Capitolo 1

Spazio euclideo

1.1 $\mathbb{R}^3$: 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 ovvero le proiezioni lungo gli assi x, y e z (Figura 1.1.1) Se indichiamo con $\mathbb{R}^3$ l'insieme ordinato di terne di numeri reali, i.e,

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

allora 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. Quindi possiamo identificare lo spazio euclideo con $\mathbb{R}^3$.

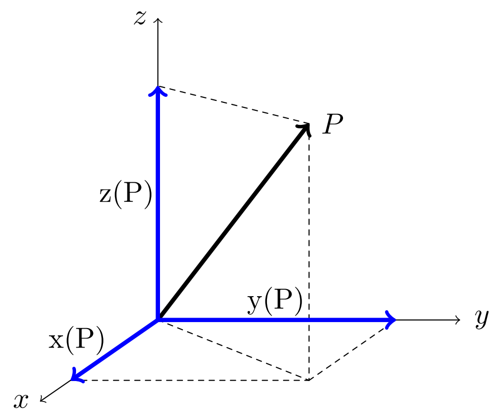


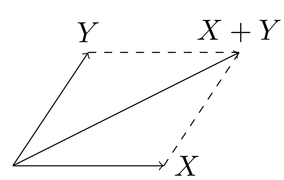
Figura 1.1.1.

Se $X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$, $Y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$, allora $X = Y$ se $x_1 = y_1,\ x_2 = y_2,\ x_3 = y_3$.
Possiamo definire due operazioni su $\mathbb{R}^3$:

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 si interpreta geometricamente con la regola del parallelogramma, ovvero se $X, Y \in \mathbb{R}^3$, allora La moltiplicazione per scalare si può pensare



come nella seguente figura:

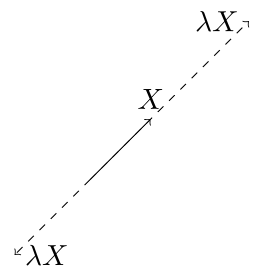


Figura 1.1.2.

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

Proposizione 1.2. *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 $0 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$, *allora* $0 + X = X + 0 = X$*;*

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

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

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

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

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

Definizione 1.3. *Siano* $X_1, \ldots, X_k \in \mathbb{R}^3$

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

- $X_1, \ldots X_k$ *si dicono linearmente dipendenti se esistono* $\lambda_1, \ldots, \lambda_k$ *non tutti nulli, tali che*

$$\lambda_1 X_1 + \cdots + \lambda_k X_k = 0;$$

- $X_1, \dots X_k$ *si dicono linearmente indipendenti se non sono linearmente dipendenti, ovvero se, comunque scelti* $\lambda_1, \dots, \lambda_k$ *non tutti nulli, si ha*

$$\lambda_1 X_1 + \cdots + \lambda_k X_k \neq 0;$$

ovvero

$$\lambda_1 X_1 + \cdots + \lambda_k X_k = 0;$$

se e solamente se $\lambda_1 = \cdots = \lambda_k = 0$.

Esempio 1.4. *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 esistono* $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}$ *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 1.5. *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.

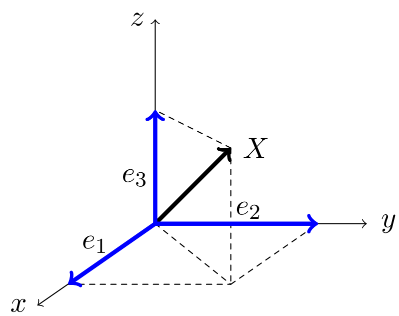


Figura 1.5.1.

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

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = x_1e_1 + x_2e_2 + x_3e_3.$$

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 1.6. *Siano* $\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 5 \end{bmatrix}$. *Stabilire se i vettori sono linearmente dipendenti, rispettivamente indipendenti, equivale a studiare le combinazioni lineari*

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

uguale al vettore nullo. 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, rispettivamente dipendenti, equivale a stabilire se il seguente sistema lineare*

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

ammette come unica soluzione $\alpha_1 = \alpha_2 = \alpha_3 = 0$*, rispettivamente ammette soluzioni* $\alpha_1, \alpha_2, \alpha_3$ *non tutti nulli.*

Proposizione 1.7. *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 $X - \lambda Y = 0$, ovvero 0 è combinazione lineare non banale dei vettori X e Y. Analogamente se $Y = \lambda X$. Viceversa, supponiamo che X e Y sono linearmente dipendenti. Allora esistono α e β non entrambi nulli, tali che

$$\alpha X + \beta Y = 0.$$

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

Diremo che due vettori X e Y sono proporzionali se esiste $\lambda \in \mathbb{R}$ tali $X = \lambda Y$ oppure $Y = \lambda X$.

Osservazione 1.8.

- $X \in \mathbb{R}^3$ *è linearmente indipendente se e solamente se* $X \neq 0$*;*
- $0, X_1, \ldots, X_k$ *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.*

1.9 $\mathbb{R}^3$: struttura metrica

In questa sezione investigheremo la struttura metrica dello spazio euclideo.

Definizione 1.10. Il prodotto scalare canonico *è una funzione che associa ad ogni coppia di vettori un numero reale come segue:*

$$\mathbb{R}^3 \times \mathbb{R}^3 \longrightarrow \mathbb{R} \qquad (X, Y) \mapsto \langle X, Y \rangle = x_1 y_1 + x_2 y_2 + x_3 y_3,$$

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

(a) $\langle X, X \rangle = x_1^2 + x_2^2 + x_3^2 \geq 0$ *e* $\langle X, X \rangle = 0$ *se e solamente se* $X = 0$*;*

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

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

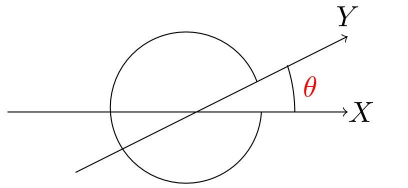
(d) $\langle \lambda X, Y \rangle = \lambda \langle X, Y \rangle$*;*

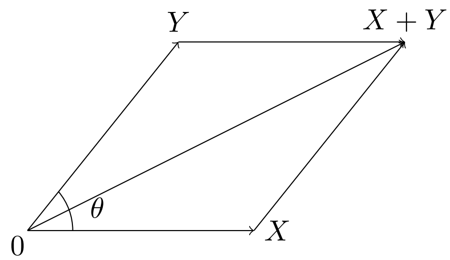
(e) $\langle X, Y + Z \rangle = \langle X, Y \rangle + \langle X, Z \rangle$*;*

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

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. Vogliamo dimostrare che il prodotto scalare $\langle X, Y \rangle$ descrive, in maniera algebrica, il $\cos \theta$, dove θ è l'angolo fra X e Y. Calcoliamo la norma del vettore $X + Y$:





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

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

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

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

Vediamo alcune conseguenze immediate della formula anteriore.

