathrm{Sol}((\mathcal{M}_{\mathcal{C},\mathcal{B}}(T)|0))=\mathcal{L}\left(\begin{bmatrix}1\\1\\-2\end{bmatrix}\right).$$

Quindi una base per il $\operatorname{Ker} T$ è dato dal vettore

$$1\begin{bmatrix}1\\0\\1\end{bmatrix}+1\begin{bmatrix}0\\2\\1\end{bmatrix}-2\begin{bmatrix}1\\1\\1\end{bmatrix}=\begin{bmatrix}-1\\0\\0\end{bmatrix}.$$

Un'altro metodo per calcolare il nucelo è ricavare

$$\begin{aligned}T\left(\begin{bmatrix}x\\y\\z\end{bmatrix}\right)&=\mathcal{M}_{\mathcal{C},\mathcal{B}}(T)\begin{bmatrix}x\\y\\z\end{bmatrix}_{\mathcal{B}}\\&=\begin{bmatrix}1&1&1\\0&2&1\\2&4&3\end{bmatrix}\begin{bmatrix}-x-y+2z\\-x+z\\2x+y-2z\end{bmatrix}\\&=\begin{bmatrix}z\\y\\y+2z\end{bmatrix},\end{aligned}$$

da cui segue facilmente che $\operatorname{Ker} T = \mathcal{L}\left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right)$.

È possibile raffinare i risultati anteriori provando il seguente risultato.

Proposizione 7.18. *Siano* V *e* W *spazi vettoriali su* $\mathbb{K}$ *e siano* $\mathcal{B} = \{v_1, \ldots, v_n\}$ *e* $\mathcal{C} = \{w_1, \ldots, w_m\}$ *basi di* V *e* W *rispettivamente. L'applicazione*

$$\mathcal{M} : \operatorname{Lin}(V, W) \longrightarrow M_{m\times n}(\mathbb{K}), \qquad T \mapsto \mathcal{M}_{\mathcal{C},\mathcal{B}}(T),$$

è un isomorfismo di spazi vettoriali. In particolare $\dim \operatorname{Lin}(V, W) = \dim V \dim W = mn$.

Facoltativa. Dimostriamo che $\mathcal{M}$ è lineare. Siano $T, L \in \operatorname{Lin}(V, W)$. La matrice $\mathcal{M}_{\mathcal{C},\mathcal{B}}(T)$, rispettivamente $\mathcal{M}_{\mathcal{C},\mathcal{B}}(L)$, è l'unica matrice di formato $m \times n$ tale che

$$[T(v)]_{\mathcal{C}} = \mathcal{M}_{\mathcal{C},\mathcal{B}}(T)[v]_{\mathcal{B}},$$

rispettivamente

$$[L(v)]_{\mathcal{C}} = \mathcal{M}_{\mathcal{C},\mathcal{B}}(L)[v]_{\mathcal{B}},$$

per ogni $v \in V$. Quindi

$$\begin{aligned} [(T+L)(v)]_{\mathcal{C}} &= [T(v)]_{\mathcal{C}} + [L(v)]_{\mathcal{C}} = \mathcal{M}_{\mathcal{C},\mathcal{B}}(T)[v]_{\mathcal{B}} + \mathcal{M}_{\mathcal{C},\mathcal{B}}(L)[v]_{\mathcal{B}} \\ &= (\mathcal{M}_{\mathcal{C},\mathcal{B}}(T) + \mathcal{M}_{\mathcal{C},\mathcal{B}}(L))[v]_{\mathcal{B}}, \end{aligned}$$

ovvero $\mathcal{M}(T+L) = \mathcal{M}(T) + \mathcal{M}(L)$. Analogamente vale $\mathcal{M}(\lambda T) = \lambda \mathcal{M}(T)$. La suriettività può essere dimostrata come segue. Sia $A \in \mathrm{M}_{m\times n}(\mathbb{K})$ e sia $T : V \longrightarrow W$ l'unica applicazione lineare definita da

$$T(v_1) = \sum_{i=1}^{m} a_{i1} w_i, \ldots, T(v_n) = \sum_{i=1}^{m} a_{in} w_i.$$

Quindi la i-esima colonna della matrice $\mathcal{M}_{\mathcal{C},\mathcal{B}}(T)$ è, per definizione, $[T(v_i)]_{\mathcal{C}} = \begin{bmatrix} a_{1i} \\ \vdots \\ a_{mi} \end{bmatrix}$, per $i = 1, \ldots, n$, ovvero la i-esima colonna della matrice A, da cui segue che $\mathcal{M}_{\mathcal{C},\mathcal{B}}(T) = A$, dimostrando che $\mathcal{M}$ è suriettiva.

Poiché $T = 0$ se e solamente se $T(v) = 0$ per ogni $v \in V$, quindi se e solamente se

$$[T(v)]_{\mathcal{C}} = \mathcal{M}_{\mathcal{C},\mathcal{B}}(T)[v]_{\mathcal{B}} = 0$$

per ogni $v \in V$. Quindi se e solamente se $\mathcal{M}_{\mathcal{C},\mathcal{B}}(T) = 0$. □

Come nel caso di $\text{Lin}(\mathbb{K}^n, \mathbb{K}^m)$, la corrispondenza biunivoca sopra definita verifica altre proprietà.

Proposizione 7.19. *Siano $L, T : V \longrightarrow W$, $H : W \longrightarrow Z$ e $G : N \longrightarrow V$. Siano $\mathcal{B}$ una base di V, $\mathcal{C}$ una base di W, $\mathcal{H}$ una base di Z ed infine $\mathcal{X}$ una base di N. Allora:*

a) $\mathcal{M}_{\mathcal{C},\mathcal{X}}(T \circ G) = \mathcal{M}_{\mathcal{C},\mathcal{B}}(T)\mathcal{M}_{\mathcal{B},\mathcal{X}}(G)$;

b) $\mathcal{M}_{\mathcal{H},\mathcal{B}}(H \circ (L+T)) = \mathcal{M}_{\mathcal{H},\mathcal{C}}(H)(\mathcal{M}_{\mathcal{C},\mathcal{B}}(L) + \mathcal{M}_{\mathcal{C},\mathcal{B}}(T))$;

c) $\mathcal{M}_{\mathcal{C},\mathcal{X}}((L+T) \circ G) = (\mathcal{M}_{\mathcal{C},\mathcal{B}}(L) + \mathcal{M}_{\mathcal{C},\mathcal{B}}(T))\mathcal{M}_{\mathcal{B},\mathcal{X}}(G)$;

d) sia $Id_V : V \longrightarrow V$. Allora $\mathcal{M}_{\mathcal{B},\mathcal{B}}(Id_V) = \text{Id}_{\dim V}$;

e) se $T : V \longrightarrow W$ è invertibile, allora $\mathcal{M}_{\mathcal{B},\mathcal{C}}(T^{-1}) = \mathcal{M}_{\mathcal{C},\mathcal{B}}(T)^{-1}$.

Facoltativa. Dimostriamo solo il punto a). Sia $n \in N$. Allora

$$\begin{aligned}[T(G(n))]_{\mathcal{C}} &= \mathcal{M}_{\mathcal{C},\mathcal{B}}(T)[G(n)]_{\mathcal{B}} \\ &= \mathcal{M}_{\mathcal{C},\mathcal{B}}(T)\mathcal{M}_{\mathcal{B},\mathcal{X}}(G)[n]_{\mathcal{X}}\end{aligned}$$

da cui segue

$$\mathcal{M}_{\mathcal{C},\mathcal{X}}(T \circ G) = \mathcal{M}_{\mathcal{C},\mathcal{B}}(T)\mathcal{M}_{\mathcal{B},\mathcal{X}}(G).$$

□

7.20 Matrice del cambiamento di base

Sia V uno spazio vettoriale su $\mathbb{K}$. Se $\mathcal{B} = \{v_1, \ldots, v_n\}$ è una base di V, allora $v = x_1v_1 + \cdots + x_nv_n$. Se $\mathcal{C} = \{w_1, \ldots, w_n\}$ è un altra base di V, allora $v = x_1'w_1 + \cdots + x_n'w_n$. Mi domando se esiste una relazione fra i vettori $[v]_{\mathcal{B}}$ e $[v]_{\mathcal{C}}$.

Poiché $\mathcal{F}_{\mathcal{B}}$ è un isomorfismo lineare si ha

$$[v]_{\mathcal{B}} = [x'_1w_1 + \cdots + x'_nw_n]_{\mathcal{B}} = x_1'[w_1]_{\mathcal{B}} + \cdots + x_n'[w_n]_{\mathcal{B}} = \mathcal{M}(\mathcal{B}, \mathcal{C})[v]_{\mathcal{C}},$$

dove

$$\mathcal{M}(\mathcal{B}, \mathcal{C}) = ([w_1]_{\mathcal{B}}, \ldots, [w_n]_{\mathcal{B}}) \in M_{n \times n}(\mathbb{K}),$$

si chiama *matrice del cambiamento di base* da $\mathcal{B}$ a $\mathcal{C}$. La matrice del cambiamento di base è l'unica matrice $\mathcal{M}(\mathcal{B}, \mathcal{C})$ tale che per ogni $v \in V$, si ha

$$[v]_{\mathcal{B}} = \mathcal{M}(\mathcal{B}, \mathcal{C})[v]_{\mathcal{C}}.$$

È facile verificare che $\mathcal{M}(\mathcal{B},\mathcal{C}) = \mathcal{M}_{\mathcal{B},\mathcal{C}}(Id_V)$ da cui segue che la matrice del cambiamento di base è invertibile. Un altra maniera per provare cje $\mathcal{M}(\mathcal{B},\mathcal{C})$ è invertibile è la seguente. Poiché $w_1,\dots,w_n$ formano una base di V, anche i vettori $[w_1]_{\mathcal{B}},\dots,[w_n]_{\mathcal{B}}$ formano una base di $\mathbb{K}^n$. Quindi $\mathcal{M}(\mathcal{B},\mathcal{C})$ è invertibile. Vediamo altre proprietà della matrice di cambiamento di base.

Proposizione 7.21. *Siano $\mathcal{B}$, $\mathcal{C}$ e $\mathcal{D}$ basi di V. Allora*

- $\mathcal{M}(\mathcal{B},\mathcal{B}) = Id_n$*;*
- $\mathcal{M}(\mathcal{B},\mathcal{C})\mathcal{M}(\mathcal{C},\mathcal{D}) = \mathcal{M}(\mathcal{B},\mathcal{D})$*;*
- $\mathcal{M}(\mathcal{B},\mathcal{C})$ *è invertibile e la sua inversa* $(\mathcal{M}(\mathcal{B},\mathcal{C}))^{-1} = \mathcal{M}(\mathcal{C},\mathcal{B})$.

Facoltativa. Sia $\mathcal{B}$ una base poiché $[v_1]_{\mathcal{B}} = e_1,\dots,[v_n]_{\mathcal{B}} = e_n$, la matrice $\mathcal{M}(\mathcal{B},\mathcal{B}) = Id_n$.

Sia $v \in V$. Allora

$$\begin{aligned}[v]_{\mathcal{B}} &= \mathcal{M}(\mathcal{B},\mathcal{C})[v]_{\mathcal{C}} = \mathcal{M}(\mathcal{B},\mathcal{C})\mathcal{M}(\mathcal{C},\mathcal{D})[v]_{\mathcal{D}} \\ &= \mathcal{M}(\mathcal{B},\mathcal{D})[v]_{\mathcal{D}}.\end{aligned}$$

Poiché vale per ogni $v \in V$, si ha $\mathcal{M}(\mathcal{B},\mathcal{D}) = \mathcal{M}(\mathcal{B},\mathcal{C})\mathcal{M}(\mathcal{C},\mathcal{D})$. In particolare

$$\mathcal{M}(\mathcal{B},\mathcal{C})\mathcal{M}(\mathcal{C},\mathcal{B}) = \mathcal{M}(\mathcal{B},\mathcal{B}) = Id_n.$$

□

Esempio 7.22. *Siano*

$$\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}, \mathcal{B}' = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$$

basi di $\mathbb{R}^3$. Allora

$$\mathcal{M}(\mathcal{B},\mathcal{B}') = \left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}_{\mathcal{B}} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}_{\mathcal{B}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}_{\mathcal{B}} \right).$$

Poiché

$$\begin{aligned}\begin{bmatrix} x \\ y \\ z \end{bmatrix}_{\mathcal{B}} &= \mathrm{Sol}\left(\left[\begin{array}{ccc|c} 1 & 0 & 1 & x \\ 0 & 2 & 1 & y \\ 1 & 1 & 1 & z \end{array}\right] \right) \\ &= \begin{bmatrix} -x-y+2z \\ -x+z \\ 2x+y-2z \end{bmatrix},\end{aligned}$$

si ha

$$\mathcal{M}(\mathcal{B},\mathcal{B}') = \begin{bmatrix} -1 & -2 & 0 \\ -1 & -1 & 0 \\ 2 & 3 & 1 \end{bmatrix}.$$

Analogamente

$$\mathcal{M}(\mathcal{B}',\mathcal{B}) = \left(\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}_{\mathcal{B}'} \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}_{\mathcal{B}'} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}_{\mathcal{B}'} \right).$$

Poiché

$$\begin{aligned} \begin{bmatrix} x \\ y \\ z \end{bmatrix}_{\mathcal{B}'} &= \mathrm{Sol}\left(\left[\begin{array}{ccc|c} 1 & 1 & 1 & x \\ 0 & 1 & 1 & y \\ 0 & 0 & 1 & z \end{array}\right] \right) \\ &= \begin{bmatrix} x-y \\ y-z \\ z \end{bmatrix}, \end{aligned}$$

si ha

$$\mathcal{M}(\mathcal{B}',\mathcal{B}) = \begin{bmatrix} 1 & -2 & 0 \\ -1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}.$$

Esempio 7.23. *Sia* $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$ *una base di* $\mathbb{R}^3$ *e sia* $\mathcal{C}$ *la base canonica di* $\mathbb{R}^3$*. Allora*

$$\mathcal{M}(\mathcal{C},\mathcal{B}) = \left(\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}_{\mathcal{C}} \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}_{\mathcal{C}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}_{\mathcal{C}} \right) = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

Invece, poiché

$$\begin{aligned} \begin{bmatrix} x \\ y \\ z \end{bmatrix}_{\mathcal{B}} &= \mathrm{Sol}\left(\left[\begin{array}{ccc|c} 1 & 0 & 1 & x \\ 0 & 2 & 1 & y \\ 1 & 1 & 1 & z \end{array}\right] \right) \\ &= \begin{bmatrix} -x-y+2z \\ -x+z \\ 2x+y-2z \end{bmatrix}, \end{aligned}$$

si ha

$$\mathcal{M}(\mathcal{B},\mathcal{C}) = \begin{bmatrix} -1 & -1 & 2 \\ -1 & 0 & 1 \\ 2 & 1 & -2 \end{bmatrix} = \mathcal{M}(\mathcal{C},\mathcal{B})^{-1}.$$

Teorema 7.24. *Sia $T : V \longrightarrow W$ una applicazione lineare. Siano $\mathcal{B}$, $\mathcal{B}'$ e $\mathcal{C}$ e $\mathcal{C}'$ basi di V e W rispettivamente. Allora*

$$\mathcal{M}_{\mathcal{C}',\mathcal{B}'}(T) = \mathcal{M}(\mathcal{C}',\mathcal{C})\mathcal{M}_{\mathcal{C},\mathcal{B}}(T)\mathcal{M}(\mathcal{B},\mathcal{B}').$$

Facoltativa. Sia $v \in V$. Allora

$$[T(v)]_{\mathcal{C}'} = \mathcal{M}(\mathcal{C}',\mathcal{C})[T(v)]_{\mathcal{C}}$$

Adesso, $[T(v)]_{\mathcal{C}} = \mathcal{M}_{\mathcal{C},\mathcal{B}}(T)[v]_{\mathcal{B}}$, e $[v]_{\mathcal{B}} = \mathcal{M}(\mathcal{B},\mathcal{B}')[v]_{\mathcal{B}'}$, da cui segue che per ogni $v \in V$ si ha

$$\begin{aligned}[T(v)]_{\mathcal{C}'} &= \mathcal{M}(\mathcal{C}',\mathcal{C})[T(v)]_{\mathcal{C}} \\ &= \big(\mathcal{M}(\mathcal{C}',\mathcal{C})\mathcal{M}_{\mathcal{C},\mathcal{B}}(T)\mathcal{M}(\mathcal{B},\mathcal{B}')\big)\,[v]_{\mathcal{B}'}.\end{aligned}$$

Per definizione di matrice associata a T rispetto alle basi $\mathcal{B}'$ in partenza e $\mathcal{C}'$ in arrivo, si ha

$$\begin{aligned}[T(v)]_{\mathcal{C}'} &= \mathcal{M}_{\mathcal{C}',\mathcal{B}'}(T)[v]_{\mathcal{B}'} \\ &= \big(\mathcal{M}(\mathcal{C}',\mathcal{C})\mathcal{M}_{\mathcal{C},\mathcal{B}}(T)\mathcal{M}(\mathcal{B},\mathcal{B}')\big)\,[v]_{\mathcal{B}'}.\end{aligned}$$

per ogni $v \in V$. Quindi

$$\mathcal{M}_{\mathcal{C}',\mathcal{B}'}(T) = \mathcal{M}(\mathcal{C}',\mathcal{C})\mathcal{M}_{\mathcal{C},\mathcal{B}}(T)\mathcal{M}(\mathcal{B},\mathcal{B}'),$$

concludendo la dimostrazione. □

Corollario 7.25. *Sia $T : V \longrightarrow W$ una applicazione lineare. Siano $\mathcal{B}$ e $\mathcal{B}'$ basi di V, rispettivamente $\mathcal{C}$ e $\mathcal{C}'$ basi di W. Allora*

$$\mathcal{M}_{\mathcal{C}',\mathcal{B}}(T) = \mathcal{M}(\mathcal{C}',\mathcal{C})\mathcal{M}_{\mathcal{C},\mathcal{B}}(T),$$

rispettivamente,

$$\mathcal{M}_{\mathcal{C},\mathcal{B}'}(T) = \mathcal{M}_{\mathcal{C},\mathcal{B}}(T)\mathcal{M}(\mathcal{B},\mathcal{B}').$$

Dimostrazione. Per il Teorema anteriore si ha

$$\mathcal{M}_{\mathcal{C}',\mathcal{B}}(T) = \mathcal{M}(\mathcal{C}',\mathcal{C})\mathcal{M}_{\mathcal{C},\mathcal{B}}(T)\mathcal{M}(\mathcal{B},\mathcal{B}),$$

rispettivamente,

$$\mathcal{M}_{\mathcal{C},\mathcal{B}'}(T) = \mathcal{M}(\mathcal{C},\mathcal{C})\mathcal{M}_{\mathcal{C},\mathcal{B}}(T)\mathcal{M}(\mathcal{B},\mathcal{B}').$$

Tenendo in mente che $\mathcal{M}(\mathcal{B},\mathcal{B}) = \mathrm{Id}_{\dim V}$, rispettivamente $\mathcal{M}(\mathcal{C},\mathcal{C}) = \mathrm{Id}_{\dim W}$, si ha la tesi. □

Esempio 7.26. *Sia $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ l'applicazione lineare definita da:*

$$T\left(\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, T\left(\begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}, T\left(\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}.$$

Allora

$$\mathcal{M}_{\mathcal{C},\mathcal{B}}(T) = \begin{bmatrix} 1 & 0 & -2 \\ 1 & 2 & 0 \\ 0 & 1 & 1 \end{bmatrix},$$

è la matrice associata a T rispetto alle basi $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$ in partenza e la base canonica $\mathcal{C}$ in arrivo. Poiché

$$T(X) = \mathcal{M}_{\mathcal{C},\mathcal{C}}(T)X,$$

calcolare $T(X)$ è la stessa cosa che calcolare $\mathcal{M}_{\mathcal{C},\mathcal{C}}(T)$. Applicando il corollario anteriore, si ha

$$\mathcal{M}_{\mathcal{C},\mathcal{C}}(T) = \mathcal{M}_{\mathcal{C},\mathcal{B}}(T)\mathcal{M}(\mathcal{B},\mathcal{C}).$$

Nell'esempio 7.23 abbiamo calcolato

$$\mathcal{M}(\mathcal{B},\mathcal{C}) = \begin{bmatrix} -1 & -1 & 2 \\ -1 & 0 & 1 \\ 2 & 1 & -2 \end{bmatrix},$$

da cui segue che

$$\begin{aligned} \mathcal{M}_{\mathcal{C},\mathcal{C}}(T) &= \begin{bmatrix} 1 & 0 & -2 \\ 1 & 2 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} -1 & -1 & 2 \\ -1 & 0 & 1 \\ 2 & 1 & -2 \end{bmatrix} \\ &= \begin{bmatrix} -5 & -3 & 6 \\ -3 & -1 & 4 \\ 1 & 1 & -1 \end{bmatrix}, \end{aligned}$$

quindi

$$T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} -5 & -3 & 6 \\ -3 & -1 & 4 \\ 1 & 1 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -5x - 3y + 6z \\ -3x - y + 4z \\ x + y - z \end{bmatrix}.$$

Esempio 7.27. *Sia $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ l'applicazione lineare definita da:*

$$T\left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, T\left(\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}, T\left(\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} -2 \\ 5 \\ 2 \end{bmatrix}.$$

Allora

$$\mathcal{M}_{\mathcal{B},\mathcal{B}}(T) = \left(\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}_{\mathcal{B}} \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}_{\mathcal{B}} \begin{bmatrix} -2 \\ 5 \\ 2 \end{bmatrix}_{\mathcal{B}}\right).$$

Poiché

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix}_{\mathcal{B}} = \begin{bmatrix} x-y \\ y-z \\ z \end{bmatrix},$$

si ha

$$\mathcal{M}_{\mathcal{B},\mathcal{B}}(T) = \begin{bmatrix} 1 & 0 & -7 \\ 1 & 2 & 3 \\ -1 & -1 & 2 \end{bmatrix}.$$

Applicando il Teorema 7.24, si ha

$$\mathcal{M}_{\mathcal{C},\mathcal{C}}(T) = \mathcal{M}(\mathcal{C},\mathcal{B})\mathcal{M}_{\mathcal{B},\mathcal{B}}(T)\mathcal{M}(\mathcal{B},\mathcal{C}).$$

Poiché

$$\mathcal{M}(\mathcal{C},\mathcal{B}) = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}, \qquad \mathcal{M}(\mathcal{B},\mathcal{C}) = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix},$$

si ha

$$\begin{aligned} \mathcal{M}_{\mathcal{C},\mathcal{C}}(T) &= \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -7 \\ -1 & 2 & 3 \\ 1 & -1 & 2 \end{bmatrix} \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & -3 \\ 0 & 1 & 4 \\ -1 & 0 & 3 \end{bmatrix}. \end{aligned}$$

Quindi

$$T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} 1 & 0 & -3 \\ 0 & 1 & 4 \\ -1 & 0 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x-3z \\ y+4z \\ -x+3z \end{bmatrix}.$$

Corollario 7.28. *Sia $T : V \longrightarrow V$ una applicazione lineare. Siano $\mathcal{B}$ e $\mathcal{C}$ basi di V. Allora*

$$\mathcal{M}_{\mathcal{C},\mathcal{C}}(T) = (\mathcal{M}(\mathcal{B},\mathcal{C}))^{-1}\mathcal{M}_{\mathcal{B},\mathcal{B}}(T)\mathcal{M}(\mathcal{B},\mathcal{C}),$$

Dimostrazione. Applicando il Teorema anteriore si ha

$$\mathcal{M}_{\mathcal{C},\mathcal{C}}(T) = \mathcal{M}(\mathcal{C},\mathcal{B})\mathcal{M}_{\mathcal{B},\mathcal{B}}(T)\mathcal{M}(\mathcal{B},\mathcal{C}).$$

Poiché $\mathcal{M}(\mathcal{C},\mathcal{B}) = \mathcal{M}(\mathcal{B},\mathcal{C})^{-1}$, si ha la tesi. $\square$

Il risultato anteriore suggerisce la seguente definizione.

Definizione 7.29. *Siano $A, B \in \mathrm{M}_{n\times n}(\mathbb{K})$. Diremo che A e B sono* simili *se esiste una matrice invertibile $P \in \mathrm{M}_{n\times n}(\mathbb{K})$ tale che $A = P^{-1}BP$.*

La relazione $A \sim B$, se A e B sono matrici simili, è una relazione di equivalenza, poiché $A \sim A$; $A \sim B$ allora $B \sim A$; infine se $A \sim B$ e $B \sim C$, allora $A \sim C$. Se $A \sim B$, allora $\mathrm{Tr}(A) = \mathrm{Tr}(B)$, $\det A = \det B$ e $\mathrm{rg}(A) = \mathrm{rg}(B)$. Tali condizioni sono necessarie ma non suffcienti (verificare).

Il corollario anteriore dimostra che se $T : V \longrightarrow V$ è una applicazione lineare, $\mathcal{B}$ e $\mathcal{C}$ basi di V, allora le matrici $\mathcal{M}_{\mathcal{C},\mathcal{C}}(T)$ e $\mathcal{M}_{\mathcal{B},\mathcal{B}}(T)$ sono matrici simili. Viceversa, due matrici simili A e B simili sono matrici associate ad una applicazione lineare $T : V \longrightarrow V$ rispetto a due basi di V.

Infatti, siano $A = (a_{ij})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}}$ e $B = (b_{ij})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}}$ matrici simili. Esiste $P \in M_{n\times n}(\mathbb{K})$ invertibile tale che $A = P^{-1}BP$. Sia $\mathcal{B} = \{v_1, \ldots, v_n\}$ una base di V. Sia $T : V \longrightarrow V$ l'unico endomorfimo che verifica $T(v_i) = \sum_{m=1}^{n} a_{mi}v_m$, per $i = 1, \ldots, n$. Vogliamo dimostrare che $\mathcal{M}_{\mathcal{B},\mathcal{B}}(T) = A$. Infatti la $k-$esima colonna di $\mathcal{M}_{\mathcal{B},\mathcal{B}}(T)$ è per definizione $[T(v_i)]_{\mathcal{B}} = \begin{bmatrix} a_{1i} \\ \vdots \\ a_{ni} \end{bmatrix}$, ovvero $\mathcal{M}_{\mathcal{B},\mathcal{B}}(T)^k = A^k$. Quindi $\mathcal{M}_{\mathcal{B},\mathcal{B}}(T) = A$.

Sia $\mathcal{C} = \{w_1, \ldots, w_n\}$ dove $w_i = \sum_{m=1}^{n} p_{mi}v_m$, $i = 1, \ldots, n$, per $i = 1, \ldots, n$. L'insieme $\mathcal{C} = \{w_1, \ldots, w_n\}$ è una base di V poiché la matrice

$$([w_1]_{\mathcal{B}}, \ldots, [w_n]_{\mathcal{B}}) = P,$$

è inveritibile. Inoltre $\mathcal{M}(\mathcal{B},\mathcal{C}) = P$. Applicando il teorema anteriore otteniamo che

$$\mathcal{M}_{\mathcal{C},\mathcal{C}}(T) = \mathcal{M}(\mathcal{C},\mathcal{B})\mathcal{M}_{\mathcal{B},\mathcal{B}}(T)\mathcal{M}(\mathcal{B},\mathcal{C}) = P^{-1}AP = B.$$

Capitolo 8

Struttura Metrica

8.1 Prodotto scalare canonica di $\mathbb{R}^n$

Dati $X, Y \in \mathbb{R}^n$ con $X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ e $Y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$, definiamo lo scalare:

$$\langle X, Y \rangle = x_1 y_1 + \cdots + x_n y_n = X^T Y.$$

$\langle X, Y \rangle$ è detto *prodotto scalare standard* oppure *prodotto scalare canonico*. Il prodotto scalare canonico è una applicazione

$$\mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R} \qquad (X, Y) \mapsto \langle X, Y \rangle.$$

che gode delle seguenti proprietà.

Proposizione 8.2. *Se $X, Y, Z \in \mathbb{R}^n$ e $\lambda \in \mathbb{R}$, allora*

a) $\langle X, X \rangle \geq 0$ *con uguaglianza se e solamente se* $X = 0_{\mathbb{R}^n}$ *(definito positivo);*

b) $\langle X, Y \rangle = \langle Y, X \rangle$ *(simmetrico);*

c) $\langle X + Y, Z \rangle = \langle X, Z \rangle + \langle Y, Z \rangle$*;*

d) $\langle \lambda X, Z \rangle = \lambda \langle X, Z \rangle$*;*

e) $\langle Z, X + Y \rangle = \langle Z, X \rangle + \langle Z, Y \rangle$*;*

f) $\langle Z, \lambda X \rangle = \lambda \langle Z, X \rangle$.

Dimostrazione. Verifichiamo solamente la prima proprietà. Le altre sono conseguenza delle proprietà del prodotto di matrici.

$\langle X, X\rangle = x_1^2 + \cdots + x_n^2 \geq 0$. Inoltre $\langle X, X\rangle = 0$ se e solamente se $X = 0_{\mathbb{R}^n}$. □

Sia X un vettore. Chiameremo *norma* o *lunghezza* di X il numero reale $\| X \| = \sqrt{\langle X, X\rangle}$. Dunque se $X \neq 0_{\mathbb{R}^n}$, allora $\| X \| > 0$. La *distanza* fra due vettori X, Y è definita come il numero non negativo $\| X - Y \|$. Quindi la lunghezza di un vettore è la distanza dall'origine 0.

Siano $v, w \in \mathbb{R}^n$, $w \neq 0$. Diremo *proiezione ortogonale* di v su w è il vettore

$$\mathrm{pr}_w(v) = \frac{\langle v, w\rangle}{\| w \|^2} w.$$

Proposizione 8.3. *(Disuguaglianza di Cauchy-Schwartz) Siano $X, Y \in \mathbb{R}^n$. Allora*

$$|\langle X, Y\rangle \leq \| X \| \| Y \|,$$

e l'uguaglianza vale se e solamente se X e Y sono linearmente dipendenti.

Dimostrazione. Se $Y = 0_{\mathbb{R}^n}$, la disuguaglianza è vera. Supponiamo che $Y \neq 0_{\mathbb{R}^n}$. Allora per ogni $t \in \mathbb{R}$ si ha

$$\begin{aligned}
0 \leq \langle X + tY, X + tY\rangle &= \underbrace{\langle X, X + tY\rangle} + \underbrace{\langle tY, X + tY\rangle} \\
&= \underbrace{\langle X, X\rangle + \langle X, tY\rangle} + \underbrace{\langle tY, X\rangle + \langle tY, tY\rangle} \\
&= \langle X, X\rangle + t\langle X, Y\rangle + t\langle Y, X\rangle + t^2\langle Y, Y\rangle \\
&= \langle X, X\rangle + 2\langle X, Y\rangle t + \langle Y, Y\rangle t^2.
\end{aligned}$$

L'equazione anteriore definisce una parabola il cui disciminante

$$0 \geq \Delta = b^2 - 4ac = 4\langle X, Y\rangle^2 - 4\langle X, X\rangle\langle Y, Y\rangle,$$

è non positivo. Quindi

$$|\langle X, Y\rangle| \leq \| X \| \| Y \| .$$

Se $|\langle X, Y\rangle| = \| X \| \| Y \|$, allora $\Delta = 0$. Quindi esiste $t_o \in \mathbb{R}$ tale che $0 = \langle X + t_o Y, X + t_o Y\rangle$. Per le proprietà del prodotto scalare si ha $X + t_o Y = 0_{\mathbb{R}^n}$, ovvero X e Y sono linearmente dipendenti. □

Definizione 8.4. *Siano X, Y due vettori non nulli. Definiamo l'angolo fra X e Y come l'unico valore $\theta \in [0, \pi]$ tale che*

$$\cos\theta = \frac{\langle X, Y\rangle}{\| X \| \| Y \|}.$$

Proposizione 8.5. *Siano $X, Y \in \mathbb{R}^n$ vettori non nulli. L'angolo fra X e Y è acuto, rispettivamente ottuso, se e solamente se $\langle X, Y\rangle > 0$. rispettivamente $\langle X, Y\rangle < 0$.*

Definizione 8.6. *Diremo che due vettori X, Y sono ortogonali se il loro prodotto scalare è nullo.*

Teorema 8.7. *Siano $X, Y \in \mathbb{R}^n$. Allora $\| X \|^2 + \| Y \|^2 = \| X + Y \|^2$ se e solamente se X, Y sono ortogonali.*

Dimostrazione. Siano $X, Y \in \mathbb{R}^n$. Utilizzando le proprietà del prodotto scalare si ha

$$\| X + Y \|^2 = \langle X + Y, X + Y\rangle = \langle X, X\rangle + 2\langle X, Y\rangle + \langle Y, Y\rangle,$$

dalla quale segue la tesi. □

Proposizione 8.8. *Siano $v_1, \dots, v_k \in \mathbb{R}^n$ non nulli e ortogonali a due a due. Allora $v_1, \dots, v_k$ sono linearmente indipendenti.*

Dimostrazione. Sia $\alpha_1 v_1 + \cdots + \alpha_k v_k = 0_{\mathbb{R}^n}$ una combinazione lineare dei vettori $v_1, \dots, v_k$ a coefficienti $\alpha_1, \dots, \alpha_k$ uguale al vettore nullo. Fissiamo $1 \leq j \leq k$. Allora

$$\begin{aligned} 0 &= \langle 0_{\mathbb{R}^n}, v_j\rangle \\ &= \langle \alpha_1 v_1 + \cdots + \alpha_k v_k, v_j\rangle \\ &= \alpha_1\langle v_1, v_j\rangle + \cdots + \alpha_k\langle v_k, v_j\rangle \\ &= \alpha_j\langle v_j, v_j\rangle \text{ (essendo } v_1, \dots, v_k \text{ ortogonali a due a due)} \end{aligned}$$

Poiché $v_j \neq 0_{\mathbb{R}^n}$, $\langle v_j, v_j\rangle > 0$ da cui segue $\alpha_j = 0$. Questo vale per $j = 1, \dots, k$, per cui i vettori $v_1, \dots, v_k$ sono linearmente indipendenti. □

Definizione 8.9. *Una base $\mathcal{B} = \{v_1, \dots, v_n\}$ si dice ortogonale se i vettori $v_1, \dots, v_n$ sono a due a due ortogonali. Diremo che $\mathcal{B} = \{v_1, \dots, v_n\}$ è una base ortonormale se $\mathcal{B} = \{v_1, \dots, v_n\}$ è una base ortogonale e i vettori hanno norma unitaria.*

È facile vericare che la base canonica è una base ortonormale di $\mathbb{R}^n$. Le coordinate di un vettore rispetto alla base canonica sono molto semplici da determinare. Infatti

$$X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1 e_1 + \cdots + x_n e_n.$$

Poiché $x_i = \langle X, e_i \rangle$ per $i = 1, \ldots, n$, si ha

$$X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \langle X, e_1 \rangle e_1 + \cdots + \langle X, e_n \rangle e_n.$$

Il prossimo risultato mostra che la formula precedente vale per ogni base ortonormale.

Proposizione 8.10. *Sia $\mathcal{B} = \{v_1, \ldots, v_n\}$ una base ortogonale di $\mathbb{R}^n$. Sia $v \in \mathbb{R}^n$. Allora*

$$v = \frac{\langle v, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1 + \cdots + \frac{\langle v, v_n \rangle}{\langle v_n, v_n \rangle} v_n.$$

In particolare se $\mathcal{B} = \{v_1, \ldots, v_n\}$ è un base ortonormale, allora:

- $v = \langle v, v_1 \rangle v_1 + \cdots + \langle v, v_n \rangle v_n$, *ovvero* $[v]_{\mathcal{B}} = \begin{bmatrix} \langle v, v_1 \rangle \\ \vdots \\ \langle v, v_n \rangle \end{bmatrix}$;
- $\| v \| = \sqrt{\langle v, v_1 \rangle^2 + \cdots + \langle v, v_n \rangle^2}$.
- $\langle v, w \rangle = [v]_{\mathcal{B}}^T [w]_{\mathcal{B}}$.

Facoltativa. Sia $v = x_1 v_1 + \cdots + x_n v_n$ e sia $1 \leq j \leq n$. Allora

$$\langle v, v_j \rangle = \langle \sum_{m=1}^{n} x_m v_m, v_j \rangle = \sum_{m=1}^{n} x_m \langle v_m, v_j \rangle = x_j \langle v_j, v_j \rangle.$$

Poiché $v_j \neq 0_{\mathbb{R}^n}$ si ha

$$x_j = \frac{\langle v, v_j \rangle}{\langle v_j, v_j \rangle},$$

da cui segue la tesi.

Sia $\mathcal{B} = \{v_1, \ldots, v_n\}$ una base ortonormale. Applicando il risultato anteriore si ha $v = \langle v, v_1 \rangle v_1 + \cdots + \langle v, v_n \rangle v_n$.

Siano $v, w \in V$ e sia $\mathcal{B} = \{v_1, \ldots, v_n\}$ una base ortonormale. Allora

$$\begin{aligned} \langle v, w \rangle &= \langle \sum_{m=1}^{n} \langle v, v_m \rangle v_m, \sum_{l=1}^{n} \langle w, v_l \rangle v_l \rangle \\ &= \sum_{m,l=1}^{n} \langle v, v_m \rangle \langle w, v_l \rangle \langle v_m, v_l \rangle. \end{aligned}$$

Adesso, tenendo in mente che $\langle v_m, v_l\rangle = 0$ se $m \neq l$ e 1 se $l = m$ si ha

$$\langle v, w\rangle = \sum_{m=1}^{n} \langle v, v_m\rangle\langle w, v_m\rangle = [v]_{\mathcal{B}}^T [w]_{\mathcal{B}}.$$

In particolare, se $v = w$, allora $\langle v, v\rangle = \sum_{m=1}^{n}\langle v, v_m\rangle^2$ e quindi $\| \ v \ \|= \sqrt{\langle v, v_1\rangle^2 + \cdots + \langle v, v_n\rangle^2}$. $\square$

Esempio 8.11. *Sia* $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} \end{bmatrix} \right\}$ *una base ortonormale di* $\mathbb{R}^3$. *Allora*

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = x \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \frac{y+z}{\sqrt{2}} \begin{bmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} + \frac{y-z}{\sqrt{2}} \begin{bmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} \end{bmatrix},$$

ovvero

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix}_{\mathcal{B}} = \begin{bmatrix} x \\ \frac{y+z}{\sqrt{2}} \\ \frac{y-z}{\sqrt{2}} \end{bmatrix}.$$

Il prossimo risultato fornisce un algoritmo per costruire basi ortogonali, e ortonormali, di sottospazi vettoriali di $\mathbb{R}^n$.

Proposizione 8.12 (Procedimento di ortogonalizzazione di Gram-Schmidt)**.** *Siano* $v_1, \ldots, v_k \in \mathbb{R}^n$ *vettori linearmente indipendenti. I vettori*

$$\begin{cases} w_1 = v_1 \\ w_2 = v_2 - \frac{\langle v_2, w_1\rangle}{\langle w_1, w_1\rangle} w_1 \\ w_3 = v_3 - \frac{\langle v_3, w_1\rangle}{\langle w_1, w_1\rangle} w_1 - \frac{\langle v_3, w_2\rangle}{\langle w_2, w_2\rangle} w_2 \\ \vdots \\ w_k = v_k - \sum_{j=1}^{k-1} \frac{\langle v_k, w_j\rangle}{\langle w_j, w_j\rangle} w_j \end{cases}$$

sono non nulli, a due a due ortogonali, quindi linearmente indipendenti, e verificano $\mathcal{L}(w_1, \ldots, w_j) = \mathcal{L}(v_1, \ldots, v_j)$ *per* $j = 1, \ldots, k$.

Facoltativa. Dimostreremo il risultato per induzione su k.

Se $k = 1$ non c'è niente da dimostrare. Supponiamo vero per $k-1$ e dimostriamolo per k. I vettori $v_1, \ldots, v_{k-1}$ sono linearemente indipendenti. Per ipotesi induttiva i vettori $w_1, \ldots, w_{k-1}$ sono non nulli, a due a due ortogonali e per ogni $1 \leq j \leq k-1$, si ha $\mathcal{L}(v_1, \dots, v_j) = \mathcal{L}(w_1, \dots, w_j)$. Dobbiamo dimostrare che w_k è non nullo, ortogonale a $w_1, \dots, w_{k-1}$ e $\mathcal{L}(v_1, \dots, v_k) = \mathcal{L}(w_1, \dots, w_k)$. Se $w_k = 0_{\mathbb{R}^n}$ allora

$$v_k = \sum_{i=1}^{k-1} \frac{\langle w_j, v_k \rangle}{\langle w_j, w_j \rangle} w_j.$$

Poiché $\mathcal{L}(w_1, \dots, w_{k-1}) = \mathcal{L}(v_1, \dots, v_{k-1})$, i vettori $v_1, \dots, v_k$ sarebbero linearmente dipendenti. Quindi $w_k \neq 0_{\mathbb{R}^n}$. Adesso, dimostriamo che i vettori $w_1, \dots, w_k$ sono ortogonali. Poiché i vettori $w_1, \dots, w_{k-1}$ sono a due a due ortogonali, è sufficiente dimostrare che

$$\langle w_k, w_s \rangle = 0,$$

per $s = 1, \dots, k-1$. Infatti

$$\begin{aligned}
\langle w_k, w_s \rangle &= \langle v_k - \sum_{j=1}^{k-1} \frac{\langle v_k, w_j \rangle}{\langle w_j, w_j \rangle} w_j, w_s \rangle \\
&= \langle v_k, w_s \rangle - \sum_{j=1}^{k-1} \Big\langle \frac{\langle v_k, w_j \rangle}{\langle w_j, w_j \rangle} w_j, w_s \Big\rangle \\
&= \langle v_k, w_s \rangle - \sum_{j=1}^{k-1} \frac{\langle v_k, w_j \rangle}{\langle w_j, w_j \rangle} \langle w_j, w_s \rangle \\
&= \langle v_k, w_s \rangle - \frac{\langle v_k, w_s \rangle}{\langle w_s, w_s \rangle} \langle w_s, w_s \rangle \\
&= \langle v_k, w_s \rangle - \langle v_k, w_s \rangle \\
&= 0.
\end{aligned}$$

Poiché i vettori $w_1, \dots, w_k$ sono a due a due ortogonali, sono linearmente indipendenti. Dimostriamo che $\mathcal{L}(v_1, \dots, v_k) = \mathcal{L}(w_1, \dots, w_k)$. Per ipotesi induttiva $w_1, \dots, w_{k-1} \in \mathcal{L}(v_1, \dots, v_{k-1}) \subseteq \mathcal{L}(v_1, \dots, v_k)$. Inoltre

$$w_k = v_k - \sum_{j=1}^{k-1} \frac{\langle v_k, w_j \rangle}{\langle w_j, w_j \rangle} w_j \in \mathcal{L}(w_1, \dots, w_{k-1}, v_k) = \mathcal{L}(v_1, \dots, v_k).$$

da cui segue che i vettori $w_1, \dots, w_k \in \mathcal{L}(v_1, \dots, v_k)$. Quindi $\mathcal{L}(w_1, \dots, w_k) \subseteq \mathcal{L}(v_1, \dots, v_k)$. Poiché $\dim \mathcal{L}(w_1, \dots, w_k) = \dim \mathcal{L}(v_1, \dots, v_k) = k$ si ha $\mathcal{L}(w_1, \dots, w_k) = \mathcal{L}(v_1, \dots, v_k)$ $\square$

Esempio 8.13. *Siano* $v_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, v_2 = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}, v_3 = \begin{bmatrix} 1 \\ -1 \\ -3 \end{bmatrix} \in \mathbb{R}^3$*. Allora*

$$w_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix},$$

$$w_2 = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} - \frac{3}{2} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} \\ 1 \\ \frac{1}{2} \end{bmatrix},$$

$$w_3 = \begin{bmatrix} 1 \\ -1 \\ -3 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + 2 \begin{bmatrix} -\frac{1}{2} \\ 1 \\ \frac{1}{2} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$$

Corollario 8.14. *Esistono basi ortonormali differenti dalla base canonica.*

Dimostrazione. Sia $\{v_1, \dots, v_n\}$ una base di $\mathbb{R}^n$. Applicando il procedimento di ortogonalizzazione di Gram-Schimdt ottengo una base $\mathcal{C} = \{w_1, \dots, w_n\}$ ortogonale. Dividendo ciascun vettore per la sua norma, i.e.,

$$\left\{ \frac{w_1}{\| w_1 \|}, \dots, \frac{w_n}{\| w_n \|} \right\}$$

ottengo una base ortonormale. □

Corollario 8.15. *Siano $v_1, \dots, v_k \in \mathbb{R}^n$ vettori non nulli e a due a due ortogonali. Allora è possibile completare $v_1, \dots, v_k$ a base ortogonale di $\mathbb{R}^n$.*

Dimostrazione. Applicando il teorema di completamento a base è possibile completare i vettori $v_1, \dots, v_k$ a una base di $\mathbb{R}^n$, che indicheremo con $v_1, \dots, v_k, v_{k+1}, \dots, v_n$. Applicando il procedimento di Gram-Schimdt alla base $\{v_1, \dots, v_n\}$ troviamo una base ortogonale di $\mathbb{R}^n$ i cui primi k vettori sono $v_1, \dots, v_k$ (perché?). □

Esempio 8.16. *Siano* $\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ -1 \end{bmatrix} \in \mathbb{R}^3$*. Possiamo completarli a base aggiungendo* $\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$*. Quindi*

$$\left\{ \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$$

è una base di $\mathbb{R}^3$. *Applicando il procedimento di Gram-Scmidt, si ottiene la base ortogonale*

$$\left\{ \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 1/15 \\ -1/3 \\ 2/15 \end{bmatrix} \right\},$$

di $\mathbb{R}^3$.

Vediamo infine un legame fra matrici ortogonali e basi ortonormali.