Corollario 1.12. *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)*

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

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

La discussione anteriore suggerisce la seguente definizione.

Definizione 1.13. *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

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

Proposizione 1.14. *Siano* $v, w \in S^\perp$ *e* $\lambda\mathbb{R}$*. Allora* $v + w \in S^\perp$ *e* $\lambda v \in S^\perp$*.*

Dimostrazione. Poiché $v, w \in S^\perp$, si ha $\langle v, s \rangle = \langle w, s \rangle = 0$ per ogni $s \in S$. Noi dobbiamo dimostrare che

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

e

$$\langle \lambda v, s \rangle = 0$$

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

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

Analogamente

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

□

Teorema 1.15. *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 $\langle X,Y\rangle = 0$ se e solamente se X, Y sono vettori perpendicolari. □

1.16 $\mathbb{R}^3$: 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

$$X \times Y = \begin{bmatrix} x_2y_3 - x_3y_2 \\ -y_3x_1 + x_3y_1 \\ x_1y_2 - y_1x_2 \end{bmatrix}.$$

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

Proposizione 1.17. *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 la verifiche delle 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

$$\begin{cases} x_2y_3 - x_3y_2 = 0\\ -y_3x_1 + x_3y_1 = 0\\ x_1y_2 - y_1x_2 = 0\end{cases}$$

Se $X = 0$ allora ho finito. Supponiamo quindi $X \neq 0$. Se $x_1 \neq 0$, allora

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

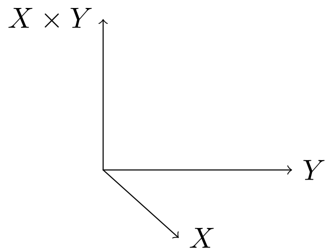
ovvero $Y = \frac{y_1}{x_1}X$. Analogamente gli altri casi.

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

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

$\square$

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



descrive il prodotto vettoriale.

1.18 Rette e piani

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 1.18.1).

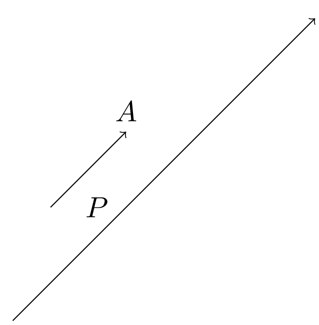


Figura 1.18.1.

Come insieme

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

Una retta ha molte possibili equazioni parametriche. Possiamo scegliere un qualsiasi punto $Q \in r$ ed come vettore direttore $B = kA$ con $k \neq 0$. Infatti se $Q = P + t_o A$, allora

$$P + tA = Q - t_o A + tA = Q + (t - t_o)A = Q + \frac{(t - t_0)}{k} B,$$

da cui segue

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

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 $A \in \mathbb{R}^3$ 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

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

Proviamo che effettivamente 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 $P_2 = Q + t_2 A$. Poiché $P_1 \neq P_2$, ne segue che $t_1 \neq t_2$. Quindi $P_1 - P_2 = (t_1 - t_2)A$ e

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

Siano $r_1 = P_1 + tA_1$ e $r_2 = P_2 + tA_2$ due rette nello spazio. Diremo che le due rette sono *ortogonali* o *perpendicolari* se $\langle A_1, A_2 \rangle = 0$. Le mutue posizione di due rette nello spazio sono le seguenti:

Le rette r_1 e r_2 sono *parallele* se non hanno punti in comune ed hanno la stessa direzione.

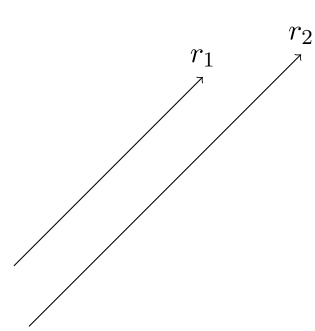


Figura 1.18.2.

Quindi i vettori direttori delle rette r_1 e r_2 sono proporzionali, ovvero A_1 e A_2 sono linearmente dipendenti. Quindi se due rette sono parallele allora $A_1 \times A_2 = 0$. Se r_1 ed r_2, oltre ad essere parallele, hanno anche un punto in comune, allora esse sono *coincidenti*. Quindi, due rette $r_1 = P_1 + tA_1$ ed $r_2 = P_2 + tA_2$ sono *coincidenti* se e solo se $P_1 \in r_2$ e r_1 ed r_2 sono parallele.

Le rette r_1 e r_2 sono *incidenti* se hanno esattamente un punto in comune.

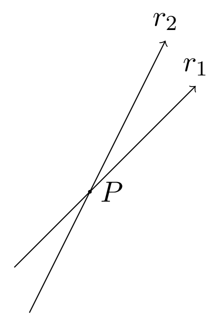


Figura 1.18.3.

Quindi r_1 e r_2 sono incidenti se e solamente se $r_1 \cap r_2 = \{P\}$ e $A_1 \times A_2 \neq 0$. Infine, due rette si dicono *sghembe* se non sono ne incidenti ne parallele, ovvero r_1 e r_2 sono sghembe se e solamente se non hanno punti in comune ed i vettori direttori sono linearmente indipendenti.

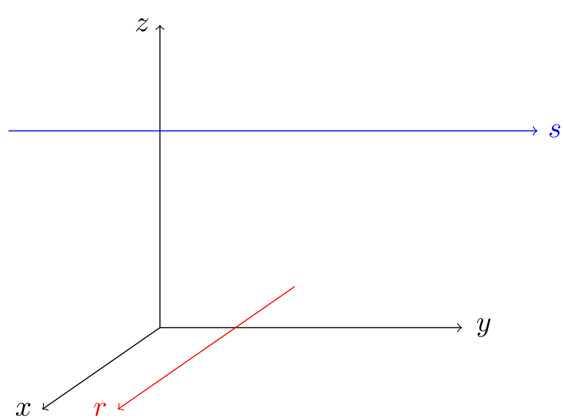
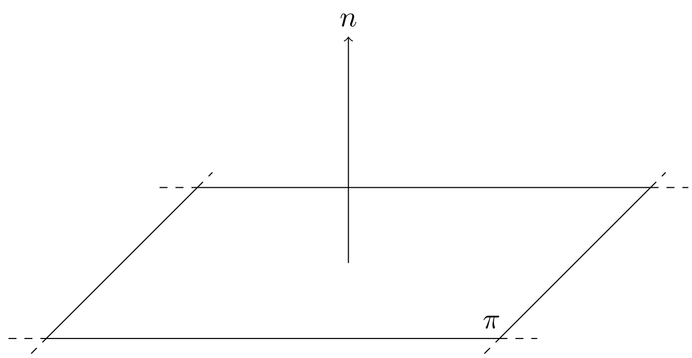


Figura 1.18.4.

Possiamo pensare ad un piano come all'insieme ortogonale ad una direzione fissata. Sia $n = \begin{bmatrix} a \\ b \\ c \end{bmatrix} \neq 0$. Il piano passante per l'origine ed ortogonale alla direzione n è

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

Il vettore n si chiama *vettore normale al piano.* Il piano passante per il



punto P e ortogonale alla direzione n 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 $d = -\langle P, n\rangle$. Quindi un'equazione cartesiane di un piano nello spazio è

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

dove $a, b, c \in \mathbb{R}$ non tutti nulli.

Il piano π passa per l'origine se e solamente se $d = 0$. Osserviamo che il piano π con vettore normale n e passante per P non è altro che il piano passante l'origine, ortogonale a n e traslato lungo il vettore P.

Osservazione 1.19.

a) Esistono infiniti piani passanti per due punti;

b) se $P_1, P_2 \in \pi$, *allora* $r = \{P_1 + t(P_2 - P_1), t \in \mathbb{R}\} \subset \pi$;

c) Esiste un unico piano passante per tre punti non allineati. Ricordiamo che P_1, P_2 *e* P_3 *vettori di* $\mathbb{R}^3$ *si dicono allineati se* $P_2 - P_1$ *e* $P_3 - P_1$ *sono linearmente dipendenti; cioé se e solamente se appartengono ad una stessa retta. Se non sono allineati possiamo trovare* $n = (P_2 - P_1) \times (P_3 - P_1)$ *e poi imponiamo il passaggio per* P_1, *ovvero il piano* π *ha equazioni*

$$\pi := \left\{X \in \mathbb{R}^3 : \langle X - P_1, n\rangle = 0\right\}.$$

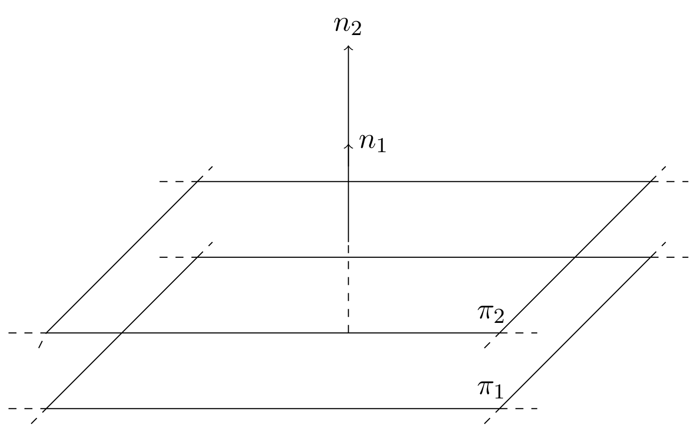
Siano $\pi_1 : a_1x+b_1y+c_1z+d_1 = 0$ e $\pi_2 : a_2x+b_2y+c_2z+d_2 = 0$ due piani. Indichiamo con $n_1 = \begin{bmatrix} a_1 \\ b_1 \\ c_1 \end{bmatrix}$ e $n_2 = \begin{bmatrix} a_2 \\ b_2 \\ c_2 \end{bmatrix}$ i loro vettori normali. Diremo che π_1 e π_2 sono *ortogonali* se n_1 e n_2 sono ortogonali, ovvero $\langle n_1, n_2 \rangle = 0$. Adesso studiamo la mutua posizione di due piani nello spazio.

Diremo che i piani π_1 e π_2 sono:

- *coincidenti* se $\pi_1 = \pi_2$, e quindi se e solamente se esiste $k \in \mathbb{R}$ non nullo tale che

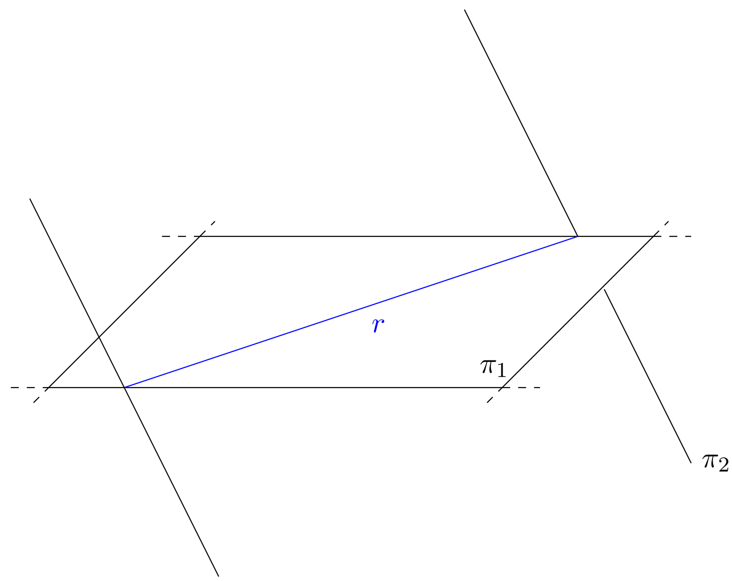
$$n_1 = kn_2 \text{ (linearmente dipendenti) e } d_1 = kd_2;$$

- *paralleli* se π_1 e π_2 non si intersecano. In questo caso si ha $n_1 \times n_2 = 0$, ovvero n_1 e n_2 sono linearmente dipendenti;



- *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 $n_1 \times n_2 \neq 0$.

Figura 1.19.1.



Calcolare le equazioni parametriche di un piano equivale a risolvere il sistema lineare che definisce il piano.

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.