Proposizione 8.17. *Sia* $A = (A^1, \ldots, A^n) \in \mathrm{M}_{n\times n}(\mathbb{R})$. *$A$ è una matrice ortogonale se e solamente se* $A^1, \ldots, A^n$ *formano una base ortonormale.*

Dimostrazione. La matrice A è ortogonale se e solamente se $AA^T = A^T A = \mathrm{Id}_n$. Ricordiamo che se $A^T A = \mathrm{Id}_n$, rispettivamente $AA^T = \mathrm{Id}_n$, allora $AA^T = \mathrm{Id}_n$, rispettivamente $A^T A = \mathrm{Id}_n$.

Sia $A = (a_{ij})_{\substack{1 \le i \le n \\ 1 \le j \le n}}$ e sia $A^T = ({a^T}_{ij})_{\substack{1 \le i \le n \\ 1 \le j \le n}}$. Infine, sia $C = A^T A = (c_{ij})_{\substack{1 \le i \le n \\ 1 \le j \le n}}$. Tenendo in mente che $a^T_{ij} = a_{ji}$, si ha

$$c_{ij} = \sum_{m=1}^{n} a^T_{im} a_{mj} = \sum_{m=1}^{n} a_{mi} a_{mj} = \langle A^i, A^j \rangle.$$

Quindi, la matrice A è ortogonale se e solamente se $C = \mathrm{Id}_n$, ovvero se e solamente se $\langle A^i, A^j \rangle = 0$ se $i \neq j$ e $\langle A^i, A^i \rangle = 1$, quindi se e solamente se i vettori $\left\{A^1, \ldots, A^n\right\}$ formano una base ortonormale. □

8.18 Sottospazio Ortogonale

Sia $W \subseteq \mathbb{R}^n$. Porremo

$$W^\perp := \{X \in \mathbb{R}^n : \langle X, s \rangle = 0 \ \forall s \in W\},$$

ovvero l'insieme dei vettori di $\mathbb{R}^n$ che sono ortogonali a ogni elemento di S.

Proposizione 8.19. *Se* $W \subseteq \mathbb{R}^n$, *allora* $W^\perp$ *è un sottospazio vettoriale di* $\mathbb{R}^n$.

Dimostrazione. Dobbiamo dimostrare che $W^\perp$ è chiuso rispetto alla somma e la moltiplicazione per scalare.

Siano $u, v \in W^\perp$. Vogliamo dimostrare che $u + v \in W^\perp$, ovvero che per ogni $s \in W$, si ha

$$\langle u + v, s \rangle = 0.$$

Poiché $\langle u, s\rangle = \langle v, s\rangle = 0$, si ha

$$\langle u + v, s\rangle = \langle u, s\rangle + \langle v, s\rangle = 0.$$

Analogamente, si dimostra che $W^\perp$ è chiuso rispetto alla moltiplicazione per scalare. Infatti, sia $v \in W^\perp$ e sia $\lambda \in \mathbb{R}$. Se $s \in W$, si ha

$$\langle \lambda v, s\rangle = \lambda\langle v, s\rangle = 0,$$

ovvero $\lambda v \in W^\perp$. □

Proposizione 8.20. *Sia* $W = \mathcal{L}(v_1, \ldots, v_k)$ *un sottospazio vettoriale di* $\mathbb{R}^n$. *Allora*

$$W^\perp = \{v \in \mathbb{R}^n : \langle v, v_1\rangle = \cdots = \langle v, v_k\rangle = 0\}.$$

Dimostrazione. Sia $W = \mathcal{L}(v_1, \ldots, v_k)$ e sia

$$U = \{v \in \mathbb{R}^n : \langle v, v_1\rangle = \cdots = \langle v, v_k\rangle = 0\}.$$

Ricordiamo che $W^\perp = \{X \in \mathbb{R}^n : \langle X, s\rangle = 0, \forall s \in W\}$. Vogliamo dimostrare che $W^\perp = U$.

L'inclusione $W^\perp \subseteq U$ è immediata poiché se $v \in W^\perp$, allora v è ortogonale ai vettori $v_1, \ldots, v_k \in W$ e quindi $v \in U$.

Sia $u \in U$. Vogliamo dimostrare che $u \in W^\perp$, ovvero che

$$\langle u, s\rangle = 0, \ \forall s \in W.$$

Poiché $s \in W = \mathcal{L}(v_1, \ldots, v_k)$, esistono $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$ tali che

$$s = \alpha_1 v_1 + \cdots + \alpha_k v_k.$$

Quindi

$$\begin{aligned}\langle u, s\rangle &= \langle w, \alpha_1 v_1 + \cdots + \alpha_k v_k\rangle \\ &= \alpha_1\langle u, v_1\rangle + \cdots + \alpha_k\langle u, v_k\rangle \\ &= 0 \text{ (essendo } u \in U)\end{aligned}$$

da cui segue che $u \in W^\perp$, ovvero $U \subseteq W^\perp$ concludendo la dimostrazione □

Esempio 8.21. *Sia dato* $W = \mathcal{L}\left(\begin{bmatrix}1\\0\\1\\1\end{bmatrix}, \begin{bmatrix}-1\\2\\1\\3\end{bmatrix}\right)$ *sottospazio di* $\mathbb{R}^4$.

Allora

$$\begin{aligned} W^\perp &= \left\{\begin{bmatrix}x_1\\x_2\\x_3\\x_4\end{bmatrix} \in \mathbb{R}^4 : x_1 + x_3 + x_4 = 0,\ -x_1 + 2x_2 + x_3 + 3x_4 = 0\right\} \\ &= \mathcal{L}\left(\begin{bmatrix}1\\1\\-1\\0\end{bmatrix}, \begin{bmatrix}1\\2\\0\\-1\end{bmatrix}\right) \end{aligned}$$

Sia $W \subseteq \mathbb{R}^n$ un sottospazio vettoriale.

Proposizione 8.22. $\mathbb{R}^n$ *è in somma diretta di* W *e* $W^\perp$.

Dimostrazione. Daremo due dimostrazioni differenti.

Sia $u \in W \cap W^\perp$. Allora $\langle u, u\rangle = 0$ da cui segue $u = 0_{\mathbb{R}^n}$. Quindi $W \cap W^\perp = \{0_{\mathbb{R}^n}\}$.

Sia $\mathcal{B} = \{v_1, \ldots, v_k\}$ una base di W e sia $A = (v_1, \ldots, v_k) \in M_{n\times k}(\mathbb{R})$. Osserviamo che $\dim W = \mathrm{rg}(A) = k$. Applicando la Proposizione anteriore, si ha

$$W^\perp = \{X \in \mathbb{R}^n : \langle X, v_1\rangle = \cdots = \langle X, v_k\rangle = 0\},$$

ovvero $W^\perp = \mathrm{Sol}(A^T | 0_{\mathbb{R}^k})$. Applicando il Corollario 7.3 si ha $\dim W^\perp = n - \mathrm{rg}(A^T) = n - \mathrm{rg}(A) = n - \dim W$. Quindi

$$\dim(W + W^\perp) = \dim W + \dim W^\perp = n.$$

Applicando il Corollario 5.31 si ha che $\mathbb{R}^n$ è in somma diretta di W e $W^\perp$.

Un'altra dimostrazione è la seguente.

Sia $\{v_1, \ldots, v_k\}$ una base di W. Per il Teorema di completamente a base, possiamo completarla a base $\mathcal{B} = \{v_1, \ldots, v_n\}$ di $\mathbb{R}^n$. Applicando il procedimento di Gram-Schimdt alla base $\mathcal{B} = \{v_1, \ldots, v_n\}$, ottengo una base ortogonale di $\mathbb{R}^n$ che indicheremo con $\mathcal{C} = \{w_1, \ldots, w_n\}$. Poiché

$$\mathcal{L}(v_1, \ldots, v_k) = \mathcal{L}(w_1, \ldots, w_k),$$

i vettori $w_1, \ldots, w_k$ formano una base ortogonale di W (Perché?). Inoltre, i vettori $w_{k+1}, \ldots, w_n$ sono non nulli, a due a due ortogonali, e sono ortogonali ad una base di W. Applicando la Proposizione 8.20 si ha $w_{k+1}, \ldots, w_n \in W^\perp$. Affermiamo che i vettori $w_{k+1}, \ldots, w_n$ formano una base ortogonale di $W^\perp$. Poiché sono linearmente indipendenti, è sufficiente dimostrare che formano un sistema di generatori di $W^\perp$.

Sia $z \in W^\perp$. Poiché $\mathcal{C} = \{w_1, \ldots, w_n\}$ è una base ortogonale di $\mathbb{R}^n$, si ha

$$z = \frac{\langle z, w_1 \rangle}{\langle w_1, w_1 \rangle} w_1 + \cdots + \frac{\langle z, w_n \rangle}{\langle w_n, w_n \rangle} w_n.$$

Poiché $z \in W^\perp$ e $W = \mathcal{L}(w_1, \ldots, w_k)$, si ha

$$\frac{\langle z, v_j \rangle}{\langle v_j, v_j \rangle} = 0, \text{ per } j = 1, \ldots, k,$$

ovvero

$$z = \frac{\langle z, w_{k+1} \rangle}{\langle w_{k+1}, w_{k+1} \rangle} w_{k+1} + \cdots + \frac{\langle z, w_n \rangle}{\langle w_n, w_n \rangle} w_n \in \mathcal{L}(v_{k+1}, \ldots, v_n).$$

Quindi $W^\perp = \mathcal{L}(v_{w+1}, \ldots, w_n)$, In particolare, $\dim W^\perp = n - k = n - \dim W$, ovvero $\dim W + \dim W^\perp = n$. Inoltre

$$W + W^\perp = \mathcal{L}(w_1, \ldots, w_n) = \mathbb{R}^n.$$

Riassumendo, abbiamo provato che

$$n = \dim(W + W^\perp) = \dim W + \dim W^\perp.$$

Applicando il Corollario 5.31 si ha $\mathbb{R}^n$ è in somma diretta di W e $W^\perp$. $\square$

Esempio 8.23. *Sia dato* $W = \mathcal{L}\left(\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix} \right)$ *sottospazio di* $\mathbb{R}^3$. *Poiché* $\dim W = 2$, $\dim W^\perp = 1$. *Inoltre,*

$$W^\perp = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \in \mathbb{R}^3 : x + z = 0, \, -x + 2y + z = 0 \right\} = \mathcal{L}\left(\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \right)$$

Esempio 8.24. *Sia dato* $W = \mathcal{L}\left(\begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 1 \\ 0 \end{bmatrix} \right)$ *sottospazio di* $\mathbb{R}^4$. *La dimensione di* W *è due (verificare!) e equazioni cartesiane per il* *sottospazio $W^\perp$ sono:*

$$W^\perp = \begin{cases} x_1 + x_2 + x_4 = 0 \\ x_1 + x_3 - x_4 = 0 \\ 2x_1 + x_2 + x_3 = 0 \end{cases}$$

Applicando il metodo di Gauss ed il metodo della risoluzione all'indietro si ha

$$W^\perp = \mathcal{L}\left(\begin{bmatrix} 1 \\ -1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -1 \\ -1 \end{bmatrix} \right)$$

8.25 Proiezione Ortogonale

Definizione 8.26. *Sia W un sottospazio vettoriale dello spazio vettoriale $\mathbb{R}^n$. Sia $\mathcal{C} = \{w_1, \ldots, w_k\}$ una base ortonormale di W. Si dice proiezione ortogonale di $\mathbb{R}^n$ su W l'applicazione lineare $P_W : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ così definita:*

$$P_W(v) = \langle v, w_1 \rangle w_1 + \cdots + \langle v, w_k \rangle w_k.$$

Esempio 8.27. *Sia $W = \mathcal{L}(w)$, $w \neq 0_{\mathbb{R}^n}$. Una base ortonormale di W è il vettore $\dfrac{w}{\| w \|}$. La proiezione ortogonale di $\mathbb{R}^n$ su W è l'applicazione*

$$P_W : \mathbb{R}^n \longrightarrow \mathbb{R}^n, \qquad v \mapsto \langle v, \frac{w}{\| w \|} \rangle \frac{w}{\| w \|} = \frac{\langle v, w \rangle}{\langle w, w \rangle} w = \mathrm{pr}_w(v).$$

Proposizione 8.28. *Sia $P_W : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ la proiezione ortogonale di $\mathbb{R}^n$ su W. Allora $\operatorname{Ker} P_W = W^\perp$ e $\operatorname{Im} P_W = W$.*

Dimostrazione. Un vettore $v \in \operatorname{Ker} P_W$ se e solamente se $P_W(v) = 0_{\mathbb{R}^n}$, ovvero se e solamente se

$$P_W(v) = \langle v, w_1 \rangle w_1 + \cdots + \langle v, w_k \rangle w_k = 0_{\mathbb{R}^n}.$$

Poiché i vettori $w_1, \ldots, w_k \in \mathbb{R}^n$ sono linearmente indipendenti, si ha che $v \in \operatorname{Ker} P_W$ se e solamente se $\langle v, w_1 \rangle = \cdots = \langle v, w_k \rangle = 0$. Applicando la Proposizione 8.20, si ha che $v \in \operatorname{Ker} P_W$ se e solamente se $v \in W^\perp$.

L'immagine di P_W è contenuta in W (perché?). Poiché $\dim \operatorname{Im} P_W = n - \dim W^\perp = \dim W$ si ha $\operatorname{Im} P_W = W$. □

Il prossimo risultato garantisce che $P_W(v)$ non dipende dalla base ortonormale scelta di W.

Proposizione 8.29. *Sia $v \in \mathbb{R}^n$. Allora esiste un unico $w \in W$ tale che $v - w \in W^\perp$. Inoltre, se $\mathcal{C} = \{w_1, \dots, w_k\}$ è una base ortonormale di W, allora $u = P_W(v) = \langle v, w_1 \rangle w_1 + \cdots + \langle v, w_k \rangle w_k$.*

Facoltativa. Sia $\mathcal{C} = \{w_1, \dots, w_k\}$ una base ortonormale di W e sia $w \in W$. Possiamo scrivere $w = \alpha_1 w_1 + \cdots + \alpha_k w_k$. Il vettore $v - w \in W^\perp$ se e solamente se $\langle v - w, w_j \rangle = 0$ per ogni $j = 1, \dots, k$, se e solamente se $\alpha_j = \langle v, w_j \rangle$. Quindi $w = \langle v, w_1 \rangle w_1 + \cdots + \langle v, w_k \rangle$ da cui segue l'esistenza e l'unicità. $\square$

Esempio 8.30. *Sia dato $W = \mathcal{L}\left(\begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 2 \\ 1 \end{bmatrix} \right)$ un sottospazio di $\mathbb{R}^4$. Vogliamo calcolare la proiezione ortogonale su W. Per definizione di proiezione ortogonale, bisogna calcolare una base ortonornale di W. Poiché la matrice*

$$\begin{bmatrix} 1 & 2 & 2 \\ 1 & 0 & 1 \\ 1 & 1 & 2 \\ 0 & 1 & 1 \end{bmatrix},$$

ha rango 2 (verificare!) la dimensione di W è 2 ed una base di W è formata da

$$\begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 1 \\ 1 \end{bmatrix}.$$

Applicando il procedimento di Gram-schimdt e poi dividendo ciascun vettore per la sua norma ottengo una base ortonormale così fatta:

$$\begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \\ 0 \end{bmatrix}, \begin{bmatrix} 1/\sqrt{3} \\ -1/\sqrt{3} \\ 0 \\ 1/\sqrt{3} \end{bmatrix}.$$

Quindi

$$P_W\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}\right) = \langle \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}, \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \\ 0 \end{bmatrix} \rangle \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \\ 0 \end{bmatrix}$$

$$+ \langle \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}, \begin{bmatrix} 1/\sqrt{3} \\ -1/\sqrt{3} \\ 0 \\ 1/\sqrt{3} \end{bmatrix} \rangle \begin{bmatrix} 1/\sqrt{3} \\ -1/\sqrt{3} \\ 0 \\ 1/\sqrt{3} \end{bmatrix}$$

$$= \frac{(x_1 + x_2 + x_3)}{\sqrt{3}} \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 0 \\ 1/\sqrt{3} \end{bmatrix} + \frac{(x_1 - x_2 + x_4)}{\sqrt{3}} \begin{bmatrix} 1/\sqrt{3} \\ -1/\sqrt{3} \\ 0 \\ 1/\sqrt{3} \end{bmatrix}$$

$$= \begin{bmatrix} 2x_1 + x_3 + x_4/3 \\ 2x_2 + x_3 - x_4/3 \\ x_1 + x_2 + x_3/3 \\ x_1 - x_2 + x_4/3 \end{bmatrix}$$

8.31 Prodotto Hermitiano canonico di $\mathbb{C}^n$

Dati $Z = \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix}$ e $W = \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix}$ in $\mathbb{C}^n$ definiamo il prodotto *Hermitiano canonico*:

$$\langle Z, W \rangle = z_1\overline{w_1} + \cdots + z_n\overline{w_n} = Z^T\overline{W}.$$

Il prodotto Hermitiano canonica è una applicazione

$$\mathbb{C}^n \times \mathbb{C}^n \longrightarrow \mathbb{C} \qquad (Z, W) \mapsto \langle Z, W \rangle$$

che soddisfa alle seguenti proprietà:

Proposizione 8.32. *Siano $Z, W, U \in \mathbb{C}^n$ e $\lambda \in \mathbb{C}$. Allora*

a) $\langle Z, W \rangle = \overline{\langle W, Z \rangle}$.

b) $\langle Z, Z \rangle \geq 0$ *e l'uguaglianza vale se e solamente se* $Z = 0_{\mathbb{C}^n}$*;*

c) $\langle Z + W, U \rangle = \langle Z, U \rangle + \langle W, U \rangle$*;*

d) $\langle U, Z + W \rangle = \langle U, Z \rangle + \langle U, W \rangle$*;*

e) $\langle \lambda Z, W\rangle = \lambda\langle Z, W\rangle$*;*

f) $\langle Z, \lambda W\rangle = \overline{\lambda}\langle Z, W\rangle$.

Dimostrazione. Esercizio. □

Sia X un vettore. Chiameremo *norma* o *lunghezza* di X il numero reale $\| X \| = \sqrt{\langle X, X\rangle}$. Dunque se $X \neq 0_{\mathbb{C}^n}$, allora $\| X \| > 0$. Chiamaremo *distanza* fra due vettori X, Y il numero non negativo $\| X - Y \|$. Quindi la lunghezza di un vettore è la distanza dall'origine 0. Inoltre vale la seguente disuguaglianza.

Teorema 8.33 (Disuguaglianza di Cauchy-Schwartz)**.**

$$|\langle Z, W\rangle| \leq \| Z \| \| W \|,$$

dove $\| Z \| := \sqrt{\langle Z, Z\rangle}$*, e l'uguaglianza vale se e solamente se* Z *e* W *sono linearmente dipendenti.*

Facoltativa. Siano $\alpha, \beta \in \mathbb{C}$. Se $Z = 0_{\mathbb{C}^n}$, allora la disuguaglianza è banalmente verificata. Supponiamo che $Z \neq 0_{\mathbb{C}^n}$. Allora

$$\begin{aligned} 0 &\leq \langle \alpha Z + \beta W, \alpha Z + \beta W\rangle \\ &\leq |\alpha|^2\langle Z, Z\rangle + |\beta|^2\langle W, W\rangle + \alpha\overline{\beta}\langle Z, W\rangle + \overline{\alpha}\beta\langle W, Z\rangle \\ . \end{aligned}$$

Se poniamo $\alpha = -\langle W, Z\rangle$ e $\beta = \langle Z, Z\rangle$, otteniamo

$$|\langle Z, W\rangle|^2\langle Z, Z\rangle + |\langle Z, Z\rangle|^2\langle W, W\rangle - 2\langle Z, Z\rangle|\langle Z, W\rangle|^2 \geq 0,$$

ovvero, dividendo per $\langle Z, Z\rangle$, si ha

$$|\langle Z, W\rangle|^2 \leq \langle Z, Z\rangle\langle W, W\rangle.$$

Se $|\langle Z, W\rangle|^2 = \langle Z, Z\rangle\langle W, W\rangle$, allora se poniamo $\alpha = -\langle W, Z\rangle$ e $\beta = \langle Z, Z\rangle$, si ha $\langle \alpha Z + \beta W, \alpha Z + \beta W\rangle = 0$ da cui segue $\alpha Z + \beta W = 0_{\mathbb{C}^n}$. □

Definizione 8.34. *Siano* $X, Y \in \mathbb{C}^n$*. Diremo che* X *e* Y *sono vettori ortogonali se* $\langle X, Y\rangle = 0$*.*

Come per il prodotto scalare standard, possiamo dimostrare il seguente risultato.

Proposizione 8.35. *Siano* $v_1, \ldots, v_k \in \mathbb{C}^n$ *non nulli e ortogonali a due a due. Allora* $v_1, \ldots, v_k$ *sono linearmente indipendenti.*

Una base $\mathcal{B} = \{v_1, \dots, v_n\}$ si dice ortogonale, rispettivamente ortonormale, se i vettori $v_1, \dots, v_n$ sono a due a due ortogonali, rispettivamente se i vettori sono a due a due ortogonali ed hanno norma unitaria. L'esistenza di basi ortogonali è conseguenza del seguente risultato.

Proposizione 8.36 (Procedimento di ortogonalizzazione di Gram-Schmidt)**.** *Siano $v_1, \dots, v_k \in V$ vettori linearmente indipendenti. I vettori*

$$\begin{cases} w_1 = v_1 \\ w_2 = v_2 - \frac{\langle v_2, w_1\rangle}{\langle w_1, w_1\rangle} w_1 \\ w_3 = v_3 - \frac{\langle v_3, w_1\rangle}{\langle w_1, w_1\rangle} w_1 - \frac{\langle v_3, w_2\rangle}{\langle w_2, w_2\rangle} w_2 \\ \vdots \\ w_k = v_k - \sum_{j=1}^{k-1} \frac{\langle v_k, w_j\rangle}{\langle w_j, w_j\rangle} w_j \end{cases}$$

sono non nulli, a due a due ortogonali, quindi linearmente indipendenti, e verificano $\mathcal{L}(w_1, \dots, w_j) = \mathcal{L}(v_1, \dots, v_j)$ per $j = 1, \dots, k$.

Proposizione 8.37. *Sia $A = (A^1, \dots, A^n) \in \mathrm{M}_{n\times n}(\mathbb{C})$. La matrice A è unitaria se e solamente se $\{A^1, \dots, A^n\}$ è una base ortonormale.*

Sia $S \subset \mathbb{C}^n$. Definiamo

$$S^\perp := \{X \in \mathbb{C}^n : \langle X, s\rangle = 0 \ \forall s \in S\}.$$

Proposizione 8.38. *$S^\perp$ è un sottospazio vettoriale di $\mathbb{C}^n$. Se $S = \mathcal{L}(v_1, \dots, v_k)$, allora*

$$S^\perp = \{v \in \mathbb{C}^n : \langle v, v_1\rangle = \dots = \langle v, v_k\rangle = 0\}$$

Capitolo 9

Endomorfismi Diagonalizzabili e Teorema spettrale

9.1 Autovalori e autovettori

Sia V uno spazio vettoriale su $\mathbb{K}$.

Definizione 9.2. *Un endomorfismo, o operatore, di V è una applicazione lineare $T : V \longrightarrow V$, ovvero una applicazione lineare di V in se stesso.*

Definizione 9.3. *Sia $T : V \longrightarrow V$ un endomorfismo. Diremo che un vettore $v \in V$ non nullo è un autovettore di T se $T(v) = \lambda v$ per un certo $\lambda \in \mathbb{K}$. Un autovalore di T è uno scalare $\lambda \in \mathbb{K}$ tale che esiste un vettore non nullo $v \in V$ tale che $T(v) = \lambda v$.*

Se $v \in V$ è un vettore non nullo tale che $T(v) = \lambda v$, allora diremo che v è un autovettore di T relativo, o corrispondente, all'autovalore λ. L'insieme degli autovalori di T si chiama *lo spettro di T*. Dato $\lambda \in \mathbb{K}$ definiamo

$$V_\lambda = \{v \in V : T(v) = \lambda v\} = \{v \in V : (T - \lambda \mathrm{Id}_V)(v) = 0_V\}.$$

Poiché $V_\lambda = \mathrm{Ker}\,(T - \lambda \mathrm{Id}_V)$, V_λ è un sottospazio vettoriale di V. Osserviamo che $V_0 = \mathrm{Ker}\, T$. Dalla definizione di autovalore segue che $V_\lambda \neq \{0_V\}$ se e solamente se λ è un autovalore. In tal caso

$$V_\lambda = \big\{v \in V : v \text{ è un autovettore di } T \text{ relativo a } \lambda\big\} \cup \{0_V\}.$$

che chiameremo *autospazio relativo a λ*.

Proposizione 9.4. *Sia $T : V \longrightarrow V$ un endomorfismo. Uno scalare $\lambda \in \mathbb{K}$ è un autovalore di T se e solamente se $T - \lambda Id_V$ non è iniettivo. In particolare T è iniettiva, e quindi biunivoca, se e solamente se 0 non è un autovalore di T*

Dimostrazione. $\lambda \in \mathbb{K}$ è una autovalore se e solamente se $V_\lambda \neq \{0_V\}$. Poiché $V_\lambda = \mathrm{Ker}\,(T - \lambda \mathrm{Id}_V)$, e $T - \lambda \mathrm{Id}_V$ è un operatore, applicando il Corollario 6.16 si ha che $\lambda \in \mathbb{K}$ è un autovalore se e solamente se $T - \lambda \mathrm{Id}_V$ non è iniettivo e quindi non è biunivoca. □

Definizione 9.5. *Sia $T : V \longrightarrow V$ un endomorfismo. Diremo che T è diagonalizzabile se esiste una base di V formata da autovettori di T.*

Sia $T : V \longrightarrow V$ un endomorfismo di V diagonalizzabile. Esiste una base $\mathcal{B} = \{v_1, \ldots, v_n\}$ di V formata da autovettori di T, ovvero $T(v_1) = \lambda_1 v_1, \ldots, T(v_n) = \lambda_n v_n$. Ricordiamo che se $\{e_1, \ldots, e_n\}$ è la base canonica di $\mathbb{K}^n$, allora $[v_i]_\mathcal{B} = e_i$, per $i = 1, \ldots, n$. Calcoliamo le coordinate di $T(v_1)$ rispetto alla base $\mathcal{B} = \{v_1, \ldots, v_n\}$.

$$[T(v_1)]_\mathcal{B} = [\lambda_1 v_1]_\mathcal{B} = \lambda_1 [v_1]_\mathcal{B} = \begin{bmatrix} \lambda_1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \lambda_1 e_1.$$

Analogamente,

$$[T(v_i)]_\mathcal{B} = \lambda_i [v_i]_\mathcal{B} = \lambda_i e_i,$$

per $i = 1, \ldots, n$. Quindi la matrice associata a T rispetto alla base $\mathcal{B}$, in partenza e in arrivo, è una matrice diagonale

$$\mathcal{M}_{\mathcal{B},\mathcal{B}}(T) = \begin{pmatrix} \lambda_1 & & & \\ & \ddots & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix},$$

dove gli elementi sulla diagonale principale sono gli autovalori di T. Viceversa, supponiamo che la matrice associata a T rispetto a una base $\mathcal{B} = \{v_1, \ldots, v_n\}$ in partenza ed in arrivo sia una matrice

$$\mathcal{M}_{\mathcal{B},\mathcal{B}}(T) = \begin{pmatrix} \lambda_1 & & & \\ & \ddots & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix},$$

diagonale. Allora

$$\begin{aligned}
[T(v_i)]_{\mathcal{B}} &= \mathcal{M}_{\mathcal{B},\mathcal{B}}(T)[v_i]_{\mathcal{B}} \\
&= \mathcal{M}_{\mathcal{B},\mathcal{B}}(T)e_i \\
&= \lambda_i e_i \\
&= \lambda_i [v_i]_{\mathcal{B}} \\
&= [\lambda_i v_i]_{\mathcal{B}}.
\end{aligned}$$

Tendendo in mente che $v = w$ se e solamente se $[v]_{\mathcal{B}} = [w]_{\mathcal{B}}$, si ha che $T(v_i) = \lambda_i v_i$ per $i = 1, \ldots, n$ da cui segue che v_i è autovettore relativo all'autovalore λ_i. Quindi $\{v_1, \ldots, v_n\}$ è una base di V formata da autovettori di T.

Proposizione 9.6. *Sia $T : V \longrightarrow V$ un endomorfismo. Allora T è diagonalizzabile se e solamente se esiste una base $\mathcal{B}$ di V tale che $\mathcal{M}_{\mathcal{B},\mathcal{B}}(T)$ è una matrice diagonale.*

Sia $\mathcal{C}$ una base di V. Poiché $\mathcal{M}_{\mathcal{C},\mathcal{C}}(T)$ è simile alla matrice $\mathcal{M}_{\mathcal{B},\mathcal{B}}(T)$ si ha il seguente risultato.

Proposizione 9.7. *Sia $T : V \longrightarrow V$ un endomorfismo e sia $\mathcal{C}$ una base di V. T è diagonalizzabile se e solamente se $\mathcal{M}_{\mathcal{C},\mathcal{C}}(T)$ è simile a una matrice diagonale.*

Il prossimo risultato prova che autovettori corrispondenti ad autovalori distinti sono linearmente indipendenti.

Proposizione 9.8. *Sia $T : V \longrightarrow V$ una applicazione lineare e siano $v_1, \ldots, v_m$ autovettori di T corrispondenti agli autovalori $\lambda_1, \ldots, \lambda_m$. Se $\lambda_1, \ldots, \lambda_m$ sono distinti, i.e., $\lambda_i \neq \lambda_j$ se $i \neq j$, allora gli autovettori $v_1, \ldots, v_m$ di T sono linearmente indipendenti.*

Facoltativa. la dimostrazione sarà fatta per induzione sul numero di autovettori.

Se $m = 1$ la proposizione è banalmente verificata poiché un autovettore è un vettore non nullo e quindi linearmente indipendente.

Supponiamo che la proposizione sia vera per m autovettori di T corrispondenti a m autovalori distinti e dimostriamolo per $m + 1$ autovalori distinti.

Siano $v_1, \ldots, v_{m+1}$ autovettori corrispondenti ad autovalori $\lambda_1, \ldots, \lambda_{m+1}$ distinti. Per ipotesi induttiva i vettori $v_1, \ldots, v_m$ sono linearmente indipendenti. Supponiamo per assurdo che $v_1, \ldots, v_{m+1}$ fossero linearmente dipendenti. Poiché $v_1, \ldots, v_m$ sono linearmente indipendenti, il vettore v_{m+1}

sarebbe combinazione lineare di $v_1, \dots, v_m$ (perché?). Quindi esistono degli scalari $\alpha_1, \dots, \alpha_m \in \mathbb{K}$ tale che

$$v_{m+1} = \alpha_1 v_1 + \cdots + \alpha_m v_m.$$

Applicando T si ha

$$\begin{aligned} T(v_{m+1}) = \lambda_{m+1} v_{m+1} &= T(\alpha_1 v_1 + \cdots + \alpha_m v_m) \\ &= \alpha_1 T(v_1) + \cdots + \alpha_m T(v_m) \\ &= \lambda_1 \alpha_1 v_1 + \cdots + \lambda_m \alpha_m v_m. \end{aligned}$$

Dall'altro lato

$$\lambda_{m+1} v_{m+1} = \lambda_{m+1} \alpha_1 v_1 + \cdots + \lambda_{m+1} \alpha_m v_m.$$

Quindi

$$\lambda_{m+1} \alpha_1 v_1 + \cdots + \lambda_{m+1} \alpha_m v_m = \lambda_1 \alpha_1 v_1 + \cdots + \lambda_m \alpha_m v_m,$$

ovvero

$$(\lambda_{m+1} - \lambda_1) \alpha_1 v_1 + \cdots + (\lambda_{m+1} - \lambda_m) \alpha_m v_m = 0.$$

Poiché i vettori $v_1, \dots, v_m$ sono linearmente indipendenti, ne segue che i coefficienti sono tutti nulli:

$$\alpha_1 (\lambda_{m+1} - \lambda_1) = \cdots = \alpha_m (\lambda_{m+1} - \lambda_m) = 0$$

Per ipotesi gli autovalori sono distinti. Quindi $\alpha_1 = \cdots = \alpha_m = 0$, ovvero $v_{m+1} = 0$. Assurdo poiché v_{m+1} è un vettore non nullo essendo un autovettore. Quindi i vettori $v_1, \dots, v_{m+1}$ sono linearmente indipendenti. □

Corollario 9.9. *Sia V uno spazio vettoriale su $\mathbb{K}$ di dimensione n e sia $T : V \longrightarrow V$ un endomorfismo. Se T ammette n autovalori distinti, allora T è diagonalizzabile.*

Dimostrazione. Poiché T ha $n = \dim V$ autovalori distinti esitono n autovettori $v_1, \dots, v_n$ corrispondenti ad autovalori distinti. Per il risultato anteriore $v_1, \dots, v_n$ sono linearmente indipendenti e quindi formano una base di V concludendo la dimostrazione. □

9.10 Polinomio caratteristico

Sia V uno spazio vettoriale su $\mathbb{K}$ di dimensione n. Vogliamo calcolare gli autovalori di un endomorfismo $T : V \longrightarrow V$.

Sia $\mathcal{B}$ una base di V. Uno scalare $\lambda \in \mathbb{K}$ è un autovalore di T se e solamente se $T - \lambda Id_V$ non è iniettiva ovvero se e solamente se la matrice $\mathcal{M}_{\mathcal{B},\mathcal{B}}(T - \lambda \mathrm{Id}_V)$ non è invertibile . Poiché $\mathcal{M}_{\mathcal{B},\mathcal{B}}(T - \lambda \mathrm{Id}_V) = \mathcal{M}_{\mathcal{B},\mathcal{B}}(T) - \lambda \mathrm{Id}_n$, si ha che $\lambda \in \mathbb{K}$ è un autovalore di T se e solamente se

$$\det(\mathcal{M}_{\mathcal{B},\mathcal{B}}(T) - \lambda \mathrm{Id}_n) = 0.$$

Definiamo

$$p_{\mathcal{B}} : \mathbb{K} \longrightarrow \mathbb{K}, \qquad p_{\mathcal{B}}(t) := \det(\mathcal{M}_{\mathcal{B},\mathcal{B}}(T) - t\mathrm{Id}_n)$$

Proposizione 9.11. *$p_{\mathcal{B}}(t)$ è un polinomio di grado $n = \dim V$ ed è ha la seguente espressione*

$$(-1)^n t^n + (-1)^{n-1}(\mathrm{Tr}(\mathcal{M}_{\mathcal{B},\mathcal{B}}(T))t^{n-1} + \cdots + det(\mathcal{M}_{\mathcal{B},\mathcal{B}}(T)).$$

Inoltre, $p_{\mathcal{B}}(t)$ non dipende dalla base scelta.

Dimostrazione. Noi proveremo solamente che se $\mathcal{C}$ è una base di V, allora $p_{\mathcal{C}} = p_{\mathcal{B}}$.

Sia $\mathcal{C} = \{w_1, \ldots, w_n\}$ una base di V. Allora

$$\mathcal{M}_{\mathcal{C},\mathcal{C}}(T) = \mathcal{M}(\mathcal{B},\mathcal{C})^{-1}\mathcal{M}_{\mathcal{B},\mathcal{B}}(T)\mathcal{M}(\mathcal{B},\mathcal{C}),$$

da cui segue

$$\begin{aligned}
p_{\mathcal{C}}(t) &= \det(\mathcal{M}_{\mathcal{C},\mathcal{C}}(T) - t\mathrm{Id}_n) \\
&= det\left(\mathcal{M}(\mathcal{B},\mathcal{C})^{-1}\mathcal{M}_{\mathcal{B},\mathcal{B}}(T)\mathcal{M}(\mathcal{B},\mathcal{C}) - t\mathrm{Id}_n\right) \\
&= \det\left(\mathcal{M}(\mathcal{B},\mathcal{C})^{-1}\mathcal{M}_{\mathcal{B},\mathcal{B}}(T)\mathcal{M}(\mathcal{B},\mathcal{C}) - t\mathcal{M}(\mathcal{B},\mathcal{C})^{-1}\mathcal{M}(\mathcal{B},\mathcal{C})\right) \\
&= \det\left((\mathcal{M}(\mathcal{B},\mathcal{C})^{-1}(\mathcal{M}_{\mathcal{B},\mathcal{B}}(T) - t\mathrm{Id}_n)\mathcal{M}(\mathcal{B},\mathcal{C})\right) \\
&= \det((\mathcal{M}(\mathcal{B},\mathcal{C})^{-1})det(\mathcal{M}_{\mathcal{B},\mathcal{B}}(T) - t\mathrm{Id}_n)det(\mathcal{M}(\mathcal{B},\mathcal{C})) \\
&= \det(\mathcal{M}_{\mathcal{B},\mathcal{B}}(T) - t\mathrm{Id}_n) \\
&= p_{\mathcal{B}}(t).
\end{aligned}$$

□

Definizione 9.12. *Sia $T : V \longrightarrow V$ un endomorfismo. Sia $\mathcal{B}$ un base di V. Il polinomio $p_T(t) = p_{\mathcal{B}}(t)$ è un polinomio di grado $n = \dim V$ ed è chiamato* il polinomio caratteristico di *T.*

Corollario 9.13. *Sia V uno spazio vettoriale su $\mathbb{K}$ di dimensione n e sia $T : V \longrightarrow V$ un endomorfismo. Sia $\lambda \in \mathbb{K}$. Allora λ è un autovalore di T se e solamente se λ è una radice del polinomio caratteristico di T. In particolare T ha al massimo n autovalori.*

Dimostrazione. Abbiamo dimostrato che $\lambda \in \mathbb{K}$ è un autovalore di T se e solamente se $det(\mathcal{M}_{\mathcal{B},\mathcal{B}}(T) - \lambda \mathrm{Id}_n) = 0$, ovvero se e solamente se $p_T(\lambda) = 0$. Poiché un polinomio di grado n ha al massimo n radice, si ha che T ha al massimo n autovalori. $\square$

Esempio 9.14. *Si consideri l'applicazione lineare $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ definita da*

$$T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} x+y+2z \\ 2x+y+3z \\ x+2y+3z \end{bmatrix},$$

Sia $\mathcal{C} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$ *la base canonica di $\mathbb{R}^3$. Allora*

$$T\left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, T\left(\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}, T\left(\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 2 \\ 3 \\ 3 \end{bmatrix},$$

e quindi

$$M_T = \mathcal{M}_{\mathcal{C},\mathcal{C}}(T) = \begin{bmatrix} 1 & 1 & 2 \\ 2 & 1 & 3 \\ 1 & 2 & 3 \end{bmatrix}.$$

Il polinomio caratteristico di T è dato da

$$p_T(t) = \det(\mathcal{M}_{\mathcal{C},\mathcal{C}}(T) - t\mathrm{Id}_3) = \det \begin{bmatrix} 1-t & 1 & 2 \\ 2 & 1-t & 3 \\ 1 & 2 & 3-t \end{bmatrix}.$$

Sviluppa in determinante rispetto alla 1 colonna si ha

$$\begin{aligned} p_T(t) &= (1-t)(t^2 - 4t - 3) - 2((3-t) - 4) + (3 - (2(1-t)) \\ &= -t^3 + 5t^2 + 3t = -t(t^2 - 5t - 3). \end{aligned}$$

Quindi gli autovalori di T sono: $0, \frac{5+\sqrt{37}}{2}, \frac{5-\sqrt{37}}{2}$.

Esempio 9.15. *Sia* $A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$ *e sia Sia* $T : M_{2\times 2}(\mathbb{R}) \longrightarrow M_{2\times 2}(\mathbb{R})$ *l'applicazione lineare così definita:*

$$T(X) = X - \mathrm{Tr}(X)A.$$

Sia

$$\mathcal{B} = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\},$$

Nell'esempio 7.16 abbiamo visto che

$$\mathcal{M}_{\mathcal{B},\mathcal{B}}(T) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & -1 \\ -1 & 0 & 1 & -1 \\ -1 & 0 & 0 & 0 \end{bmatrix}.$$

Quindi

$$p_T(t) = \det(\mathcal{M}_{\mathcal{B},\mathcal{B}}(T) - t\mathrm{Id}_4) = \det \begin{bmatrix} 1-t & 0 & 0 & 0 \\ -1 & 1-t & 0 & -1 \\ -1 & 0 & 1-t & -1 \\ -1 & 0 & 0 & -t \end{bmatrix} = -t(1-t)^3,$$

ovvero gli autovalori si T *sono* 0 *e* 1*.*

9.16 Molteplicità geometrica e algebrica

Sia V uno spazio vettoriale su $\mathbb{K}$ di dimensione n e sia $T : V \longrightarrow V$ un endomorfismo.

Definizione 9.17. *Sia* $\lambda \in \mathbb{K}$ *un autovalore di* T*. Definiamo:*

- *la* molteplicità algebrica *di* λ *è la molteplicità di* λ *come radice di* $p_T(t)$*, ovvero quante volte* λ *è radice di* $p_T(t)$*, che indicheremo con* $m_a(\lambda)$*.*

- *la* molteplicità geometrica *di* λ *è la dimensione dell'autospazio* V_λ *e si indicherà con* $m_g(\lambda) = \dim V_\lambda$*.*

Se $\lambda \in \mathbb{K}$ è un autovalore, allora $m_a(\lambda)$ è il più grande numero naturale m tale che $(x - \lambda)^m$ divide $p_T(t)$.

Proposizione 9.18. *Sia V uno spazio vettoriale su $\mathbb{K}$ di dimensione n e sia $T : V \longrightarrow V$ un endomorfismo. Sia $\lambda \in \mathbb{K}$ un autovalore. Se $\mathcal{B}$ è una base di V, allora*

$$m_g(\lambda) = n - \mathrm{rg}(\mathcal{M}_{\mathcal{B},\mathcal{B}}(T) - \lambda \mathrm{Id}_n).$$

Dimostrazione. Poiché $V_\lambda = \mathrm{Ker}\,(T - \lambda \mathrm{Id}_V)$, applicando il Teorema della dimensione, tenendo in mente che $\dim V = n$, si ha

$$m_g(\lambda) = \dim \mathrm{Ker}\,(T - \lambda \mathrm{Id}_V) = n - \dim \mathrm{Im}\,(T - \lambda \mathrm{Id}_V).$$

Poiché $\dim \mathrm{Im}\,(T - \lambda \mathrm{Id}_V) = \mathrm{rg}(\mathcal{M}_{\mathcal{B},\mathcal{B}}(T - \lambda \mathrm{Id}_V))$, e tenendo in mente che $\mathcal{M}_{\mathcal{B},\mathcal{B}}(T - \lambda \mathrm{Id}_V) = \mathcal{M}_{\mathcal{B},\mathcal{B}}(T) - \lambda \mathrm{Id}_n$, si ha

$$\begin{aligned} m_g(\lambda) &= n - \dim \mathrm{Im}\,(T - \lambda \mathrm{Id}_V) \\ &= n - \mathrm{rg}(\mathcal{M}_{\mathcal{B},\mathcal{B}}(T - \lambda \mathrm{Id}_V)) \\ &= n - \mathrm{rg}(\mathcal{M}_{\mathcal{B},\mathcal{B}}(T) - \lambda \mathrm{Id}_n). \end{aligned}$$

□

Il prossimo risultato garantisce che la molteplicità algebrica è maggiore oppure uguale della molteplicità geometrica. La dimostrazione è facoltativa.

Proposizione 9.19. *Sia $\lambda \in \mathbb{K}$ un autovalore di $T : V \longrightarrow V$. Allora $m_g(\lambda) \leq m_a(\lambda)$.*

Facoltativa. Sia $\mathcal{B} = \{v_1, \ldots, v_{m_g(\lambda)}\}$ una base di V_λ. Possiamo completarla a base di V che indicheremo con $\mathcal{B} = \{v_1, \ldots, v_n\}$. Allora

$$\mathcal{M}_{\mathcal{B},\mathcal{B}}(T) = \left(\begin{array}{c|c} \lambda Id_{m_g(\lambda) \times m_g(\lambda)} & * \\ \hline 0 & D \end{array} \right).$$

Quindi

$$\begin{aligned} p_T(t) &= \det(\mathcal{M}_{\mathcal{B},\mathcal{B}}(T) - t \mathrm{Id}_n) \\ &= (\lambda - t)^{m_g(\lambda)} \det(D - t \mathrm{Id}_{(n - m_g(\lambda)) \times (n - m_g(\lambda))}), \end{aligned}$$

da cui segue che $m_g(\lambda) \leq m_a(\lambda)$. □

Corollario 9.20. *Sia $T : V \longrightarrow V$ un endomorfismo e sia $\lambda \in \mathbb{K}$ autovalore di $T : V \longrightarrow V$. Se $m_a(\lambda) = 1$, allora $m_g(\lambda) = 1$.*

Dimostrazione. Poiché $m_g(\lambda) \geq 1$ (perché ?), si ha

$$1 \leq m_g(\lambda) \leq m_a(\lambda) = 1,$$

da cui segue la tesi. $\square$

Esempio 9.21. *Si consideri l'applicazione lineare $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ definita da*

$$T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} 2x - y + 3z \\ x + 3z \\ 2x - 2y + 7z \end{bmatrix},$$

Sia $\mathcal{C} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$ *la base canonica di* $\mathbb{R}^3$*. Allora*

$$T\left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}, T\left(\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} -1 \\ 0 \\ -2 \end{bmatrix}, T\left(\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 3 \\ 3 \\ 7 \end{bmatrix},$$

e quindi

$$M_T = \mathcal{M}_{\mathcal{C},\mathcal{C}}(T) = \begin{bmatrix} 2 & -1 & 3 \\ 1 & 0 & 3 \\ 2 & -2 & 7 \end{bmatrix}.$$

Il polinomio caratteristico di T è dato da

$$p_T(t) = \det(\mathcal{M}_{\mathcal{C},\mathcal{C}}(T) - t\mathrm{Id}_3) = \det \begin{bmatrix} 2-t & -1 & 3 \\ 1 & -t & 3 \\ 2 & -2 & 7-t \end{bmatrix} = -(t-1)^2(t-7).$$

Quindi gli autovalori di T sono 1 con molteplicità algebrica 2 e 7 con molteplicità algebrica 1. Poiché

$$\mathcal{M}_{\mathcal{C},\mathcal{C}}(T) - \mathrm{Id}_3 = \begin{bmatrix} 1 & -1 & 3 \\ 1 & -1 & 3 \\ 2 & -2 & 6 \end{bmatrix},$$

la molteplicità geometrica dell'autovalore 1 è $m_g(1) = 3 - \mathrm{rg}(\mathcal{M}_{\mathcal{C},\mathcal{C}}(T) - \mathrm{Id}_3) = 2$*. La molteplicità geometrica dell'autovalore 7 è uno.*

Il prossimo Teorema fornisce un criterio necessario e sufficiente affinché un endomorfismo sia diagonalizzabile. La dimostrazione è facoltativa come è facoltativa la dimostrazione del eguente lemma.