- Se $a \neq 0$, allora $x = -\frac{d}{a} - \frac{b}{a}y - \frac{c}{a}z$, da cui segue

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

  Quindi π ha equazioni paramentriche: $\pi : P + yv + zw,\ y, z \in \mathbb{R}$, dove $P = \begin{bmatrix} -d/a \\ 0 \\ 0 \end{bmatrix} \in \pi$, $v = \begin{bmatrix} -b/a \\ 1 \\ 0 \end{bmatrix}$, $w = \begin{bmatrix} -c/a \\ 0 \\ 1 \end{bmatrix}$ sono vettori linearmenti indipendenti e ortogonali a n;

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

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

;

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

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

Esempio 1.20. *Determinare un'equazione parametrica per il piano* $\pi : 2x - y + z = 4$. *Poiché* $y = 2x + z - 4$ *si ha*

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

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

Abbiamo dimostrato che ogni piano π ha equazioni paramentriche $\pi : P + tv + sw$, $t, s \in \mathbb{R}$, dove v, w sono vettori linearmente indipendenti ed ortogonali a n e $P \in \pi$. Se $P_1, P_2, P_3 \in \mathbb{R}^3$ sono sono tre punti non allineati, ovvero $(P_2 - P_1) \times (P_3 - P_1) \neq 0$, allora l'unico piano passante per P_1, P_2, P_3 ha equzioni parametriche

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

Sia π un piano e sia

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

con v e w linearmente indipendenti, un'equazione parametrica di π. Allora π ha equazioni cartesiane:

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

dove $n = \begin{bmatrix} a \\ b \\ c \end{bmatrix} = v \times w$ e $d = -\langle P, n \rangle$.

Esempio 1.21. *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 1.22. *Determinare un'equazione parametrica e cartesiane del piano* π *contenente la retta*

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

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

Poiché $Q = \begin{bmatrix} 1 \\ 2 \\ -2 \end{bmatrix} \in r$, *allora* $Q \in \pi$. *Quindi i vettori* $P, Q \in \pi$, *da cui segue che il vettore* $P - Q = \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}$ *è ortogonale al vettore normale al piano. La stessa considerazione si applica al vettore direttore della retta* 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

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

Abbiamo detto che due piani non paralleli si intersecano lungo una retta. Vogliamo provare il viceversa, ovvero che ogni retta è intersezione di due piani non paralleli.

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 fosse differente da zero, allora potremmo ricavare t

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

e sostituendo nelle altre due, 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

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

ed 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, intersecandosi, individuano la stessa retta.

Esempio 1.23.

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

Posso ricavare $t = x - 1$ *e quindi*

$$r = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \in \mathbb{R}^3 : y - 2x = 0,\ z - 3x = -3 \right\}.$$

Se la retta r ha equazioni cartesiane

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

allora determinare le equazioni parametriche della retta r equivale a risolvere un sistema di 2 equazioni in 3 incognite. Poiché un vettore direttore della retta r è ortogonale ai vettori normali dei due piani, può essere calcolato facendo 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 1.24.

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

Dalla seconda equazione ricaviamo $y = 2 + 2z$ *e poi* $x = 3 + z$*. Quindi*

$$r := \begin{cases} x = 3 + z \\ y = 2 + 2z \\ z = z \end{cases} = \begin{bmatrix} 3 \\ 2 \\ 0 \end{bmatrix} + z \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix},\ z \in \mathbb{R}$$

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 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 esercizi.

Esempio 1.25. *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 1.26. *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.

Un piano contenente s_1, applichiamo il metodo del fascio, ha equazioni:

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

Imponendo il passaggio per P si ha:

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

ovvero $\alpha = -2\beta$. Se $\beta = -1$, allora $\alpha = 2$ il piano $\pi_1 : y + 2z = 0$ contiene la retta r. 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 1.27. *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 e 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 il cui vettore normale è ortogonale ad A contiene 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*

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

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

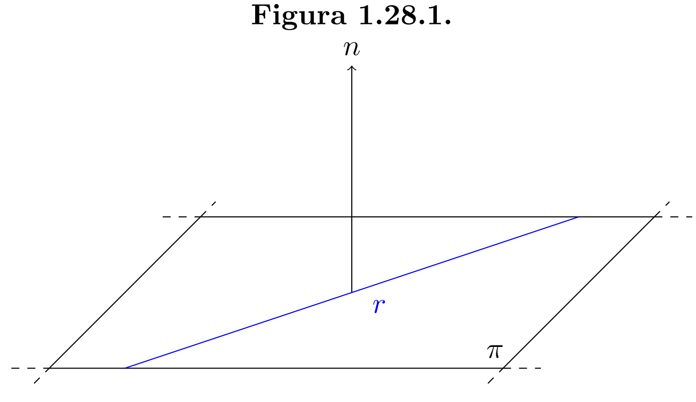
Proposizione 1.28. *Due rette sono complanari se e solamente se sono parallele oppure incidenti.*

Infine 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:

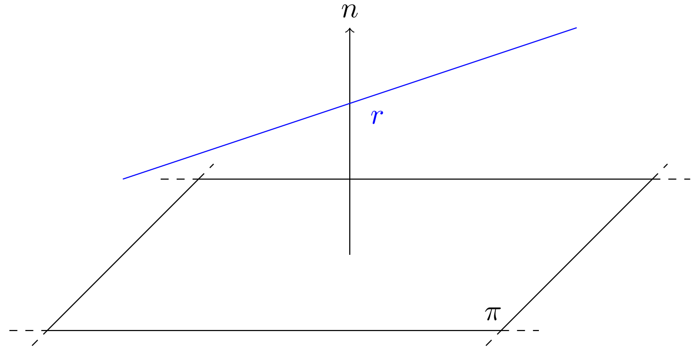
- $r \subset \pi$ se e solamente se $P \in \pi$ e $\langle A, n\rangle = 0$;

Figura 1.28.1.



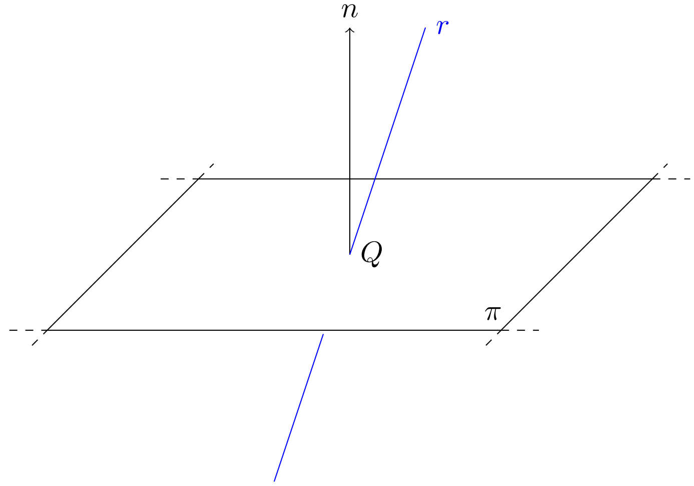
- la retta r è parallela a π e questo succede se e solamente se $r \cap \pi = \emptyset$ ed $\langle A, n \rangle = 0$;

Figura 1.28.2.