Lemma 9.22. *Sia V uno spazio vettoriale su $\mathbb{K}$ e sia $T : V \longrightarrow V$ un endomorfismo. Siano $\lambda_1, \ldots, \lambda_k \in \mathbb{K}$ autovalori distinti. Indichiamo con $\mathcal{B}_{\lambda_i}$ una base per l'autospazio V_{λ_i} per $i = 1, \ldots, k$. Allora*

$$\mathcal{B} = \mathcal{B}_{\lambda_1} \bigcup \cdots \bigcup \mathcal{B}_{\lambda_k},$$

è formato da vettori linearmente indipendenti.

Dimostrazione. A meno di riordinare i vettori di $\mathcal{B}$, esistono $1 \leq j_1 < j_2 < \cdots < j_k \leq n =$ tale che

$$\begin{array}{ll} v_1, \ldots, v_{j_1} & \text{formano una base di } V_{\lambda_1} \\ v_{j_1+1}, \ldots, v_{j_2} & \text{formano una base di } V_{\lambda_2} \\ \vdots & \\ v_{j_k+1}, \ldots, v_n & \text{formano una base di } V_{\lambda_k} \end{array}$$

Siano $\alpha_1, \ldots, \alpha_{j_1}, \ldots, \alpha_{j_k}, \ldots, \alpha_n \in \mathbb{K}$ e sia $\alpha_1 v_1 + \cdots + \alpha_n v_n = 0_V$ una combinazione lineare uguale al vettore nullo. Poniamo

$$\begin{array}{l} w_1 = \alpha_1 v_1 + \cdots + \alpha_{j_1} v_{j_1} \in V_{\lambda_1} \\ w_2 = \alpha_{j_1+1} v_{j_1+1} + \cdots + \alpha_{j_2} v_{j_2} \in V_{\lambda_2} \\ \vdots \\ w_k = \alpha_{j_k+1} v_{j_k+1} + \cdots + \alpha_n v_n \in V_{\lambda_k} \end{array}$$

Quindi

$$w_1 + \cdots + w_k = 0_V$$

Per la Proposizione 9.8, si ha $w_1 = \cdots = w_k = 0_V$ e quindi

$$\begin{array}{l} 0_V = \alpha_1 v_1 + \cdots + \alpha_{j_1} v_{j_1} \\ 0_V = \alpha_{j_1+1} v_{j_1+1} + \cdots + \alpha_{j_2} v_{j_2} \\ \vdots \\ 0_V = \alpha_{j_k+1} v_{j_k+1} + \cdots + \alpha_n v_n. \end{array}$$

Tenendo in mente che

$$\begin{array}{ll} v_1, \dots, v_{j_1} & \text{formano una base di } V_{\lambda_1} \\ v_{j_1+1}, \dots, v_{j_2} & \text{formano una base di } V_{\lambda_2} \\ \vdots & \\ v_{j_k+1}, \dots, v_n & \text{formano una base di } V_{\lambda_k} \end{array}$$

si ha $\alpha_1 = \cdots = \alpha_n = 0$. $\square$

Teorema 9.23. *Sia V uno spazio vettoriale su $\mathbb{K}$ e sia $T : V \longrightarrow V$ un endomorfismo. Le seguenti condizioni sono equivalenti:*

a) T è diagonalizzabile.

b) Tutti gli autovalori di T sono in $\mathbb{K}$. Inoltre, per ogni $\lambda \in sp(T)$ si ha $m_a(\lambda) = m_g(\lambda)$.

Facoltativa. Supponiamo che T sia diagonalizzabile. Sia $\mathcal{B} = \{v_1, \dots, v_n\}$ una base formata da autovettori di T. Siano $\lambda_1, \dots, \lambda_r$ gli autovalori distinti di T. La matrice associata a T rispetto alla base $\mathcal{B} = \{v_1, \dots, v_n\}$ in partenza ed in arrivo ha la forma

$$\mathcal{M}_{\mathcal{B},\mathcal{B}}(T) = \begin{pmatrix} \lambda_1 \mathrm{Id}_{m_g(\lambda_1)\times m_g(\lambda_1)} & & \\ & \ddots & \\ & & \lambda_k \mathrm{Id}_{m_g(\lambda_k)\times m_g(\lambda_k)} \end{pmatrix}.$$

e quindi

$$\begin{aligned} p_T(t) &= \det(\mathcal{M}_{\mathcal{B},\mathcal{B}}(T) - t\mathrm{Id}_n) \\ &= \det \begin{pmatrix} (\lambda_1 - t)\mathrm{Id}_{m_g(\lambda_1)\times m_g(\lambda_1)} & & \\ & \ddots & \\ & & (\lambda_k - t)\mathrm{Id}_{m_g(\lambda_k)\times m_g(\lambda_k)} \end{pmatrix} \\ &= (\lambda_1 - t)^{m_g(\lambda_1)} \cdots (\lambda_k - t)^{m_g(\lambda_k)}. \end{aligned}$$

da cui segue che tutti gli autovalori stanno in $\mathbb{K}$ ed $m_a(\lambda_j) = m_g(\lambda_j)$ per $j = 1, \dots, k$.

Viceversa, supponiamo che tutti gli autovalori di T stanno in $\mathbb{K}$, ovvero T ha n autovalori non necessariamente distinti, e che la molteplicità algebrica e geometrica coincidono per ogni autovalore. Siano $\lambda_1, \dots, \lambda_k \in \mathbb{K}$ gli autovalori distinti di T. Sia $\mathcal{B}_{V_{\lambda_j}}$ una base di V_{λ_j} per $j = 1, \dots, k$. L'insieme $\mathcal{B} = \mathcal{B}_{V_{\lambda_1}} \bigcup \cdots \bigcup \mathcal{B}_{V_{\lambda_k}}$ è formato da $n = \dim V$ autovettori. Per il Lemma 9.22 i vettori di $\mathcal{B}$ sono linearmente indipendenti e quindi una base, concludendo la dimostrazione. $\square$

Esempio 9.24. *Sia* $T : M_{2\times 2}(\mathbb{R}) \longrightarrow M_{2\times 2}(\mathbb{R})$ *l'applicazione lineare così definita:*

$$T\left(\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\right) = \begin{bmatrix} a_{11} + a_{12} & a_{12} \\ a_{11} + a_{21} + a_{22} & a_{11} + a_{21} + a_{22} \end{bmatrix}.$$

Sia

$$\mathcal{B} = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\},$$

Poiché

$$\mathcal{M}_{\mathcal{B},\mathcal{B}}(T) = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix},$$

si ha

$$P_T(t) = \det(\mathcal{M}_{\mathcal{B},\mathcal{B}}(T) - t\mathrm{Id}_4) = \det \begin{bmatrix} 1-t & 1 & 0 & 0 \\ 0 & 1-t & 0 & 0 \\ 1 & 0 & 1-t & 1 \\ 1 & 1 & 1 & 1-t \end{bmatrix} = t(t-2)(t-1)^2.$$

Quindi gli autovalori si T *sono* $0, 1, 2$. *La molteplicità algebrica di* 0 *e* 2 *è uno mentre la molteplicità algebrica di* 1 *è due. Poiché*

$$m_g(1) = 4 - \mathrm{rg}(\mathcal{M}_{\mathcal{B},\mathcal{B}}(T) - \mathrm{Id}_4) = 4 - \mathrm{rg}\left(\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix}\right) = 1,$$

T *non è diagonalizzabile.*

Esempio 9.25. *Sia* $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ *l'applicazione lineare definita da:*

$$T\left(\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 0 \\ -2 \\ 2 \end{bmatrix}, T\left(\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} -1 \\ -3 \\ 3 \end{bmatrix}, T\left(\begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} -1 \\ -3 \\ -4 \end{bmatrix}.$$

Allora
$$\mathcal{M}_{\mathcal{C},\mathcal{B}}(T) = \begin{bmatrix} 0 & -1 & -1 \\ -2 & -3 & -3 \\ 2 & 3 & -4 \end{bmatrix},$$
è la matrice associata a T rispetto alle basi $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} \right\}$ *in partenza e la base canonica $\mathcal{C}$ in arrivo. Per stabilire se T è diagonalizzabile non posso utilizzare la matrice $\mathcal{M}_{\mathcal{C},\mathcal{B}}(T)$ (perché?). Svolgeremo due procedimenti.*

I metodo. Il polinomio caratteristico di T è
$$p_T(t) = \det(\mathcal{M}_{\mathcal{C},\mathcal{C}}(T) - t\mathrm{Id}_3).$$
La matrice
$$\mathcal{M}_{\mathcal{C},\mathcal{C}}(T) = \mathcal{M}_{\mathcal{C},\mathcal{B}}(T)\mathcal{M}(\mathcal{B},\mathcal{C}).$$
Poiché
$$\begin{bmatrix} x \\ y \\ z \end{bmatrix}_{\mathcal{B}} = \mathrm{Sol}\left(\left[\begin{array}{ccc|c} 1 & 1 & 0 & x \\ 0 & 1 & -1 & y \\ 1 & 1 & 1 & z \end{array}\right] \right) = \begin{bmatrix} 2x - y - z \\ -x + y + z \\ -x + z \end{bmatrix}$$
si ha
$$\mathcal{M}(\mathcal{B},\mathcal{C}) = \begin{bmatrix} 2 & 1 & -1 \\ -1 & 1 & 1 \\ -1 & 0 & 1 \end{bmatrix}$$
e quindi
$$\mathcal{M}_{\mathcal{C},\mathcal{C}}(T) = \begin{bmatrix} 0 & -1 & -1 \\ -2 & -3 & -3 \\ 2 & 3 & -4 \end{bmatrix} \begin{bmatrix} 2 & -1 & -1 \\ -1 & 1 & 1 \\ -1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & -1 & -2 \\ 2 & -1 & -4 \\ -1 & 1 & 3 \end{bmatrix}.$$
Il polinomio caratteristico di T è dato da
$$p_T(t) = \det(\mathcal{M}_{\mathcal{C},\mathcal{C}}(T) - t\mathrm{Id}_3) = \det\left(\begin{bmatrix} 2-t & -1 & -2 \\ 2 & -1-t & -4 \\ -1 & 1 & 3-t \end{bmatrix} \right) = -(t-1)^2(t-2),$$

per cui gli autovalori di T sono 1 con molteplicità algebrica due e 2 con molteplicità algebrica uno, e quindi anche geometrica uno. Poiché

$$\begin{aligned} m_g(1) &= 3 - \mathrm{rg}(\mathcal{M}_{\mathcal{C},\mathcal{C}}(T) - \mathrm{Id}_3) \\ &= 3 - \mathrm{rg}\left(\begin{bmatrix} 1 & -1 & -2 \\ 2 & -2 & -4 \\ -1 & 1 & 2 \end{bmatrix}\right) \\ &= 2, \end{aligned}$$

l'endomorfismo T è diagonalizzabile.
II metodo. Vogliamo calcolare il polinomio caratteristico $p_T(t) = \det(\mathcal{M}_{\mathcal{B},\mathcal{B}}(T) - t\mathrm{Id}_3)$. Poiché

$$T\left(\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 0 \\ -2 \\ 2 \end{bmatrix}, T\left(\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} -1 \\ -3 \\ 3 \end{bmatrix}, T\left(\begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} -1 \\ -3 \\ -4 \end{bmatrix},$$

si ha

$$\mathcal{M}_{\mathcal{B},\mathcal{B}}(T) = \left(\begin{bmatrix} 0 \\ -2 \\ -2 \end{bmatrix}_{\mathcal{B}}, \begin{bmatrix} -1 \\ -3 \\ 3 \end{bmatrix}_{\mathcal{B}}, \begin{bmatrix} -1 \\ -3 \\ 2 \end{bmatrix}_{\mathcal{B}}\right).$$

Le coordinate di un vettore rispetto alla base $\mathcal{B}$ è l'unica soluzione del seguente sistema lineare

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix}_{\mathcal{B}} = \mathrm{Sol}\left(\left[\begin{array}{ccc|c} 1 & 1 & 0 & x \\ 0 & 1 & -1 & y \\ 1 & 1 & 1 & z \end{array}\right]\right) = \begin{bmatrix} 2x - y - z \\ -x + y + z \\ -x + z \end{bmatrix}$$

da cui segue che

$$\mathcal{M}_{\mathcal{B},\mathcal{B}}(T) = \begin{bmatrix} 0 & -2 & -1 \\ 0 & 1 & 0 \\ 2 & 4 & 3 \end{bmatrix}.$$

Il polinomio caratteristico di T è

$$p_T(t) = \det(\mathcal{M}_{\mathcal{B},\mathcal{B}}(T) - t\mathrm{id}_3) = \det\left(\begin{bmatrix} -t & -2 & -1 \\ 0 & 1-t & 0 \\ 2 & 4 & 3-t \end{bmatrix}\right) = (2-t)(1-t)^2,$$

per cui gli autovalori di T sono 2 con molteplicità algebrica uno, e quindi anche la molteplicità geometrica è uno, ed 1 con molteplicità algebrica 2.

Inoltre,

$$\begin{aligned} m_g(1) &= 3 - \operatorname{rg}(\mathcal{M}_{\mathcal{B},\mathcal{B}}(T) - \mathrm{Id}_3) \\ &= 3 - \operatorname{rg}\left(\begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 2 & 4 & 2 \end{bmatrix}\right) \\ &= 2, \end{aligned}$$

quindi T è diagonalizzabile.

9.26 Diagonalizzazione di matrici

Sia $A \in M_{n\times n}(\mathbb{K})$ dove $\mathbb{K} = \mathbb{R}$ oppure $\mathbb{C}$. Sia $L_A : \mathbb{K}^n \longrightarrow \mathbb{K}^n$ l'endomorfismo associato alla matrice A, ovvero

$$L_A(X) = AX.$$

Definiamo autovettori e autovalori di A, gli autovettori e autovalori di L_A.

Uno scalare $\lambda \in \mathbb{K}$ è un *autovalore* di A se esiste un vettore $v \in \mathbb{K}^n$ non nullo tale che $Av = \lambda v$, ovvero $L_A(v) = \lambda v$.

Un vettore $v \in \mathbb{K}^n$ non nullo si dice un *autovettore* di A se esiste $\lambda \in \mathbb{K}$ tale che $Av = \lambda v$, ovvero

$$(A - \lambda \mathrm{Id}_n)v = 0.$$

Quindi $\lambda \in \mathbb{K}$ è una autovalore di A se e solamente se $\det(A - \lambda \mathrm{Id}_n) = 0$. Chiamaremo $p_A(t) = \det(A - t\mathrm{Id}_n)$ il *polinomio caratteristico* di A. Ricordiamo che $\mathcal{M}_{\mathcal{C},\mathcal{C}}(L_A) = A$, dove $\mathcal{C}$ è la base canonica di $\mathbb{K}^n$. Infatti,

$$L_A(e_i) = Ae_i = A^i,$$

per $i = 1, \ldots, n$. Quindi il polinomio caratteristico di A è il polinomio caratteristico di L_A. Analogamento possiamo definire l'autospazio relativo all'autovalore λ essendo

$$\begin{aligned} V_\lambda &= \{v \in \mathbb{K}^n : L_A(v) = \lambda v\} \\ &= \{v \in \mathbb{K}^n : Av = \lambda v\} \\ &= \{v \in \mathbb{K}^n : (A - \lambda \mathrm{Id}_n)v = 0\} \\ &= \mathrm{Sol}(A - \lambda \mathrm{Id}_n | 0_{\mathbb{K}^n}). \end{aligned}$$

La dimensione di V_λ è la *molteplicità geometrica* di λ. Applicando il Teorema di Rouché-Capelli, si ha

$$m_g(\lambda) = n - \mathrm{rg}(A - \lambda \mathrm{Id}_n).$$

La molteplicità algebrica di λ, che indicheremo con $m_a(\lambda)$, è il numero di volte che λ è radice del polinomio $p_A(t)$, ovvero il più grande numero naturale N tale che $(t-\lambda)^N$ divide $p_A(t)$. Poiché un autovalore di A è un autovalore di L_A, si ha

$$m_g(\lambda) \leq m_a(\lambda).$$

Inoltre, se $m_a(\lambda) = 1$, allora anche $m_g(\lambda) = 1$ (esercizio!).

Definizione 9.27. *Sia $A \in M_{n\times n}(\mathbb{K})$ dove $\mathbb{K} = \mathbb{R}$ oppure $\mathbb{C}$. Diremo che A è diagonalizzabile su $\mathbb{K}$ se esiste una matrice $P \in \mathrm{M}_{n\times n}(\mathbb{K})$ invertibile tale che $P^{-1}AP$ è una matrice diagonale.*

Quindi, una matrice quadrata A è diagonalizzabile se e solamente se A è simile ad una matrice diagonale.

Proposizione 9.28. *A è diagonalizzabile se e solamente se l'endomorfismo $L_A : \mathbb{K}^n \longrightarrow \mathbb{K}^n$ è diagonalizzabile.*

Dimostrazione. Poiché $\mathcal{M}_{\mathcal{C},\mathcal{C}}(L_A) = A$, dove $\mathcal{C}$ è la base canonica, L_A è diagonalizzabile se e solamente se A è simile ad una matrice diagonale e quindi se e solamente se A è diagonalizzabile.

Vediamo un'altra dimostrazione dove si descrive esplicitamente il legame fra la matrice invertibile P, la matrice diagonale D tale che $P^{-1}AP = D$ e autovettori e autovalori dell'endomorfismo L_A e quindi di A.

Supponiamo che A sia diagonalizzabile. Quindi esiste una matrice invertibile P e una matrice diagonale D tale che $P^{-1}AP = D$. Le colonne della matrice P formano una base di $\mathbb{K}^n$. Se indichiamo con $\mathcal{B} = \{P^1, \ldots, P^n\}$, tenendo in mente che $\mathcal{M}(\mathcal{C}, \mathcal{B}) = P$, allora

$$\begin{aligned}\mathcal{M}_{\mathcal{B},\mathcal{B}}(L_A) &= \mathcal{M}(\mathcal{B},\mathcal{C})\mathcal{M}_{\mathcal{C},\mathcal{C}}(L_A)\mathcal{M}(\mathcal{C},\mathcal{B}) \\ &= P^{-1}AP \\ &= D.\end{aligned}$$

Quindi abbiamo dimostrato che L_A è diagonalizzabile. In dettaglio abbiamo dimostrato che P^i è una autovettore di L_A relativo all'autovalore d_{ii}, dove $D = (d_{ij})_{\substack{1\leq i\leq n\\ 1\leq j\leq n}}$, per $i = 1, \ldots, n$.

Viceversa supponiamo che L_A sia diagonalizzabile. Sia $\mathcal{B} = \{v_1, \dots, v_n\}$ una base di $\mathbb{K}^n$ formata da autovettori di L_A. La matrice $P = (v_1, \dots, v_n)$ è invertibile poiché $\mathcal{B} = \{v_1, \dots, v_n\}$ è una base di $\mathbb{K}^n$. Vogliamo dimostrare che $P^{-1}AP$ è una matrice diagonale. Infatti

$$\begin{aligned} P^{-1}AP &= \mathcal{M}(\mathcal{B}, \mathcal{C})\mathcal{M}_{\mathcal{C},\mathcal{C}}(L_A)\mathcal{M}(\mathcal{C}, \mathcal{B}) \\ &= \mathcal{M}_{\mathcal{B},\mathcal{B}}(L_A). \end{aligned}$$

Quindi P è una matrice invertibile le cui colonne formano una base di autovettori di L_A e D è una matrice diagonale i cui elementi sulla diagonale principale sono gli autovalori di L_A e quindi di A. □

Applicando il Teorema 9.23 e la Proposizione 9.28 si ha il seguente criterio di diagonalizzazione di matrici.

Teorema 9.29. *Sia $A \in M_{n\times n}(\mathbb{K})$. Le seguenti condizioni sono equivalenti:*

a) A è diagonalizzabile su $\mathbb{K}$;

b) tutti gli autovalori di A sono in $\mathbb{K}$ e per ogni autovalore $\lambda \in sp(T)$, si ha $m_a(\lambda) = m_g(\lambda)$;

Corollario 9.30. *Sia $A \in \mathrm{M}_{n\times n}(\mathbb{K})$. Se A ha n autovalori in $\mathbb{K}$ distinti allora A è diagonalizzabile.*

Dimostrazione. Siano $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ autovalori distinti di A. Poiché il polinomio caratteristico di A ha grado n, la moleplicità algebrica di ciascuno autovalore è 1. Quindi anche la molteplicità geometrica di ciascun autovalore è 1 da cui segue che A è diagonalizzabile su $\mathbb{K}$. □

Osservazione 9.31.

a) Una matrice quadrata a coefficienti reali può essere diagonalizzabile su $\mathbb{C}$ ma non su $\mathbb{R}$.

b) Sia $A \in \mathrm{M}_{n\times n}(\mathbb{C})$. Per il Teorema fondamentale dell'algebra A ha n autovalori, non necessariamente distinti, complessi. Siano $\lambda_1, \dots, \lambda_k \in \mathbb{C}$ gli autovalori distinti di A. Allora

$$\sum_{i=1}^{k} m_a(\lambda_i) = n.$$

c) Sia $A \in \mathrm{M}_{n\times n}(\mathbb{R})$ e sia $p_A(t)$ il suo polinomio caratteristico e sia $\lambda \in \mathbb{C}$ un autovalore complesso. Allora anche $\overline{\lambda}$ è una radice di $p_A(t)$. Quindi il numero delle radici complesse di un polinomio a coefficienti reali è pari. Infatti, poiché $p_A(t)$ è un polinomio a coefficienti reali, si ha $\overline{p_A(t)} = p_A(\overline{t})$ da cui segue

$$p_A(\overline{\lambda}) = \overline{p_A(\lambda)} = 0.$$

d) Sia $A \in \mathrm{M}_{n\times n}(\mathbb{C})$. La matrice A ha esattamente n autovalori, non necessariamente distinti, che indicheremo con $\lambda_1, \ldots, \lambda_n$. Si può dimostrare che

$$\det(A) = \lambda_1 \cdots \lambda_n \qquad \mathrm{Tr}(A) = \lambda_1 + \cdots + \lambda_n.$$

9.32 Metodi di calcolo

Sia $A \in \mathrm{M}_{n\times n}(\mathbb{K})$. Se $\lambda \in \mathbb{K}$ è un autovalore, allora

$$V_\lambda = \mathrm{Sol}(A - \lambda \mathrm{Id}_n | 0_{\mathbb{K}^n})$$

Indichiamo con $\mathcal{B}_\lambda$ una base di V_λ. Se $\lambda_1, \cdots, \lambda_s \in \mathbb{K}$ sono autovalori di A distinti allora si può dimostrare che

$$\mathcal{B}_{\lambda_1} \cup \cdots \cup \mathcal{B}_{\lambda_s},$$

sono vettori linearmente indipendenti.

Vediamo quali sono i passi per stabilire se una matrice quadrata $A \in \mathrm{M}_{n\times n}(\mathbb{K})$ è diagonalizzabile su $\mathbb{K}$ ed in caso affermatico come determinare una matrice invertibile P e una matrice diagonale D tale che $P^{-1}AP = D$.

- Gli autovalori della matrice A sono le radici del polinomio caratteristico $p_A(t) = \det(A - t\mathrm{Id}_n)$. Se tutte le radici del polinomio caratteristico appartengono a $\mathbb{K}$, allora continuo. Altrimenti la matrice A non è diagonalizzabile.

- Sia $\lambda \in \mathbb{K}$ un autovalore. Allora $V_\lambda = \mathrm{Sol}(A - \lambda\mathrm{Id}_n|0_{\mathbb{K}^n})$ e $m_g(\lambda) = \dim V_\lambda = n - rg(A - \lambda\mathrm{Id}_n)$. Ricordiamo che la molteplicità algebrica di λ è il numero di volte che λ è radice. Inoltre $m_g(\lambda) \leq m_a(\lambda)$. Se $m_a(\lambda) = m_g(\lambda)$ per ogni autovalore λ allora la matrice è diagonalizzabile. Altrimenti no.

Supponiamo che A sia diagonalizzabile. Siano $\lambda_1, \dots, \lambda_s \in \mathbb{K}$ autovalori distinti di A. Allora

$$\mathcal{B} = \mathcal{B}_{\lambda_1} \cup \cdots \cup \mathcal{B}_{\lambda_s} = (v_1, \dots, v_n)$$

è una base di $\mathbb{K}^n$ formata da autovettori di A. Quindi la matrice $P = (v_1, \dots, v_n)$, i.e., la matrice le cui colonne sono i vettori $v_1, \dots, v_n$, è una matrice invertibile. Da adesso in poi indicheremo con μ_i l'autovalore corrispondente all'autovettore v_i. Quindi μ_i può essere uguale a μ_j anche se $i \neq j$. Affermiamo che

$$P^{-1}AP = D = \begin{pmatrix} \mu_1 & & \\ & \ddots & \\ & & \mu_n \end{pmatrix}$$

Per provare che le due matrici sono uguali, dimostreremo che le due matrici hanno le stesse colonne.

Siano $e_1, \dots, e_n$ la base canonica di $\mathbb{K}^n$. Ricordiamo che la i-esima colonna della matrice $P^{-1}AP$ è

$$(P^{-1}AP)^i = P^{-1}APe_i,$$

per $i = 1, \dots, n$.

Per definizione di P si ha $Pe_i = v_i$ per $i = 1, \dots, n$. Moltiplicando a destra e sinistra per P^{-1}, si ha $P^{-1}v_i = e_i$ per $i = 1, \dots, n$. Quindi

$$\begin{aligned}(P^{-1}AP)^i &= P^{-1}APe_i \\ &= P^{-1}Av_i \\ &= \mu_i P^{-1}v_i \\ &= \mu_i e_i,\end{aligned}$$

ovvero $P^{-1}AP = D$.

Esempio 9.33. *Sia*

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

$$p_A(t) = \det(A - t\mathrm{Id}_3) = \det \begin{pmatrix} -t & -2 & 1 \\ 2 & -5-t & 2 \\ -1 & 2 & -2-t \end{pmatrix} = -(t+5)(t+1)^2.$$

Gli autovalori di A sono -5 con molteplicità algebrica uno, e quindi anche la molteplicità geometrica è 1, e -1 con molteplicità algebrica due. Poiché

$$A + \mathrm{Id}_3 = \begin{pmatrix} 1 & -2 & 1 \\ 2 & -4 & 2 \\ -1 & 2 & -1 \end{pmatrix}$$

La molteplicità geometrica dell'autovalore 1 è

$$m_g(-1) = 3 - \mathrm{rg}(A + \mathrm{Id}_3) = 2$$

e quindi A è diagonalizzabile. Una base per l'autospazio relativo a -5 è

$$V_{-5} = \mathrm{Sol}(A + 5\mathrm{Id}|0_{\mathbb{R}^3}) = \mathcal{L}\left(\begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}\right)$$

mentre

$$V_{-1} = \mathrm{Sol}(A + \mathrm{Id}_3|0_{\mathbb{R}^3}) = \mathcal{L}\left(\begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}\right).$$

Quindi se $P = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 0 & 2 \\ 0 & -1 & -1 \end{pmatrix}$ *e* $D = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -5 \end{pmatrix}$, *allora si ha*

$$P^{-1}AP = D.$$

Esempio 9.34. *Sia*

$$A = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 0 & -1 \\ 0 & 1 & 1 \end{pmatrix}.$$

Il polinomio caratteristico di A è

$$\begin{aligned} p_A(t) &= \det(A - t\mathrm{Id}_3) \\ &= \det \begin{pmatrix} 1-t & -1 & 0 \\ -1 & -t & -1 \\ 0 & 1 & 1-t \end{pmatrix} \\ &= (1-t)(t^2+t+1) + (t-1) = -t(t-1)^2, \end{aligned}$$

e quindi gli autovalori di A sono 0 con molteplicità algebrica uno e 1 con molteplicità algebrica due. Poiché

$$A - \mathrm{Id}_3 = \begin{pmatrix} 0 & -1 & 0 \\ -1 & -1 & -1 \\ 0 & 1 & 0 \end{pmatrix},$$

si ha

$$m_g(1) = 3 - \mathrm{rg}(A - \mathrm{Id}_3) = 1,$$

ovvero la molteplicità geometrica dell'autovalore 1 *è uno e quindi* A *non è diagonalizzabile. Una base per l'autospazio relativo a* 0 *è*

$$V_0 = \mathrm{Sol}(A|0_{\mathbb{R}^3}) = \mathcal{L}\left(\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}\right)$$

mentre

$$V_1 = \mathrm{Sol}(A - \mathrm{Id}_3|0_{\mathbb{R}^3}) = \mathcal{L}\left(\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}\right).$$

Esempio 9.35. *Sia*

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

$$p_A(t) = \det\begin{pmatrix} -2-t & -1 & 3 \\ -5 & -2-t & 5 \\ -4 & -2 & 5-t \end{pmatrix} = (1-t)(t^2+1)$$

Gli autovalori di A *sono* 0 *con molteplicità algebrica uno e gli autovalori* $i, -i$ *con molteplicità algebrica uno rispettivamente. Quindi* A *non è diagonalizzabile su* $\mathbb{R}$ *ma è diagonalizzabile su* $\mathbb{C}$*. Gli autospazi della matrice* A*, pensata come matrice a coefficienti complessi, sono:*

$$V_1 = \mathrm{Sol}(A - \mathrm{Id}_3|0_{\mathbb{R}^3}) = \mathcal{L}\left(\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}\right),$$

$$V_i = \mathrm{Sol}(A - i\mathrm{Id}_3|0_{\mathbb{R}^3}) = \mathcal{L}\left(\begin{bmatrix} -i \\ 2-i \\ 1-i \end{bmatrix}\right).$$

$$V_{-i} = \mathrm{Sol}(A + i\mathrm{Id}_3|0_{\mathbb{R}^3}) = \mathcal{L}\left(\begin{bmatrix} i \\ 2+i \\ 1+i \end{bmatrix}\right).$$

Se indichiamo con

$$P = \begin{pmatrix} 1 & -i & i \\ 0 & 2-i & 2+i \\ 1 & 1-i & 1+i \end{pmatrix}, \qquad D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & i & 0 \\ 0 & 0 & -i \end{pmatrix},$$

si ha
$$P^{-1}AP = D.$$

Sia $T : V \longrightarrow V$ e sia $\mathcal{B}$ una base di V. La matrice $\mathcal{M}_{\mathcal{B},\mathcal{B}}(T)$ è l'unica matrice $n \times n$ tale che
$$[T(v)]_{\mathcal{B}} = \mathcal{M}_{\mathcal{B},\mathcal{B}}(T)[v]_{\mathcal{B}}.$$

- $p_T(t) = p_{\mathcal{M}_{\mathcal{B},\mathcal{B}}(T)}(t)$. Quindi $\lambda \in \mathbb{K}$ è un autovalore di T se e solamente se $\lambda \in \mathbb{K}$ è un autovalore di $\mathcal{M}_{\mathcal{B},\mathcal{B}}(T)$. Inoltre la molteplicità algebrica di λ come autovalore di T coincide con la molteplicità algebrica di λ come autovalore di $\mathcal{M}_{\mathcal{B},\mathcal{B}}(T)$.

- v è autovettore relativo all'autovalore λ di T se e solamente se $T(v) = \lambda v$ se e solamente se $[T(v)]_{\mathcal{B}} = \lambda[v]_{\mathcal{B}}$. Poiché $[T(v)]_{\mathcal{B}} = \mathcal{M}_{\mathcal{B},\mathcal{B}}(T)[v]_{\mathcal{B}}$ si ha v è autovalore di T relativo a λ se e solamente se $[v]_{\mathcal{B}}$ è un autovettore di $\mathcal{M}_{\mathcal{B},\mathcal{B}}(T)$ relativo all'autovalore λ. Infine la molteplicità geometrica di λ come autovalore di T coincide con la molteplicità geometrica di λ come autovalore di $\mathcal{M}_{\mathcal{B},\mathcal{B}}(T)$.

- Per i punti precedenti T è un endomorfismo diagonalizzabile se e solamente se $\mathcal{M}_{\mathcal{B},\mathcal{B}}(T)$ è diagonalizzabile.

Se $\mathcal{M}_{\mathcal{B},\mathcal{B}}(T)$ è diagonalizzabile, allora per i punti precedenti siamo in grado di determinare una base $\mathcal{C} = \{w_1, \ldots, w_n\}$ formata da autovettori di $\mathcal{M}_{\mathcal{B},\mathcal{B}}(T)$. Se
$$w_1 = \begin{bmatrix} a_{11} \\ \vdots \\ a_{n1} \end{bmatrix}, \ldots, w_n = \begin{bmatrix} a_{1n} \\ \vdots \\ a_{nn} \end{bmatrix},$$
allora una base di V formata da autovettori di T è
$$\left\{\mathcal{F}_{\mathcal{B}}^{-1}(w_1) =, \ldots, \mathcal{F}_{\mathcal{B}}^{-1}(w_n)\right\} = \{a_{11}v_1 + \cdots + a_{n1}v_n, \ldots, a_{1n}v_1 + \cdots + a_{nn}v_n\}.$$

Esempio 9.36. *Sia $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ l'applicazione lineare definita da:*
$$T\left(\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 0 \\ -2 \\ 2 \end{bmatrix}, T\left(\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} -1 \\ -3 \\ 3 \end{bmatrix}, T\left(\begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} -1 \\ -3 \\ -4 \end{bmatrix}.$$

Nell'esempio 9.25 abbiamo dimostrato che T è diagonalizzabile. Adesso, vogliamo calcolare una base di $\mathbb{R}^3$ formata da autovettori di T seguendo i due metodi sviluppati nell'esempio citato. Ricordiamo che T ha due autovalori:

1 *con molteplicità algebrica e geometrica due e* 2 *con molteplicità algebrica e geometrica uno. Sia* $\mathcal{C}$ *la base canonica e sia* $\mathcal{B}$ *la base su cui è definita* T.

I metodo. Poiché

$$(T - 2\mathrm{Id}_{\mathbb{R}^3})X = \mathcal{M}_{\mathcal{C},\mathcal{C}}(T - 2\mathrm{Id}_{\mathbb{R}^3})X = (\mathcal{M}_{\mathcal{C},\mathcal{C}}(T) - 2\mathrm{Id}_3)X,$$

rispettivamente

$$(T - \mathrm{Id}_{\mathbb{R}^3})X = \mathcal{M}_{\mathcal{C},\mathcal{C}}(T - \mathrm{Id}_{\mathbb{R}^3})X = (\mathcal{M}_{\mathcal{C},\mathcal{C}}(T) - \mathrm{Id}_3)X,$$

si ha

$$\begin{aligned} V_2 &= \mathrm{Sol}(\mathcal{M}_{\mathcal{C},\mathcal{C}}(T) - 2\mathrm{Id}_3 | 0_{\mathbb{R}^3}) \\ &= \mathrm{Sol}\left(\left[\begin{array}{ccc|c} 0 & -1 & -2 & 0 \\ 2 & -3 & -4 & 0 \\ -1 & 1 & 1 & 0 \end{array}\right]\right) \\ &= \mathcal{L}\left(\begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}\right), \end{aligned}$$

rispettivamente

$$\begin{aligned} V_1 &= \mathrm{Sol}(\mathcal{M}_{\mathcal{C},\mathcal{C}}(T) - \mathrm{Id}_3 | 0_{\mathbb{R}^3}) \\ &= \mathrm{Sol}\left(\left[\begin{array}{ccc|c} 1 & -1 & -2 & 0 \\ 2 & -2 & -4 & 0 \\ -1 & 1 & 2 & 0 \end{array}\right]\right) \\ &= \mathcal{L}\left(\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ -1 \end{bmatrix}\right). \end{aligned}$$

Quindi

$$\left\{\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}\right\}$$

formano una base di $\mathbb{R}^3$ *formata da autovettori di* T.

II metodo. Ricordiamo che

$$\begin{aligned} [(T - \mathrm{Id}_{\mathbb{R}^3})(X)]_{\mathcal{B}} &= \mathcal{M}_{\mathcal{B},\mathcal{B}}(T - \mathrm{Id}_{\mathbb{R}^3})[X]_{\mathcal{B}} \\ &= \mathcal{M}_{\mathcal{B},\mathcal{B}}(T) - \mathrm{Id}_3)[X]_{\mathcal{B}}. \end{aligned}$$

Primo passo, calcoliamo gli autospazi della matrice $\mathcal{M}_{\mathcal{B},\mathcal{B}}(T)$.

$$\begin{aligned}\mathrm{Sol}(\mathcal{M}_{\mathcal{B},\mathcal{B}}(T) - \mathrm{Id}_3 | 0_{\mathbb{R}^3}) &= \mathrm{Sol}\left(\left[\begin{array}{ccc|c} -1 & -2 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ 2 & 4 & 2 & 0 \end{array}\right]\right) \\ &= \mathcal{L}\left(\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}\right)\end{aligned}$$

e quindi

$$\begin{aligned}V_1 &= \mathcal{L}\left(\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} - \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}\right) \\ &= \mathcal{L}\left(\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -2 \\ 1 \end{bmatrix}\right).\end{aligned}$$

Analogamente

$$\begin{aligned}\mathrm{Sol}(\mathcal{M}_{\mathcal{B},\mathcal{B}}(T) - 2\mathrm{Id}_3 | 0_{\mathbb{R}^3}) &= \mathrm{Sol}\left(\left[\begin{array}{ccc|c} -2 & -2 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ 2 & 4 & 1 & 0 \end{array}\right]\right) \\ &= \mathcal{L}\left(\begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix}\right),\end{aligned}$$

da cui segue

$$\begin{aligned}V_2 &= \mathcal{L}\left(\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} - 2\begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}\right) \\ &= \mathcal{L}\left(\begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}\right)\end{aligned}$$

Riassumendo,

$$\mathcal{B}' = \left\{\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}\right\}$$

è una base di $\mathbb{R}^3$ formata da autovettori di T e

$$\mathcal{M}_{\mathcal{B}',\mathcal{B}'}(T) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

formano una base di $\mathbb{R}^3$ formata da autovettori di T

9.37 Teorema spettrale

In questa sezione dimostreremo che una matrice simmetrica, a coefficienti reali, è sempre diagonalizzabile su $\mathbb{R}$. In dettaglio, proveremo che esiste una base ortonormale di $\mathbb{R}^n$ formata da autovettori di A. Il primo lemma prova che gli autovalori di una matrice simmetrica sono numeri reali.

Lemma 9.38. *Gli autovalori di una matrice $A \in \mathrm{M}_{n\times n}(\mathbb{R})$ simmetrica sono numeri reali.*

Dimostrazione. Poiché una matrice reale è anche una matrice complessa, A induce un endomorfismo $L_A : \mathbb{C}^n \longrightarrow \mathbb{C}^n$ così definito: $L_A(X) = AX$. Se $\mathcal{C}$ è la base canonica di $\mathbb{C}^n$, allora $\mathcal{M}_{\mathcal{C},\mathcal{C}}(L_A) = A$. Quindi gli autovalori dell'endomorfismo L_A sono gli autovalori di A.

Sia $\langle\cdot,\cdot\rangle$ il prodotto Hermitiano canonico di $\mathbb{C}^n$. Ricordiamo che

$$\langle X, Y\rangle = X^T\overline{Y}.$$

Affermiamo che per ogni $X, Y \in \mathbb{C}^n$, si ha

$$\langle AX, Y\rangle = \langle X, AY\rangle.$$

Infatti, tenendo in mente che $A = \overline{A}$, si ha

$$\begin{aligned}\langle AX, Y\rangle &= (AX)^T\overline{Y} \\ &= X^T A^T \overline{Y} \\ &= X^T A\overline{Y} \\ &= X^T \overline{AY} \\ &= \langle X, AY\rangle.\end{aligned}$$

Sia $X \in \mathbb{C}^n$ un autovettore di A relativo all'autovalore λ. Quindi $X \neq 0_{\mathbb{C}^n}$ e $AX = \lambda X$. Applicando la formula anteriore si ha

$$\langle AX, X\rangle = \langle X, AX\rangle.$$

Quindi

$$\begin{aligned}\langle AX, X\rangle &= \langle X, AX\rangle\\ \langle \lambda X, X\rangle &= \langle X, \lambda X\rangle\\ \lambda\langle X, X\rangle &= \overline{\lambda}\langle X, X\rangle\end{aligned}$$

Poiché $\langle X, X\rangle \neq 0$ si ha $\lambda = \overline{\lambda}$, ovvero $\lambda \in \mathbb{R}$. $\square$

La seguente proposizione garantisce che autospazi corrispondenti a autovettori distinti sono fra loro ortogonali.

Proposizione 9.39. *Sia A una matrice simmetrica di ordine n e siano v e w autovettori corrispondenti ad autovalori distinti $\lambda, \mu \in \mathbb{R}$. Allora $\langle v, w\rangle = 0$. Quindi $V_\lambda \subset V_\mu^\perp$ se $\lambda \neq \mu$.*

Dimostrazione. Sia $\langle X, Y\rangle = X^T Y$ il prodotto scalare canonico di $\mathbb{R}^n$. Se A è una matrice simmetrica allora vale la seguente formula:

$$\langle AX, Y\rangle = \langle X, AY\rangle.$$

Infatti

$$\begin{aligned}\langle AX, Y\rangle &= (AX)^T Y\\ &= X^T A^T Y\\ &= X^T (AY) \text{ essendo } A \text{ simmetrica)}\\ &= \langle X, AY\rangle.\end{aligned}$$

Siano v, w autovettori di A relativi agli autovalori λ e μ rispettivamente. Applicando la formula anteriore si ha

$$\langle Av, w\rangle = \langle v, Aw\rangle.$$

Poiché v, rispettivamente w, è un autovettore di A relativo a λ, rispettivamente a μ, si ha

$$\begin{aligned}\langle Av, w\rangle &= \langle v, Aw\rangle\\ \langle \lambda v, w\rangle &= \langle v, \mu w\rangle\\ \lambda\langle v, w\rangle &= \mu\langle v, w\rangle\\ (\lambda - \mu)\langle v, w\rangle &= 0\end{aligned}$$

Poiché λ e μ sono distinti, si ha $\langle v, w\rangle = 0$ concludendo la dimostrazione. $\square$

Teorema 9.40 (Teorema spettrale)**.** *Sia $A \in M_{n\times n}(\mathbb{R})$ una matrice simmetrica. Allora esiste una matrice ortogonale P tale che $P^T AP$ è diagonale. Ovvero esiste una base ortonormale di $\mathbb{R}^n$ formata da autovettori di A.*

Facoltativa. Per induzione su n. Se $n = 1$ è vero. Supponiamo di aver dimostrato il teorema per n e proviamolo per $n+1$. Per il lemma anteriore una matrice simmetrica ha tutti gli autovalori reali. Sia $\lambda \in \mathbb{R}$ un autovalore. Allora esiste $v \in \mathbb{R}^n$ non nullo tale che $Av = \lambda_1 v$. Denotiamo con $v_1 = \frac{v}{\|v\|}$. Completiamo v_1 ad una base ortonormale di $\mathbb{R}^n$ che denotiamo con $\mathcal{B} = \{v_1, \ldots, v_{n+1}\}$. Allora

- $P = \mathcal{M}(\mathcal{C}, \mathcal{B}) = (v_1, \ldots, v_{n+1})$ è una matrice ortogonale;
- $P^T AP = \left(\begin{array}{c|ccc} \lambda_1 & 0 & \cdots & 0 \\ \hline \vdots & & B & \\ 0 & & & \end{array}\right)$

dove $B \in M_{n\times n}(\mathbb{R})$ è una matrice simmetrica. Per ipotesi induttiva esiste una matrice ortogonale Q di ordine n tale che

$$Q^T BQ = \begin{pmatrix} \lambda_2 & & \\ & \ddots & \\ & & \lambda_{n+1} \end{pmatrix}.$$

Sia $\tilde{Q} = \left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline \vdots & & Q & \\ 0 & & & \end{array}\right) \in M_{n+1\times n+1}(\mathbb{R})$. $\tilde{Q}$ è una matrice ortogonale da cui segue che anche $P\tilde{Q}$ è una matrice ortogonale. Inoltre

$$(P\tilde{Q})^T A(P\tilde{Q}) = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_{n+1} \end{pmatrix} = D.$$

Quindi la base $\mathcal{C} = \{w_1, \ldots, w_{n+1}\}$ dove $w_i = (P\tilde{Q})^i$ per $i = 1, \ldots, n+1$ è una base ortonormale formata da autovettori di A. □

Corollario 9.41. *Sia $A \in M_{n\times n}(\mathbb{R})$. A è simmetrica se e solamente se esiste una matrice ortogonale P tale che $P^T AP$ è una matrice diagonale.*

Dimostrazione. Una direzione è il teorema spettrale. Viceversa se $P^T AP = D$ matrice diagonale, con P ortogonale, allora

$$A = PDP^T$$

e quindi

$$A^T = (PDP^T)^T = PD^T P^T = PDP^T = A$$

ovvero $A = A^T$ come si voleva dimostrare. □

Osservazione 9.42. *Eistono matrici $A \in \mathrm{M}_{n\times n}(\mathbb{C})$ tali che $A = A^T$ non diagonalizzabile. Infatti*

$$A = \begin{pmatrix} 1 & i \\ i & -1 \end{pmatrix}.$$

verifica $A = A^T$. $p_A(t) = t^2$ quindi ha un autovalore $\lambda = 0$ con molteplicità algebrica 2 e geometrica 1 da cui segue che A non è diagonalizzabile. Quindi il Teorema spettrale non vale per matrici complesse.