- infine $r \cap \pi = \{Q\}$, ovvero r e π sono *incidenti*. Quindi r ed π sono incidenti se e solamente se $\langle A, n \rangle \neq 0$;

Figura 1.28.3.



Esempio 1.29. *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.*

Poiché la retta passa per P ed è parallela al piano $x + y - z = 6$, la retta cercata 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

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

Capitolo 2

Matrici

2.1 $\mathbb{R}^n$, $\mathbb{C}^n$: Struttura lineare

Sia $\mathbb{K} = \mathbb{R}$ oppure $\mathbb{C}$. Il simbolo $\mathbb{K}^n$ è l'insieme delle colonne ordinate di n numeri, ovvero

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

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 solamente se $x_1 = y_1, \ldots,\ x_n = y_n$. Dati $X, Y \in \mathbb{K}^n$ e $\lambda \in \mathbb{K}$ si pone:

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

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

ovvero somma e moltiplicazione per scalare sono definite componente per componente. La somma e moltiplicazioni per scalari godono di importanti proprietà.

Proposizione 2.2. *Dati* $X, Y, Z \in \mathbb{K}^n$, $\lambda, \mu \in \mathbb{K}$ *allora* *a)* $X + (Y + Z) = (X + Y) + Z$*: proprierà associativa della somma;*

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

c) posto $0 = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$, *allora* $0 + X = X + 0 = X$*;*

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

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

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

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

Notazione: porremo $X - Y := X + (-1)Y$.

Definizione 2.3.

- *Dati* $X_1, \ldots X_k \in \mathbb{K}^n$ *si dice combinazione lineare di* $X_1, \ldots, X_k$ *ogni elemento* $Z \in \mathbb{K}^n$ *per cui esistano* $\lambda_1, \ldots, \lambda_k \in \mathbb{K}$ *tali che* $Z = \lambda_1 X_1 + \cdots + \lambda_k X_k$*. I numeri* $\lambda_1, \ldots, \lambda_k$ *si chiamano i coefficienti della combinazione lineare.*

- $X_1, \ldots X_k \in \mathbb{K}^n$ *si dicono linearmente dipendenti se esistono* $\lambda_1, \ldots, \lambda_k$ *non tutti nulli, tali che*

$$\lambda_1 X_1 + \cdots + \lambda_k X_k = 0;$$

$X_1, \ldots X_k \in \mathbb{K}^n$ *si dicono linermente indipendenti se non sono linearmente dipendenti, ovvero se, comunque scelti* $\lambda_1, \ldots, \lambda_k$ *non tutti nulli, si ha*

$$\lambda_1 X_1 + \cdots + \lambda_k X_k \neq 0;$$

ovvero

$$\lambda_1 X_1 + \cdots + \lambda_k X_k = 0;$$

se e solamente se $\lambda_1 = \cdots = \lambda_k = 0$.

Esempio 2.4. *Siano* $e_1 = \begin{bmatrix} 1 \\ \vdots \\ 0 \end{bmatrix}, \cdots, e_n = \begin{bmatrix} 0 \\ \vdots \\ 1 \end{bmatrix}$. *È facile verificare che* $(e_1, \ldots, e_n)$ *sono linearmente indipendenti e che ogni vettore* $X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n$ *è combinazione lineare di* $e_1, \ldots, e_n$. *Infatti*

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

Esempio 2.5. *Siano* $\begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 0 \\ 5 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -1 \\ 3 \end{bmatrix} \in \mathbb{R}^4$ *Stabilire se* $\begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix}$ *è combinazione lineare dei rimanenti, significa studiare l'esistenza* $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{R}$ *tale*

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

ovvero

$$\begin{cases} \alpha_1 + 2\alpha_2 = 1 \\ \alpha_1 + \alpha_2 + \alpha_3 = 1 \\ \alpha_1 - \alpha_3 = 0 \\ \alpha_1 + 5\alpha_2 + 53\alpha_3 = 1. \end{cases}$$

Quindi stabilire se $\begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix}$ *è combinazione lineare dei vettori* $\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 0 \\ 5 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -1 \\ 3 \end{bmatrix}$ *è equivalente a stabilire se il sistema lineare anteriore ammette soluzione, combinazione lineare, oppure no, ovvero non è combinazione lineare.*

Osservazione 2.6.

- $X \in \mathbb{K}^n$ *è linearmente indipendente se e solamente se* $X \neq 0$.
- *Due vettori* X *e* Y *di* $\mathbb{K}^n$ *sono linearmente dipendenti se e solamente se sono proporzionali, ovvero* $X = \lambda Y$ *oppure* $Y = \lambda X$ *per un opportuno* $\lambda \in \mathbb{K}$*;*
- $0, X_1, \ldots, X_k$ *sono linearmente dipendenti;*
- *Se* $\mathcal{B}$ *è un insieme di elementi di* $\mathbb{K}^n$ *linearmente indipendenti, allora ogni sottoinsieme di* $\mathcal{B}$ *è costituito da elementi linearmente indipendenti;*
- *se* $\mathcal{B}$ *è un insieme costituito di elementi di* $\mathbb{K}^n$ *linearmente dipendenti, allora ogni sovrainsieme di* $\mathcal{B}$ *è costituito da elementi linearmente dipendenti.*

2.7 Matrici

Definizione 2.8. *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* $a_{ij} \in \mathbb{C}$.

Due matrici A e B sono uguali se hanno lo stesso formato e vale $a_{ij} = b_{ij}$ per ogni $1 \leq i \leq m$, e $1 \leq j \leq n$. L'insieme delle matrici reali di formato $m \times n$, rispettivamente complesse, sarà indicato con $M_{m \times n}(\mathbb{R})$, rispettivamente $M_{m \times n}(\mathbb{C})$. Si osservi inoltre che $M_{m \times 1}(\mathbb{R}) = \mathbb{R}^m$ e $M_{m \times 1}(\mathbb{C}) = \mathbb{C}^m$ rispettivamente. Sia $A \in M_{m \times n}(\mathbb{K})$ dove $\mathbb{K} = \mathbb{R}$ oppure $\mathbb{K} = \mathbb{C}$. 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 $A \in \mathrm{M}_{m\times n}(\mathbb{K})$ come formata da n colonne $A = (A^1, \dots, A^n)$ oppure come formata da m righe $A = \begin{bmatrix} A_1 \\ \vdots \\ A_m \end{bmatrix}$. Se $n = m$, allora la matrice A si dirà *matrice quatrata* di ordine n.

2.8.1 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 $B = (b_{ij})_{\substack{1 \le i \le m \\ 1 \le j \le n}}$, allora

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

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

Proposizione 2.9. *Dati $A, B, C \in M_{m\times n}(\mathbb{K})$, $\lambda, \mu \in \mathbb{K}$ allora*

a) $A + (B + C) = (A + B) + C$*: proprierà associativa della somma;*

b) $A + B = B + A$*: proprietà commutativa della somma;*

c) posto $0 = \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{bmatrix}$*, allora* $0 + A = A + 0 = A$*;*

d) se $A \in \mathrm{M}_{m\times n}(\mathbb{K})$*, allora* $A+(-1)A = 0$*;* $-A := (-1)A$ *è detto opposto di* A*;*

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

f) $\lambda(A + B) = \lambda A + \lambda B$*;*

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

Notazione: $A - B := A + (-1)B$.

Sia $A \in M_{m\times n}(\mathbb{K})$. La trasposta di A è una matrice $A^T \in M_{n\times m}(\mathbb{K})$ così definita:

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

ovvero A^T si ottiene a partire da A scambiando le righe con le colonne; rispettivamente le colonne con le righe.

Proposizione 2.10. *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.

Proposizione 2.11. *Siano $A, B \in M_{m\times n}(\mathbb{C})$ e $\lambda \in \mathbb{C}$. Allora*

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

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

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

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: ovvero $A = (a_{ij})_{\substack{1\le i\le n\\ 1\le j\le n}}$ e $a_{ij} = 0$ se $i \neq j$.

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

La *matrice identità di ordine* n, Id_n, è la matrice diagonale che ha tutti gli elementi uguali ad 1 sulla diagonale principale.

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

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 $A = (a_{ij})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}}$, è triangolare superiore, i.e.,

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

se $a_{ij} = 0$ quando $i > j$, rispettivamente triangolare inferiore, i.e.,

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

se $a_{ij} = 0$ quando $i < j$. L'insieme della matrici diagonali, rispettivamente triangolari superiori e triangolari inferiori, è chiuso rispetto alla somma e moltiplicazione per scalare $M_{n\times n}(\mathbb{K})$. Ovvero se A, B sono matrici diagonali, rispettivamente triangolari superiori e triangolari inferiori e $\lambda \in \mathbb{K}$, allora $A+B$ e λA sono ancora matrici diagonali, rispettivamente triangolari superiori e triangolari inferiori.

Sia $A \in M_{n\times n}(\mathbb{R})$. Diremo che la matrice A è: *simmetrica* se $A = A^T$; *antisimmetrica* se $A = -A^T$. L'insieme delle matrici simmetriche, rispettivamente antisimmetriche, è chiuso rispetto alla somma e moltiplicazione per scalare.

Se A è una matrice quadrata, alllora A e la sua trasposta A^T hanno la stessa diagonale principale. Quindi una matrice antisimmetrica ha tutti gli elementi nulli sulla diagonale principale. È facile verificare che tale condizione è necessaria ma non sufficiente.

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 $A = -A^*$. La somma di matrici Hermitiane, rispettivamente anti-Hermitiane, è ancora una matrice Hermitiana, rispettivamente anti-Hermitiana. 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 2.12. *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 2.13. *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 $\mathrm{Tr}(A^*) = \overline{\mathrm{Tr}(A)}$.

2.13.1 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 2.14.

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

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.

Esempio 2.15. *Siano* $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, $C = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$. *Allora:*

- $CA = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = A$;
- $AC = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$;
- $AA = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$;
- $CC = C$;

Quindi il prodotto non è commutativo. Inoltre è possibile che $AC = 0$ benché A e C non siano la matrice nulla.

Se $A = (a_{ij})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}}$. Se $p \in \mathbb{N}$, allora definiamo $A^p = \underbrace{A \cdots A}_{p}$ se $p > 0$.

Il prodotto di matrici gode delle seguenti proprietà.

Proposizione 2.16. *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 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 $(AB)^T$. Osserviamo che in generale la matrice $A^T B^T$ non è nemmeno definita. Dimostriamo che coincidono ovvero hanno gli stessi elementi. Sia $C = (AB)^T$. Allora

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

dove $(AB)_{ji}$, rispettivamente $(B^T A^T)_{ij}$ indica l'elemento della matrice AB che si trova sulla j riga e i colonna, rispettivamente la i riga e la j esima colonna. Quindi $(AB)^T = B^T A^T$. $\square$

Osservazione 2.17. *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*

$$(AB)_k = A_k B.$$

Infatti

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

Osservazione 2.18. *Siano* $e_1 = \begin{bmatrix} 1 \\ \vdots \\ 0 \end{bmatrix}, \dots, e_n = \begin{bmatrix} 0 \\ \vdots \\ 1 \end{bmatrix} \in \mathbb{K}^n$ *e sia* $A \in \mathrm{M}_{m\times n}(\mathbb{K})$. *Allora*

$$Ae_i = \begin{bmatrix} a_{1i} \\ \vdots \\ a_{mi} \end{bmatrix} = A^i$$

per $= 1, \dots, n$.

Siano $e_1 = \begin{bmatrix} 1 \\ \vdots \\ 0 \end{bmatrix}, \dots, e_m = \begin{bmatrix} 0 \\ \vdots \\ 1 \end{bmatrix} \in \mathbb{K}^m$. *Allora*

$$e_i^T A = A_i,$$

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

Proposizione 2.19. *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.

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

□

2.20 Matrici invertibili

Definizione 2.21. *Una matrice* A *quadrata di ordine* n *si dice invertibile se esiste un matrice* B *quadrata di ordine* n *tale che* $AB = BA = \mathrm{Id}_n$.

Proposizione 2.22.

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 $AB = \mathrm{Id}_n$, allora

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

assurdo.

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

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

Allora

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

Le rimanenti proprietà sono lasciate per esercizio. □

Osservazione 2.23. *Si può dimostrare che se A, B sono matrici quadrate tale che $AB = \mathrm{Id}_n$, allora A è invertibile e B è l'inversa di A.*

La proposizione anteriore afferma che il prodotto di matrici invertibili è invertibile. Quindi l'insieme $GL(n, \mathbb{R}) = \{A \in M_{n\times n}(\mathbb{R}) :\ A$ è invertibile $\}$ rispettivamente $GL(n, \mathbb{C}) = \{A \in M_{n\times n}(\mathbb{R}) :\ A$ è invertibile$\}$ è un *gruppo non commutativo*, detto gruppo lineare generale.