9.43 Metodi di calcolo

Sia $A \in \mathrm{M}_{n\times n}(\mathbb{R})$ una matrice simmetrica. Il teorema spettrale garantisce che A è diagonalizzabile. Vediamo come calcolare una matrice ortogonale P e una matrice diagonale D tale che $P^T AP = D$.

Poiché A è diagonalizzabile possiamo calcolare una base $\mathcal{B} = \{v_1, \ldots, v_n\}$ di $\mathbb{R}^n$ formata da autovettori di A come nella sezione 9.32. Se indichiamo con $Q = (v_1, \ldots, v_n)$, allora $Q^{-1}AQ = D$ è diagonale. La matrice Q non è, in generale, ortogonale. Per ottenere un matrice ortogonale possiamo utlizzare il Procedimento di Gram-Schimdt. Infatti, se applichiamo il procedimento di Gram-Schimdt alla base $\mathcal{B} = \{v_1, \ldots, v_n\}$, ottengo una base ortogonale $\mathcal{C} = \{w_1, \ldots, w_n\}$ di $\mathbb{R}^n$. Infine, dividendo ciascun vettore per la sua norma, ottengo una base $\mathcal{C}' = \left\{\frac{w_1}{\|w_1\|}, \ldots, \frac{w_n}{\|w_n\|}\right\}$ ortonormale di $\mathbb{R}^n$. Poiché autovettori di A corrispondenti ad autovalori distinti sono fra loro ortogonali, la base $\mathcal{C}'$ è ancora una base formata da autovettori di A. Se la matrice A fosse diagonalizzabile ma non fosse simmetrica i vettori della base $\mathcal{C}'$ non sarebbero più autovettori di A. Quindi se applichiamo il procedimento di Gram-Schidt ad una base di autovettori di una matrice diagonalizzabile A, tale base rimane una base formata da autovettori se e solamente se A è simmetrica.

Sia $P = (\frac{w_1}{\|w_1\|}, \ldots, \frac{w_n}{\|w_n\|})$. P è una matrice ortogonale poiché le colonne formano una base ortonormale di autovettori di A. Infine, tenendo in mente che $P^{-1} = P^T$, si ha $P^T AP = D$ dove D è la stessa matrice diagonale ottenuta in precedenza. (Perché?)

Esempio 9.44. *Sia*

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

$$p_A(t) = \det(A - t\mathrm{Id}_3) = \det \begin{pmatrix} 2-t & 1 & 1 \\ 1 & 2-t & 1 \\ 1 & 1 & 2-t \end{pmatrix} = -(t-4)(t-1)^2.$$

Gli autovalori di A sono 4 con molteplicità algebrica uno e 1 con molteplicità algebrica due. Una base per l'autospazio relativo a 1 è

$$V_1 = \mathrm{Sol}(A - \mathrm{Id}|0_{\mathbb{R}^3}) = \mathcal{L}\left(\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} \right)$$

mentre

$$V_4 = \mathrm{Sol}(A - 4\mathrm{Id}_3|0_{\mathbb{R}^3}) = \mathcal{L}\left(\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right).$$

Quindi se $P = \begin{pmatrix} 1 & 0 & 1 \\ -1 & 1 & 1 \\ 0 & -1 & 1 \end{pmatrix}$ *e* $D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{pmatrix}$, *allora*

$$P^{-1}AP = D.$$

La matrice P NON E' ORTOGONALE. Per ottenere una matrice ortogonale U tale che $U^TAU = D$, posso applicare il procedimento di Gram-Schimdt alla base

$$\left\{ \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\} \rightsquigarrow \left\{ \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1/2 \\ 1/2 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$$

e poi dividiamo ciascun vettore per la sua norma

$$\left\{ \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \\ 0 \end{bmatrix}, \begin{bmatrix} 1/\sqrt{6} \\ 1/\sqrt{6} \\ -2/\sqrt{6} \end{bmatrix}, \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix} \right\}.$$

La matrice

$$U = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{6} & 1/\sqrt{3} \\ -1/\sqrt{2} & 1/\sqrt{6} & 1/\sqrt{3} \\ 0 & -2/\sqrt{6} & 1/\sqrt{3} \end{pmatrix}$$

è ortogonale, le colonne formano una base ortonormale di autovettori di A e verifica

$$U^TAU = D.$$

Capitolo 10

Matrici ortogonali

10.1 Proprietà delle matrici ortogonali

Sia $A \in \mathrm{M}_{n\times n}(\mathbb{R})$. Sia $\langle \cdot, \cdot \rangle$ il prodotto scalare canonico di $\mathbb{R}^n$.

Proposizione 10.2. *A è ortogonale se e solamente se per ogni $X, Y \in \mathbb{R}^n$ si ha $\langle AX, AY\rangle = \langle X, Y\rangle$.*

Facoltativa. Supponiamo che A sia ortogonale e siano $X, Y \in \mathbb{R}^n$. Allora

$$\langle AX, AY\rangle = (AX)^T AY = X^T A^T AY = X^T Y = \langle X, Y\rangle.$$

Viceversa, sia $\mathcal{C} = \{e_1, \ldots, e_n\}$ la base canonica. Poiché $Ae_i = A^i$, i.e., la i-esmia colonna si ha

$$\langle Ae_i, Ae_j\rangle = \langle A^i, A^j\rangle = \langle e_i, e_j\rangle$$

per ogni $1 \leq i \leq j \leq n$. Quindi $A^1, \ldots, A^n$ formano una base ortonormale ovvero A è ortogonale. □

Corollario 10.3. *Sia $A \in \mathrm{M}_{n\times n}(\mathbb{R})$ una matrice ortogonale. Allora l'endomorfimo $L_A : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ preserva la lunghezza di un vettore, l'angolo fra due vettori non nulli e infine la distanza fra due vettori.*

Facoltativa. Siano $X, Y \in \mathbb{R}^n$. Poiché

$$\langle L_A(X), L_A(Y)\rangle = \langle AX, AY\rangle = \langle X, Y\rangle,$$

si ha $\parallel L_A(X) \parallel = \parallel X \parallel$ ed

$$\langle L_A(X)-L_A(Y), L_A(X)-L_A(Y)\rangle = \langle L_A(X-Y), L_A(X-Y)\rangle = \langle X-Y, X-Y\rangle.$$

Poiché la distanza fra X e Y è per definizione $d(X,Y) = \| X - Y \|$, si ha che L_A preserva la distanza.

Siano X e Y vettori non nulli. Allora

$$\frac{\langle L_A(X), L_A(Y)\rangle}{\| L_A(X) \| \| L_A(Y) \|} = \frac{\langle X, Y\rangle}{\| X \| \| Y \|},$$

i.e., L_A preserva l'angolo fra due vettori. $\square$

Le matrici ortogonali non sono in generale diagonalizzabili su $\mathbb{R}$. Infatti, sia

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Il polinomio caratteristi di A è $t^2 + 1$ che non ammette radici su $\mathbb{R}$. Gli autovalori reali di una matrice ortogonale, se esistono, sono ± 1.

Proposizione 10.4. *Sia A un matrice ortogonale e sia $\lambda \in \mathbb{C}$ un autovalore. Allora $|\lambda| = 1$. In particolare se $\lambda \in \mathbb{R}$, allora $\lambda = \pm 1$.*

Facoltativa. Sia $L_A : \mathbb{C}^n \longrightarrow \mathbb{C}^n$. Gli autovalori di A sono gli autovalori di L_A. Affermiamo che per ogni $Z, W \in \mathbb{C}^n$ si ha

$$\langle AZ, AW\rangle = \langle Z, W\rangle,$$

dove $\langle \cdot, \cdot \rangle$ è il prodotto Hermitiano canonico. Infatti

$$\langle AZ, AW\rangle = (AZ)^T \overline{AW} = Z^T A^T A \overline{W} = Z^T \overline{W} = \langle Z, W\rangle.$$

Sia $\lambda \in \mathbb{C}$ autovalore e sia $Z \neq 0$ autovettore relativo a λ. Allora

$$\langle AZ, AZ\rangle = \langle \lambda Z, \lambda Z\rangle = |\lambda|^2 \langle Z, Z\rangle.$$

Per la formula anteriore si ha

$$\langle AZ, AZ\rangle = \langle Z, Z\rangle,$$

ovvero

$$|\lambda|^2 \langle Z, Z\rangle = \langle Z, Z\rangle.$$

Poiché $\langle Z, Z\rangle \neq 0$, si ha $|\lambda|^2 = 1$. $\square$

Una conseguenza del risultato anteriore è che il determinante di una matrice ortogonale è ± 1, cosa che avevamo già dimostrato utilizzando la formula di Binet.

Corollario 10.5. *Sia A una matrice ortogonale. Allora* $\det(A) = \pm 1$.

Facoltativa. Poiché A è una matrice a coefficienti reali, se $\lambda \in \mathbb{C}$ è autovalore, allora anche $\overline{\lambda}$ è autovalore di A. Poiché A è ortogonale. allora $\lambda\overline{\lambda} = 1$. Tenendo i mente che $\det(A)$ è il prodotto dei suoi autovalori e che gli autovalori reali possono essere ± 1, si ha che $\det(A) = \pm 1$. $\square$

Definizione 10.6. *Sia $W \subset \mathbb{R}^n$ un sottospazio vettoriale e sia $A \in \mathrm{M}_{n\times n}(\mathbb{R})$. Diremo che W è A-invariante se per ogni $w \in W$ si ha $Aw \in W$.*

È facile verificare che se A è invertibile, allora W è un sottospazio di $\mathbb{R}^n$ A-invariante se e solamente se W è A^{-1}-invariante. Se A è ortogonale vale la seguente proposizione.

Proposizione 10.7. *Sia A una matrice ortogonale e sia Sia $W \subset \mathbb{R}^n$ un sottospazio A-invariante. Allora $W^{\perp}$ è A-invariante.*

Facoltativa. Sia $Z \in W^{\perp}$. Vogliamo dimostrare che $AZ \in W^{\perp}$, ovvero $\langle AZ, w\rangle = 0$ per ogni $w \in W$. Poiché A^T è ancora una matrice ortogonale (perché?) si ha

$$\langle AZ, w\rangle = \langle A^T AZ, A^T w\rangle = \langle Z, A^T w\rangle = 0,$$

dove l'ultima uguaglianza segue dal fatto che essendo $A^T = A^{-1}$, W è A^T invariante. Riassumendo, abbiamo provato che $W^{\perp}$ è A-invariante. $\square$

10.8 Riflessioni e rotazioni del piano

Sia $L_A : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ con A matrice ortogonale. Se

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

dall'equazione $A^T A = \mathrm{Id}_2$, ne segue che

$$\begin{cases} a^2 + c^2 = 1 \\ ab + cd = 0 \\ b^2 + d^2 = 1 \end{cases}$$

Quindi esistono $\theta, \psi \in [0, 2\pi]$ tale che

$$\begin{array}{ll} a = \cos\theta & c = \sin\theta \\ b = \cos\psi & d = \sin\psi \end{array}.$$

Inoltre dalla secondo equazione si deduce che $\cos\theta\cos\psi + \sin\theta\sin\psi = \cos(\psi - \theta) = 0$, ovvero

$$\psi = \theta + \frac{\pi}{2} + k\pi.$$

Se k è pari allora $\cos\psi = \cos(\theta + \frac{\pi}{2}) = -\sin(\theta)$ e $\sin\psi = \sin(\theta + \frac{\pi}{2}) = \cos(\theta)$, ovvero

$$A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix},$$

e $\det(A) = 1$. Se invece k è dispari, allora $\cos\psi = \cos(\theta + \frac{3\pi}{2}) = \sin(\theta)$ e $\sin\psi = \sin(\theta + \frac{3\pi}{2}) = -\cos(\theta)$, ovvero

$$A = \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix},$$

e quindi A è simmetrica e $\det(A) = -1$. Riassumendo, abbiamo dimostrato che se A è una matrice ortogonale di formato due per due con $\det(A) = 1$, allora esiste $\theta \in \mathbb{R}$ tale che

$$A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

Se invece, $\det(A) = -1$, allora esiste $\theta \in \mathbb{R}$ tale che

$$A = \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix},$$

Adesso vogliamo studiare l'endomorfismo associato a A, ovvero $L_A : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$. Se $\det(A) = -1$, allora

$$A = \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix},$$

per un certo $\theta \in [0, 2\pi]$. A è simmetrica, quindi diagonalizzabile, e gli autovalori di A sono sono le radici del polinomio $p_A(t) = t^2 - 1$, cioé $1, -1$. L'autospazio relativo all'autovalore 1 sono le soluzione del sistema lineare omogeneo:

$$\begin{bmatrix} \cos\theta - 1 & \sin\theta \\ \sin\theta & -\cos\theta - 1 \end{bmatrix}.$$

Poiché $\dim V_1 = 1$ basta risolvere una delle due equazione, per esempio:

$$(\cos\theta - 1)x + (\sin\theta)y = 0.$$

Tenendo in mente che $\cos\theta - 1 = -2\sin^2(\frac{\theta}{2})$ e $\sin\theta = 2\sin\frac{\theta}{2}\cos\frac{\theta}{2}$, l'equazione diventa

$$-2\sin^2\frac{\theta}{2}x + 2\sin\frac{\theta}{2}\cos\frac{\theta}{2}y = 0,$$

e quindi

$$V_1 = \mathcal{L}(\begin{bmatrix} \cos\frac{\theta}{2} \\ \sin\frac{\theta}{2} \end{bmatrix}).$$

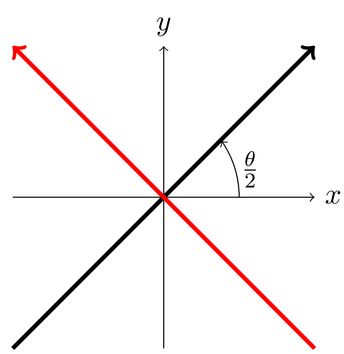
Poiché $V_{-1} = V_1^{\perp}$, si ha

$$V_{-1} = \mathcal{L}(\begin{bmatrix} -\sin\frac{\theta}{2} \\ \cos\frac{\theta}{2} \end{bmatrix}).$$

La base $\mathcal{B} = \left\{\begin{bmatrix} \cos\frac{\theta}{2} \\ \sin\frac{\theta}{2} \end{bmatrix}, \begin{bmatrix} -\sin\frac{\theta}{2} \\ \cos\frac{\theta}{2} \end{bmatrix}\right\}$ è una base ortonormale formata da autovettori di A. Quindi

$$\mathcal{M}_{\mathcal{B},\mathcal{B}}(L_A) = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix},$$

per cui l'endomorfismo L_A è la riflessione rispetto alla retta passante per l'origine e che ha come vettore direttore $\begin{bmatrix} \cos\frac{\theta}{2} \\ \sin\frac{\theta}{2} \end{bmatrix}$.



Se $\det(A) = 1$, allora

$$A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

per un certo $\theta \in [0, 2\pi]$. Poiché $p_A(t) = t^2 - 2\cos\theta + 1$, L_A non è diagonalizzabile se $\theta \neq 0, -\pi$, che corrispondono a Id_2 e $-\mathrm{Id}_2$. Vogliamo dimostrare che L_A è un rotazione di angole θ. È facile verificare che

$$\langle Ae_1, e_1\rangle = \langle Ae_2, e_2\rangle = \cos\theta,$$

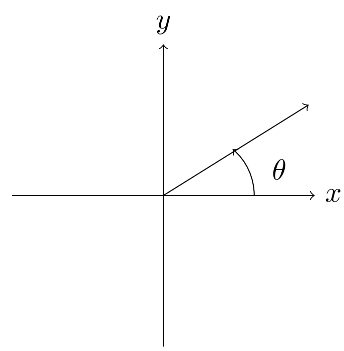
dove $\langle \cdot, \cdot \rangle$ è il prodotto scalare standard di $\mathbb{R}^2$. Quindi, se $v = \begin{bmatrix} x \\ y \end{bmatrix} = xe_1 + ye_2 \in \mathbb{R}^3$, si ha

$$\begin{aligned}\langle Av, v\rangle &= x^2\langle Ae_1, e_1\rangle + xy\langle Ae_1, e_2\rangle + xy\langle Ae_2, e_1\rangle + y^2\langle Ae_2, e_2\rangle \\ &= (x^2 + y^2)\cos\theta + xy\sin\theta - xy\sin\theta \\ &= \| v \|^2 \cos\theta.\end{aligned}$$

Quindi

$$\frac{\langle Av, v\rangle}{\| v \| \| Av \|} = \cos\theta,$$

ovvero l'endomorfismo L_A è una rotazione attorno all'origine di angolo θ.



Proposizione 10.9. *Sia* $A \in M_{2\times 2}(\mathbb{R})$ *una matrice ortogonale e sia* L_A *l'applicazione lineare indotta.* L_A *è una rotazione se e solamente se* $\det(A) = 1$*;* L_A *è una riflessione se e solamente se* $\det(A) = -1$*.*

Corollario 10.10. *La composizione di due rotazioni è ancora una rotazione; la composizione di due riflessioni è una rotazione; la composizione di una rotazione ed una riflessione è una riflessione.*

Dimostrazione. Noi sappiamo che $L_A \circ L_B = L_{AB}$ e che per il teorema di Binet $\det(AB) = \det(A)\det(B)$. Dalla discussione precedente segue la tesi. $\square$

Capitolo 11

Appendice

Gli argomenti trattati in questa capitolo non fanno parte del programma del corso di Geometria $1a$.

11.1 Matrici ortogonali di formato 3×3

Sia $A \in \mathrm{M}_{3\times 3}(\mathbb{R})$ ortogonale.

Lemma 11.2. *Se* $\det(A) = 1$, *rispettivamente* $\det(A) = -1$ *allora* $\lambda = 1$, *rispettivamente* $\lambda = -1$ *è un autovalore.*

Facoltativa. Sia A una matrice ortogonale di formato 3×3 con determinante 1. Allora

$$\begin{aligned} A - \mathrm{Id}_3 &= A - AA^T \\ &= A(\mathrm{Id}_3 - A^T) \\ &= A(\mathrm{Id}_3 - A)^T \\ &= -A(A - \mathrm{Id}_3)^T \end{aligned}$$

Quindi

$$\det(A - \mathrm{Id}_3) = \det(-A)\det(A - \mathrm{Id}_3) = -(\det(A - \mathrm{Id}_3),$$

poiché $\det(-A) = (-1)^3 \det(A) = -1$. Quindi $\det(A - \mathrm{Id}_3) = 0$, per cui 1 è un autovalore di A. Anologamente, se $\det(A) = -1$. si ha

$$\begin{aligned} A + \mathrm{Id} &= A + AA^T \\ &= A(\mathrm{Id} + A^T) \\ &= A(A + \mathrm{Id})^T \end{aligned}$$

da cui segue $\det(A + \mathrm{Id}) = 0$. $\square$

Definizione 11.3. *Una rotazione di $\mathbb{R}^3$ intorno all'origine è una applicazione lineare $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ tale che:*

- $\langle T(v), T(w)\rangle = \langle v, w\rangle$*;*
- *esiste $v \in V$ tale che $T(v) = v$;*
- *T agisce come una rotazione sul piano α passante per l'origine e normale al vettore v.*

Proposizione 11.4. *$T : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ è una rotazione attorno all'origine se e solamente se esiste una matrice ortogonale di ordine 3 con determinante uguale a 1, tale che $T = L_A$.*

Facoltativa. Sia A una matrice ortogonale di ordine 3 con determinante uguale a 1. Vogliamo dimostrare che $T = L_A$ è una rotazione. Per la proposizione 10.2 L_A soddisfa la prima proprietà ed Il lemma anteriore garantisce che 1 è un autovalore di T. Quindi esiste $v \in \mathbb{R}^3$ non nullo tale che $T(v) = L_A v = Av = v$. Completiamo $v_1 = \frac{v}{\|v\|}$ a base ortonornale di $\mathbb{R}^3$ che indichiamo con $\mathcal{B} = \{v_1, w_1, w_2\}$. La matrice $\mathcal{M}(\mathcal{C}, \mathcal{B})$ è ortogonale e quindi anche la sua inversa $\mathcal{M}(\mathcal{C}, \mathcal{B})$. In particolare

$$\mathcal{M}_{\mathcal{B},\mathcal{B}}(L_A) = \mathcal{M}(\mathcal{B}, \mathcal{C})^T A \mathcal{M}(\mathcal{B}, \mathcal{C})$$

è ortogonale essendo prodotto di matrici ortogonali. Affermiamo che

$$\mathcal{M}_{\mathcal{B},\mathcal{B}}(L_A) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \tilde{a}_{11} & \tilde{a}_{12} \\ 0 & \tilde{a}_{21} & \tilde{a}_{22} \end{bmatrix}.$$

dove $\tilde{A} = \begin{bmatrix} \tilde{a}_{11} & \tilde{a}_{12} \\ \tilde{a}_{21} & \tilde{a}_{22} \end{bmatrix}$ è una matrice ortogonale. Infatti, la prima colonna è il vettore e_1 perché v_1 è autovettore relativo all'autovalore 1. Poiché le colonne di $\mathcal{M}_{\mathcal{B},\mathcal{B}}(L_A)$ formano una base ortonormale, quindi sono a due a due ortogonali, si ha che necessariamente la prima componente del 2 e 3 vettore è nulla. Poiché la seconda e la terza colonna sono vettori ortogonali e di norma unitaria ed il primo coefficiente è zero, le colonne della matrice $\tilde{A}$ formano una base ortonormale di $\mathbb{R}^2$. Quindi $\tilde{A}$ è una matrice ortogonale. Infine, sviluppando il determinante rispetto alla prima colonna si ha $\det \mathcal{M}_{\mathcal{B},\mathcal{B}}(L_A) = \det A = \det \tilde{A} = 1$. Quindi L_A preserva il piano $W = \langle v\rangle^{\perp}$ ed induce su di esso una rotazione poiché $\det \tilde{A} = 1$.

Viceversa, sia $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ una rotazione e sia A tale che $T = L_A$. Poiché T è una rotazione allora se indichiamo con e_1, e_2, e_3 i vettori della base canonica, allora $T(e_1), T(e_2), T(e_3)$ formano una base ortonormale di $\mathbb{R}^3$. Poiché $T(e_i) = A^i$, i.e., la i-esima colonna di A, si ha che le colonne della matrice A formano una base ortonornale ovvero A è ortogonale. Vogliamo dimostrare che A ha determinante uguale a 1. Sia v un autovettore unitario relativo all'autovalore 1 di T. Completiamo v a base ortonormale di $\mathbb{R}^3$ che indichiamo $\mathcal{B} = \{v, w_1, w_2\}$. I vettori w_1, w_2 formano una base del piano normale a v.