Se A è una matrice invertibile, e $n \in \mathbb{Z}$ negativo, definiamo

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

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

Definizione 2.24. *Una matrice $A \in M_{n\times n}(\mathbb{R})$ si dice* ortogonale *se $A^T = A^{-1}$,ovvero $AA^T = A^TA = \mathrm{Id}_n$*

Definizione 2.25. *Una matrice $A \in M_{n\times n}(\mathbb{C})$ si dice* unitaria *se $AA^* = A^*A = Id_n$.*

Il prodotto di matrici ortogonali, rispettivamente unitarie, è ancora una matrice ortogonale, rispettivamente una matrice unitaria. 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.

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

Esempio 2.27.

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

3
$$\det(A^1, \dots, A^i + B^i, \dots, A^n) = \det(A^1, \dots, A^i, \dots A^n) + \det(A^1, \dots, B^i, \dots, A^n),$$

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 2.28.

- *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 2.29. *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 2.30. *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 $B = A^{-1}$. $\square$

Corollario 2.31. *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

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

poiché $\det(A) = \det(A^T)$, da cui segue la tesi. $\square$

Corollario 2.32. *Se* A *è una matrice unitaria, allora* $|\det(A)| = 1$.

Dimostrazione.

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

$\square$

2.33 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. 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 2.34. *$A \in M_{n\times n}(\mathbb{K})$ 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 2.35 (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, possiamo effettuare delle operazioni di righe, rispettivamente colonne, senza alterare il rango di una matrice.

Definizione 2.36. *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 2.37. *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. Quindi il rango per minore non cambia. $\square$

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

Teorema 2.38. *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, dare la seguente definizione.

Definizione 2.39. *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*

2.40 Riduzione a scala di Gauss

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 $s_{1j_1}, \ldots, s_{r,j_r}$ 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 2.41.

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

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

Proposizione 2.42. *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 $rg(S) \leq r$. Se consideriamo il minore di ordine r formato dalle colonne $S^{j_1}, \dots, S^{j_r}$ e dalle righe $S_1, \dots, S_r$, ovvero $S^{j_1,\dots,j_r}_{1,\dots,r}$ ha la forma:

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

Quindi il $r \leq rg(S) \leq r$ da cui segue che $rg(S) = r$.

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

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

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

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

Poiché $s_{rj_r}, \dots, s_{1j_1}$ sono tutti non nulli si ha che $\det A \neq 0$, 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 $S^{j_1}, \dots, S^{j_r}$ sono linearmente indipendenti. □

Teorema 2.43. *Ogni matrice A può essere ridotta in forma a scala mediante operazioni elementari di righe.*

Dimostrazione. Sia $A = (A^1, \dots, A^n)$.
Passo 1
sia $1 \le j_1 \le n$ il più piccolo intero affinché $A^{j_1} \neq 0$.
Passo 2
Se $a_{1j_1} \neq 0$ bene. Altrimenti scambio due righe in modo che $a_{1j_1} \neq 0$. Quindi la matrice A, dopo le precedenti operazioni elementari, 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 mio obbiettivo è quello di arrivare ad una matrice i cui elementi sulla colonna A^{j_1} sono tutti nulli tranne il primo. Quindi, se alla k-riga A_k, $k \ge 2$, sottraggo $-a_{kj_1}/a_{1,j_1}A_1$, ottengo, attraverso operazioni elementari di righe, 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 $B = 0$ 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 2.44. *Sia A una matrice ed S una sua riduzione a scala. Allora $rg(A) = rg(S)$, ovvero è uguale al numero di elementi di testa di una sua riduzione a scala; è 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 $\mathrm{rg}(A) = \mathrm{rg}(S)$. $\square$

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

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

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:*

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

Dimostrazione. Supponiamo che esistano $0, 0'$ tale che $x + 0 = 0 + x = x + 0' = 0' + x = x$ per ogni $x \in V$. Allora

$$0 = 0 + 0' = 0'.$$

Sia $\underline{0} \in \mathbb{K}$ e sia $v \in V$. Allora

$$\underline{0}v = (\underline{0} + \underline{0})v = \underline{0}v + \underline{0}v.$$

Poiché $\underline{0}v$ ammette un inverso rispetto alla somma si ha

$$\underline{0}v = 0.$$

Poiché $v + (-1)v = (1 - 1)v = \underline{0}v = 0$, $(-1)v$ è un inverso rispetto alla somma. Vediamo l'unicità. Siano $v', v'' \in V$ tale che $v' + v = v'' + v = 0$. Allora

$$v' = v' + (v + v'') = (v' + v) + v'' = 0 + v'' = v''.$$

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

$$\lambda 0 = \lambda(0 + 0) = \lambda 0 + \lambda 0,$$

ovvero $\lambda 0 = 0$. Infine sia $v \neq 0$. Sia $\lambda \neq 0$ e supponiamo che $\lambda v = 0$. Moltiplicando per λ^{-1} a destra e sinistra si ha $v = 0$. Assurdo. Quindi $\lambda v = 0$ se e solamente se $\lambda = 0$. $\square$

In seguito, denoteremo $v - w := v + (-1)w$. Vediamo alcuni esempi di spazi vettoriali.

Sia

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

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 di $\mathbb{R}^n$ è il vettore $\begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$. In maniera analoga, 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 $x_1 = y_1, \ldots,\ x_n = y_n$. Definiamo:

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

b) $\lambda X = \begin{bmatrix} \lambda x_1 \\ \vdots \\ \lambda x_n \end{bmatrix}$, ovvero somma e moltiplicazione per scalare.

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

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

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

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. Per definire la somma $v + w$, osserviamo che se $m > n$, allora possiamo scrivere $v = a_0 + a_1 x + \cdots + a_n x^n + 0x^{n+1} + \cdots + 0x^m$. Analogamente se $m < n$. Quindi possiamo supporre che $n = m$ e definire

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

L'insieme $\mathbb{K}[x]$ con le operazioni definite è uno spazio vettoriale su $\mathbb{K}$. Il vettore nullo è 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 $V = \{f : V \longrightarrow \mathbb{K}\}$ ovvero l'insieme di tutte le applicazioni a valori in un 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 V lo $0 \in \mathbb{K}$, ovvera l'applicazione che vale costantemente $0 \in \mathbb{K}$.

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

- $\forall v, w \in W$, *allora* $v + w \in W$ *(chiuso rispetto alla somma);*
- $\lambda \in \mathbb{K}$ *e* $w \in W$, *allora* $\lambda w \in W$ *(chiuso rispetto alla moltiplicazione per scalare).*

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

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. La prossima proposizione è una condizione equivalente affinché un sottoinsieme è un sottospazio vettoriale ed è lasciata per esercizio.