La matrice associata a T rispetto alla base $\mathcal{B}$ è

$$\mathcal{M}_{\mathcal{B},\mathcal{B}}(T) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \tilde{a}_{11} & \tilde{a}_{12} \\ 0 & \tilde{a}_{21} & \tilde{a}_{22} \end{bmatrix},$$

dove $\tilde{A}$ è una matrice ortogonale che rappresenta una rotazione nel piano ortogonale a v. Poiché matrici simili hanno lo stesso determinante, ne segue $\det(A) = \det(\mathcal{M}_{\mathcal{B},\mathcal{B}}(T)) = \det(\tilde{A}) = 1$. $\square$

Capitolo 12

Determinante di una matrice

12.1 Applicazioni Multilineari

Sia $\mathbb{K} = \mathbb{R}$ oppure $\mathbb{K} = \mathbb{C}$.

Definizione 12.2.

$$f : \underbrace{\mathbb{K}^n \times \cdots \times \mathbb{K}^n}_{n} \longrightarrow \mathbb{K}$$

si dice una applicazione multilineare *se per ogni* $1 \leq i \leq n$ *e per ogni scelta di vettori* $x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n \in \mathbb{K}^n$*, si ha*

$$\begin{aligned} f(x_1, \ldots, x_{i-1}, Z + W, x_{i+1}, \ldots, x_n) &= f(x_1, \ldots, x_{i-1}, Z, x_{i+1}, \ldots, x_n) \\ &\quad + f(x_1, \ldots, x_{i-1}, W, x_{i+1}, \ldots, x_n), \end{aligned}$$

e

$$f(x_1, \ldots, x_{i-1}, \lambda Z, x_{i+1}, \ldots, x_n) = \lambda f(x_1, \ldots, x_{i-1}, Z, x_{i+1}, \ldots, x_n).$$

per ogni $Z, W \in \mathbb{K}^n$ *e* $\lambda \in \mathbb{K}$*. Diremo che* f *è* una applicazione multilineare alternante *se* f *è multilineare e* $f(x_1, \ldots, x_n) = 0$ *ogni volta che esiste un indice* $1 \leq i \leq n - 1$ *tale che* $x_i = x_{i+1}$*.*

Proposizione 12.3. *Sia* $f : \underbrace{\mathbb{K}^n \times \cdots \times \mathbb{K}^n}_{n} \longrightarrow \mathbb{K}$ *una applicazione multilineare. Allora le seguenti condizioni sono equivalenti:*

a) f *è multineare alternante;*

b) per ogni $1 \leq i \leq n - 1$*, allora*

$$f(x_1, \ldots, x_i, x_{i+1}, \ldots, x_n) = -f(x_1, \ldots, x_{i+1}, x_i, \ldots, x_n);$$

c) $f(v_1, \dots, v_i, \dots, v_j, \dots, v_n) = -f(v_1, \dots, v_j, \dots, v_i, \dots, v_n)$*;*

d) $f(x_1, \dots, x_n) = 0$ *se esistono* $1 \leq i < j \leq n$ *tale che* $x_i = x_j$*;*

e) per ogni $1 \leq j \leq n$ $f(x_1, \dots, x_j + \sum_{m=1, m\neq j}^{n} \lambda_m x_m, x_{j+1}, \dots, x_n) = f(x_1, \dots, x_n)$.

Dimostrazione. $[a] \Rightarrow [b]$. Poiché f è alternante allora

$$f(x_1, \dots, x_i + x_{i+1}, x_i + x_{i+1}, \dots, x_n) = 0.$$

Poiché f è multilineare ed alternante, sviluppando il termine di sinistra si ha:

$$f(x_1, \dots, x_i, x_{i+1}, \dots, x_n) + f(x_1, \dots, x_{i+1}, x_i, \dots, x_n) = 0$$

$[b] \Rightarrow [a]$. Se $x_i = x_{i+1}$, allora

$$f(x_1, \dots, x_i, x_i, \dots, x_n) = -f(x_1, \dots, x_i, x_i, \dots, x_n),$$

ovvero $f(x_1, \dots, x_i, x_i, \dots, x_n) = 0$ da cui segue che f è alternante.

$[b] \Rightarrow [c]$.

$$\begin{aligned} f(v_1, \dots, v_i, \dots, v_j, \dots, v_n) &= (-1)^{j-i} f(v_1, \dots, v_j, v_i, \dots, v_n) \\ &= (-1)^{i-j+(i-j-1)} f(v_1, \dots, v_j, \dots, v_i, \dots, v_n) \\ &= -f(v_1, \dots, v_j, \dots, v_i, \dots, v_n) \end{aligned}$$

$[c] \Rightarrow [d]$. Supponiamo che $x_i = x_j$ con $1 \leq i < j \leq n$. Poiché

$$f(x_1, \dots, x_i, \dots, x_i, \dots, x_n) = -f(x_1, \dots, x_i, x_i, x_{i+1} \dots, x_n)$$

si ha la tesi.

$[d] \Rightarrow [e]$.

$$\begin{aligned} f(x_1, \dots, x_j + \sum_{m=1, m\neq j}^{n} \lambda_m x_m, x_{j+1}, \dots, x_n) &= f(x_1, \dots, x_n) \\ &\quad + \sum_{m=1, m\neq j}^{n} \lambda_j f(x_1, \dots, x_j, \dots, x_n) \\ &= f(x_1, \dots, x_n) \end{aligned}$$

$[e] \Rightarrow [a]$. $f(x_1, \ldots, x_i, x_i + x_{i+1}, \ldots, x_n) = f(x_1, \ldots, x_n)$. Sviluppando il termine di sinistra si ha

$$f(x_1, \ldots, x_i, x_i, \ldots, x_n) + f(x_1, \ldots, x_n) = f(x_1, \ldots, x_n).$$

Quindi $f(x_1, \ldots, x_i, x_i, \ldots, x_n) = 0$. $\square$

Siano $e_1 = \begin{bmatrix} 1 \\ \vdots \\ 0 \end{bmatrix}, \ldots, e_n = \begin{bmatrix} 0 \\ \vdots \\ 1 \end{bmatrix} \in \mathbb{K}^n$.

Teorema 12.4. *Esiste una unica applicazione multilineare alternante*

$$f : \underbrace{\mathbb{K}^n \times \cdots \times \mathbb{K}^n}_{n} \longrightarrow \mathbb{K}$$

tale che $f(e_1, \ldots, e_n) = 1$.

Prima di fornire una dimostrazione, introduciamo il concetto di permutazione.

Definizione 12.5. *Sia* $J_n = \{1, \ldots n\}$. *Una* permutazione *è una applicazione* $\sigma : J_n \longrightarrow J_n$ *biunivoca.*

Ricordiamo che una applicazione $f : X \longrightarrow X$ si dice iniettiva, se $x \neq y$, allora $f(x) \neq f(y)$; suriettiva se l'immagine di f è tutto X, ovvero per ogni $x \in X$ esiste $y \in X$ tale che $f(y) = x$; biunivoca se f è iniettiva e suriettiva.

Poiché J_n è formato da un numero finito di elementi, allora una applicazione $\sigma : J_n \longrightarrow J_n$ è biunivoca se e solamente se σ è iniettiva se e solamente se σ è suriettiva.

Indicheremo con S_n l'insieme di tutte le permutazioni di J_n. Si può dimostrare che S_n è un insieme formato da $n!$ elementi.

Siano $\sigma, \theta \in S_n$. Possiamo comporre due permutazione come usuale $(\sigma \circ \theta)(i) := \sigma(\theta(i))$. La composizione di due permutazioni è ancora una permutazione. Inoltre vale la la proprietà associativa, i.e., $\sigma \circ (\theta \circ \nu) = (\sigma \circ \theta) \circ \nu$ per ogni $\sigma, \theta, \nu \in S_n$. La mappa che manda ogni elemento in se stesso, che indicheremo con id, è una permutazione ed è l'elemento neutro rispetto all'operazione precedente.

Se $\sigma \in S_n$, allora l'applicazione $\tilde{\sigma} \in S_n$, così definita

$$\tilde{\sigma}(j) = k, \text{ se } \sigma(k) = j$$

definisce una permutazione che verifica $\sigma \circ \tilde{\sigma} = \tilde{\sigma} \circ \sigma = id$. La permutazione $\tilde{\sigma}$ è univocamente determinata. In realtà abbiamo dimostrato che S_n ammette una struttura di gruppo rispetto alla composizione.

Una *trasposizione* è una permutazione che scambia due elementi ed lascia fissi tutti gli altri. L'inversa di una trasposizione è ancora una trasposizione.

Sia $f : \underbrace{\mathbb{K}^n \times \cdots \times \mathbb{K}^n}_{n} \longrightarrow \mathbb{K}$ una funzione multilineare alternata che verifica $f(e_1, \ldots, e_n) = 1$. Siano $A^1, \ldots, A^n \in \mathbb{K}^n$ e sia $\sigma \in S_n$. L'insieme $A^1, \ldots, A^n$ differisce da $A^{\sigma(1)}, \ldots, A^{\sigma(n)}$ a meno dell'ordine. Quindi

$$f(A^{\sigma(1)}, \ldots, A^{\sigma(n)}) = \pm f(A^1, \ldots, A^n).$$

Il segno $\pm$ è completamente determintato da σ e non dai vettori $A^1, \ldots, A^n$. Se $f(A_1, \ldots, A^n) \neq 0$, allora definiamo

$$\epsilon(\sigma) = \frac{f(A^{\sigma(1)}, \ldots, A^{\sigma(n)})}{f(A^1, \ldots, A^n)}.$$

Per esempio se σ è una trasposizione, allora $\epsilon(\sigma) = -1$. Invece $\sigma(id) = 1$.

Proposizione 12.6. *Siano $\sigma, \theta \in S_n$. Allora $\epsilon(\sigma \circ \theta) = \epsilon(\sigma)\epsilon(\theta)$. In particolare $\epsilon(\sigma^{-1}) = \epsilon(\sigma)$.*

Dimostrazione. Per definizione

$$f(A^{\sigma(\theta(1))}, \ldots, A^{\sigma(\theta(n))}) = \epsilon(\sigma \circ \theta) f(A^1, \ldots, A^n).$$

Poiché il segno non dipenda dai vettori scelti, allora

$$\begin{aligned} f(A^{\sigma(\theta(1))}, \ldots, A^{\sigma(\theta(n))}) &= \epsilon(\sigma) f(A^{\theta(1)}, \ldots, A^{\theta(n)}) \\ &= \epsilon(\sigma)\epsilon(\theta) f(A^1, \ldots, A^n), \end{aligned}$$

da cui la tesi. □

Dimostrazione Teorema 12.4. Dimostriamo che se f esiste, allora è unica.

Vogliamo calcolare $f(A^1, \ldots, A^n)$. Denotiamo

$$A^j = \begin{bmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{bmatrix},$$

ovvero

$$A^j = a_{1j} e_1 + \cdots + a_{nj} e_n.$$

Poiché f è multilineare, allora si ha

$$f(A^1, \ldots, A^n) = \sum_{j_1,\ldots,j_n=1}^{n} a_{j_1 1} \cdots a_{j_n n} f(e_{j_1}, \ldots, e_{j_n}).$$

Poiché f è alternante, si ha $f(e_{j_1}, \ldots, e_{j_n}) = 0$ ogni volta che abbiamo due indici unguali. Se $\{j_1, \ldots, j_n\} = \{1, \ldots, n\} = J_n$, allora la mappa

$$\sigma : J_n \longrightarrow J_n \qquad k \mapsto j_k,$$

è una permutazione, da cui segue che

$$f(e_{j_1}, \ldots, e_{j_n}) = f(e_{\sigma(1)}, \ldots, e_{\sigma(n)}) = \epsilon(\sigma) f(e_1, \ldots, e_n) = \epsilon(\sigma).$$

Allora

$$f(A^1, \ldots, A^n) = \sum_{\sigma \in S_n} \epsilon(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(n)1}.$$

ovvero esiste un unica f multilineare alternata che verifica $f(e_1, \ldots, e_n) = 1$. Viceversa, utilizzando la notazione precedente, definiamo

$$f(A^1, \ldots, A^n) = \sum_{\sigma \in S_n} \epsilon(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(n)n}.$$

Se $B = b_{1i}e_1 + \cdots + b_{ni}e_n$, allora

$$\begin{aligned} f(A^1, \ldots, A^i + B, \ldots, A^n) &= \sum_{\sigma \in S_n} \epsilon(\sigma) a_{\sigma(1)1} \cdots (a_{\sigma(i)i} + b_{\sigma(i)i}) \cdots a_{\sigma(n)n} \\ &= \sum_{\sigma \in S_n} \epsilon(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(i)i} \cdots a_{\sigma(n)n} \\ &+ \sum_{\sigma \in S_n} \epsilon(\sigma) a_{\sigma(1)1} \cdots b_{\sigma(i)i} \cdots a_{\sigma(n)n} \\ &= f(A^1, \ldots, A^i, \ldots, A^n) + f(A^1, \ldots, B, \ldots, A^n). \end{aligned}$$

Inoltre

$$\begin{aligned} f(A^1, \ldots, \lambda A^i, \ldots, A^n) &= \sum_{\sigma \in S_n} \epsilon(\sigma) a_{\sigma(1)1} \cdots (\lambda a_{\sigma(i)i}) \cdots a_{\sigma(n)n} \\ &= \lambda \sum_{\sigma \in S_n} \epsilon(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(i)i} \cdots a_{\sigma(n)n} \\ &= \lambda f(A^1, \ldots, A^n). \end{aligned}$$

Per dimostrare che f è alternante, basta dimostrare che vale una delle condizioni equivalenti della Proposizione 12.3. Supponiamo che $A^i = A^{i+1}$. Poiché

$$A^j = a_{1j}e_1 + \cdots + a_{nj}e_n,$$

allora $a_{1i} = a_{1(i+1)}, \ldots, a_{ni} = a_{n(i+1)}$ e quindi

$$f(A^1, \ldots, A^n) = \sum_{\sigma \in S_n} \epsilon(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(i)i} a_{\sigma(i+1)(i+1)} \cdots a_{\sigma(n)n}.$$

Data σ, consideriamo τ_σ la trasposizione che scambia $\sigma(i)$, $\sigma(i+1)$. Quindi $\epsilon(\sigma) = -\epsilon(\tau_\sigma \circ \sigma)$ ed

$$\begin{aligned}&\epsilon(\tau_\sigma \circ \sigma) a_{\tau_\sigma(\sigma(1))1} \cdots a_{\tau_\sigma(\sigma(i))i} a_{\tau_\sigma(\sigma(i+1))(i+1)} \cdots a_{\tau_\sigma(\sigma(n))n} \\ &= -\epsilon(\sigma) a_{\tau_\sigma(\sigma(1))1} \cdots a_{\sigma(i)} a_{i,\sigma(i+1))(i+1)} \cdots a_{\tau_\sigma(\sigma(n))n}.\end{aligned}$$

Quindi

$$\begin{aligned}&\epsilon(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(i)i} a_{\sigma(i+1)(i+1)} \cdots a_{\sigma(n)n} \\ &+ \epsilon(\tau_\sigma \circ \sigma) a_{\tau_\sigma(\sigma(1))1} \cdots a_{\tau(\sigma(i))i} a_{\tau_\sigma(\sigma(i+1))(i+1)} \cdots a_{\tau_\sigma(\sigma(n))n} = 0\end{aligned}$$

Questo significa che i termini si cancellano due a due da cui segue che $f(A^1, \ldots, A^n) = 0$ se esiste i tale che $A^i = A^{i+1}$. $\square$

Sia $A \in M_{n \times n}(\mathbb{K})$. Se $A = (A^1, \ldots, A^n)$ ed $A^j = \begin{bmatrix} a_{1j} \\ \vdots \\ a_{jn} \end{bmatrix}$, definiamo

$$\mathrm{Det}(A) = \sum_{\sigma \in S_n} \epsilon(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(n)n}.$$

Proposizione 12.7. $\mathrm{Det}(A) = \mathrm{Det}(A^T)$

Dimostrazione. Se $A = (a_{ij})$, allora $A^T = (b_{ij})$ dove $b_{ij} = a_{ji}$. Inoltre se $k = \sigma(i)$, allora $\sigma^{-1}(k) = i$. Quindi

$$\begin{aligned}
\mathrm{Det}(A) &= \sum_{\sigma \in S_n} \epsilon(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(n)n} \\
&= \sum_{\sigma \in S_n} \epsilon(\sigma) a_{\sigma(1)\sigma^{-1}(\sigma(1))} \cdots a_{\sigma(n)\sigma^{-1}(\sigma(n))} \\
&= \sum_{\sigma \in S_n} \epsilon(\sigma) a_{1\sigma^{-1}(1)} \cdots a_{n\sigma^{-1}(n)} \\
&= \sum_{\sigma \in S_n} \epsilon(\sigma^{-1}) a_{1\sigma^{-1}(1)} \cdots a_{n\sigma^{-1}(n)} \\
&= \sum_{\sigma \in S_n} \epsilon(\sigma^{-1}) b_{\sigma^{-1}(1)1} \cdots b_{\sigma^{-1}(n)n} \\
&= \sum_{\sigma \in S_n} \epsilon(\sigma) b_{\sigma(1)1} \cdots b_{\sigma(n)n} \\
&= \mathrm{Det}(A^T).
\end{aligned}$$

□

Teorema 12.8 (Formula di Binet)**.** $\mathrm{Det}(AB) = \mathrm{Det}(A)\mathrm{Det}(B)$.

Dimostrazione. Se $C = AB$, allora $C^i = AB^i$ ed

$$C^i = b_{1i}A^1 + \cdots + b_{ni}A^n.$$

Quindi

$$\begin{aligned}
\mathrm{Det}(C) &= \mathrm{Det}(AB^1, \ldots, AB^n) \\
&= \sum_{j_1,\ldots,j_n=1}^{n} b_{j_1 1} \cdots b_{j_n 1} \mathrm{Det}(A^{j_1}, \ldots, A^{j_n}) \\
&= \sum_{\sigma \in S_n} b_{\sigma(1)1} \cdots b_{\sigma(n)n} \mathrm{Det}(A^{\sigma(1)}, \ldots, A^{\sigma(n)}) \\
&= \sum_{\sigma \in S_n} \epsilon(\sigma) b_{\sigma(1)1} \cdots b_{\sigma(n)n} \mathrm{Det}(A^1, \ldots, A^n) \\
&= \mathrm{Det}(B)\mathrm{Det}(A)
\end{aligned}$$

□

Sia $1 \leq k \leq n$. Vogliamo dimostrare che è possibile calcolare il determinante attraverso lo sviluppo di Laplace rispetto ad una qualsiasi riga oppure colonna.

Sia $1 \leq k \leq n$. Se $A \in M_{n\times n}(\mathbb{K})$. Se $A \in M_{1\times 1}(\mathbb{K})$, allora $A = (a)$ e definiamo $f(A) = a$. Supponiamo di aver definito f per $n-1$ e definiamo

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

dove A_{kj} è la matrice di formato $(n-1)\times(n-1)$ che ottengo a partire da A eliminando la k-esima riga e la j-esima colonna. La formula anteriore è lo *sviluppo di Laplace rispetto alla k-esima riga.* Vogliamo dimostrare che $f(A) = \det(A)$. Quindi è sufficiente dimostrare che la funzione f è multineare ed alternante sulle colonne di A e che verifica $f(e_1, \ldots, e_n) = 1$. La dimostrazione verrà fatta per induzione su n.

Se $n = 1$ allora $f(A) = \det(A)$. Supponiamo vero per $n-1$. Dimostriamolo per n. Sia

$$A = (A^1, \ldots, \lambda A^i + \mu B^i, \ldots, A^n).$$

Allora

$$\begin{aligned} f(A) &= f(A^1, \ldots, \lambda A^i + \mu B^i, \ldots, A^n) \\ &= \sum_{j=1}^{i-1} (-1)^{j+k} a_{kj} f(A_{kj}) \\ &+ \sum_{j=i+1}^{n} (-1)^{j+k} a_{kj} f(A_{kj}) \\ &+ (\lambda a_{ki} + \mu b_{ki}) f(A_{ki}) \end{aligned}$$

Adesso se $j \neq i$, allora

$$A_{kj} = (A^1, \ldots, \lambda A^i + \mu B^i, \ldots, A^n)_{kj} = \lambda (A^1, \ldots, A^i, \ldots, A^n)_{kj} + \mu (A^1, \ldots, B^i, \ldots, A^n)_{kj}$$

Se $i = j$, allora

$$f(A_{ki}) = f((A^1, \ldots, A^i, \ldots, A^n)_{ki}) = f((A^1, \ldots, B^i, \ldots, A^n)_{ki})$$

Quindi

$$\begin{aligned}
f(A) &= f(A^1, \dots, \lambda A^i + \mu B^i, \dots, A^n) \\
&= \sum_{j=1}^{i-1} (-1)^{j+k} a_{kj} f(A_{kj}) + \sum_{j=i+1}^{n} (-1)^{j+k} a_{kj} f(A_{kj}) \\
&+ (\lambda a_{ki} + \mu b_{ki}) f(A_{ki}) \\
&= \sum_{j=1}^{i-1} (-1)^{j+k} \big(\lambda a_{kj} f((A^1, \dots, A^i, \dots, A^n)_{kj}) + \mu a_{kj} f((A^1, \dots, B^i, \dots, A^n)_{kj})\big) \\
&+ \sum_{j=i+1}^{n} (-1)^{j+k} \big(\lambda a_{kj} f((A^1, \dots, A^i, \dots, A^n)_{kj}) + \mu a_{kj} f((A^1, \dots, B^i, \dots, A^n)_{kj})\big) \\
&+ \lambda a_{ki} f((A^1, \dots, A^i, \dots, A^n)_{ki}) + \mu b_{ki} f((A^1, \dots, B^i, \dots, A^n)_{ki}) \\
&= \lambda f(A^1, \dots, A^i, \dots, A^n) + \mu f(A^1, \dots, B^i, \dots, A^n).
\end{aligned}$$

Dimostriamo che è alternante. Supponiamo che $A^i = A^{i+1}$ per un certo $1 \leq i \leq n-1$.

$$\begin{aligned}
f(A^1, \dots, A^i, A^i, \dots, A^n) &= \sum_{j=1}^{i-1} (-1)^{j+k} a_{kj} f(A_{kj}) \\
&+ \sum_{j=i+2}^{n} (-1)^{j+k} a_{kj} f(A_{kj}) \\
&+ (-1)^{k+i} a_{ki} f(A_{ki}) + (-1)^{k+i+1} a_{ki} f(A_{k(i+1)}).
\end{aligned}$$

La matrice A_{kj} ha due colonne uguali per $k \neq i, i+1$. Per ipotesi induttiva $f(A_{kj}) = 0$. Inoltre $A_{ki} = A_{k(i+1)}$ e quindi gli ultimi due termini si cancellano.

Infine dimostriamo che $f(Id_n) = 1$. È facile verificare che

$$f(Id_n) = \sum_{j=1}^{n} a_{jk} f((Id_n)_{kj}) = f(Id_{n-1}) = \cdots = f(Id_1) = 1.$$

Applicando Il Teorema 12.4 si ha che

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

Corollario 12.9. *Sia* $A \in M_{n\times n}(\mathbb{K})$ *e sia* $1 \leq k \leq n$. *Allora*

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

Dimostrazione. Applicando la Proposizione 12.7, si ha

$$\det(A) = \det(A^T) = \sum_{j=1}^{n} (-1)^{j+k} a_{kj} \det(A^t_{jk}).$$

Poiché $A^T_{jk} = A_{kj}$, allora

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

□