Proposizione 3.8. *Sia $W \subset V$. 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 $S \subset \mathbb{R}^3$. Allora $S^\perp$ è un sottospazio vettoriale di $\mathbb{R}^3$;

b) $W = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2 : x + y = 0 \right\}$ è un sottospazio vettoriale di $\mathbb{R}^2$;

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 $V = \mathbb{K}[x]$ e sia $p = a_0 + a_1 x + \cdots + a_n x^n \in \mathbb{K}[x]$. Possiamo pensare a p come una funzione

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

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

Sia $\beta \in \mathbb{K}$. 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 $v_1, \dots, v_s \in V$. Si dice combinazione lineare di $v_1, \dots, v_s \in V$ ogni elemento $w \in V$ esprimibile nella forma*

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

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

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

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

Esempio 3.11. *Sia $V = \mathbb{R}^3$. Un retta passante per l'origine ha equazione parametrica*

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

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

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

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

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

$$v + w = (\alpha_1 + \beta_1) v_1 + \dots + (\alpha_s + \beta_s) v_s \in \mathcal{L}(v_1, \dots, v_s)$$

e

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

□

Dalla proposizione anteriore si ha che le rette e piani per l'origine sono sottospazi vettoriale di $\mathbb{R}^3$. Non è difficile convincersi che i sottospazi vettoriali di $\mathbb{R}^3$ sono: $\{0\}$, rette passanti per l'origine, piani passanti per l'origine e $\mathbb{R}^3$ stesso.

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

Dimostrazione. Siano $\alpha_1, \dots, \alpha_k \in \mathbb{K}$. Poiché W è un sottospazio vettoriale, si ha che $\alpha_1 w_1, \cdots, \alpha_k w_k \in W$ e quindi anche la somma, ovvero

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

Quindi ogni combinazione lineare di $w_1, \dots, w_k$ appartiene a W per cui

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

□

Osservazione 3.14. *Sia $w, w_1, \dots, w_k \in V$. Dire che w è combinazione lineare di $w_1, \dots, w_k$ è equivalente alla condizione $w \in \mathcal{L}(w_1, \dots, w_k)$.*

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

a) $v_1, \dots, v_s$ *sono* linearmente dipendenti *se esistono $\alpha_1, \dots, \alpha_s$ non tutti nulli tali che*

$$\alpha_1 v_1 + \cdots + \alpha_s v_s = 0;$$

b) $v_1, \dots, v_s$ linearmente indipendenti *se non sono linearmente dipendenti, ovvero se*

$$\alpha_1 v_1 + \cdots + \alpha_s v_s = 0,$$

allora necessariamente $\alpha_1 = \cdots = \alpha_s = 0$;

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

Esempio 3.16. *Sia $V = \mathbb{R}^4$ e siano* $\begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -3 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ -1 \\ 1 \end{bmatrix} \in \mathbb{R}^4$. *I vettori sono linearmente dipendenti se esistono $\alpha_1, \dots, \alpha_4 \in \mathbb{R}$ non tutti nulli tali che*

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

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

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

Quindi se il sistema anteriore ammette solzioni "non banali", allora i vettori sono linearmente dipendenti. Altrimenti i vettori sono linearmente indipendenti.

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

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

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

Esempio 3.18. *Sia $1, x, \dots, x^n \in \mathbb{K}_n[x]$. Allora $\mathcal{L}(1, x, \dots, x^n) = \mathbb{K}_n[x]$. Inoltre $1, x, \dots, x^n$ sono linearmente indipendenti.*

Osservazione 3.19. *Sia V uno spazio vettoriale su $\mathbb{K}$. Allora*

- *$v \in V$ è linearmente indipendente se e solamente se $v \neq 0$;*
- *$v, w \in V$ sono linearmenti dipendenti se e solamente se $v = \lambda w$ oppure $w = \lambda v$ per un certo $\lambda \in \mathbb{K}$;*
- *sia $\mathcal{B} = \{v_1, \dots, v_n\}$. Allora:*
  - *se $v_1, \dots v_n$ 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 e sistemi lineari

4.1 Sistemi lineari

Studiando i sistemi lineari, i problemi principali che vogliamo risolvere sono: quando un sistema ammette soluzioni; se lo 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:

$$\left\{\begin{array}{ccc} a_{11}x_1 + \cdots + a_{1n}x_n & = & b_1 \\ \vdots & & \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n & = & b_m \end{array}\right. \tag{4.1}$$

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 termnini noti sono uguali a zero, il sistema si dice *omogeneo*.

Definizione 4.2. *Una soluzione del sistema (4.1) 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

$$AX = b, \tag{4.2}$$

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 $Sol(A|b) = \{X \in \mathbb{K}^n : \ AX = b\}$ 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. Poiché

$$AX = x_1A^1 + \cdots + x_nA^n,$$

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

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

Poiché $\mathcal{L}(A^1, \ldots, A^n, b) \supseteq \mathcal{L}(A^1, \ldots, A^n)$ è sufficiente dimostrare che l'uguaglianza $\mathcal{L}(A^1, \ldots, A^n, b) \subset \mathcal{L}(A^1, \ldots, A^n)$ è equivalente la fatto che il sistema lineare $AX = b$ è compatibile. Infatti, se $\mathcal{L}(A^1, \ldots, A^n, b) \subset \mathcal{L}(A^1, \ldots, A^n)$, allora $b \in \mathcal{L}(A^1, \ldots, A^n)$ e quindi il sistema è compatibile. Viceversa, se $AX = b$ è compatibile, allora $b \in \mathcal{L}(A^1, \ldots, A^n)$. Poiché $b, A_1, \ldots, A_n \in \mathcal{L}(A^1, \ldots, A^n)$, tenendo in mente che $\mathcal{L}(A^1, \ldots, A^n)$ è un sottospazio vettoriale di $\mathbb{K}^n$, si ha

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

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

Un sistema lineare omogeneo è sempre compatibile o risolubile poiché ammette come soluzione il vettore nullo $0 \in \mathbb{K}^n$. Inoltre vale il seguente risultato.

Proposizione 4.3. *Sia $AX = 0$ un sistema lineare omogeneo. Allora $Sol(A|0) = \{X \in \mathbb{K}^n : AX = 0\}$ è un sottospazio vettoriale di $\mathbb{K}^n$.*

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

$$A(\lambda X + \mu Y) = \lambda AX + \mu AY = 0.$$

$\square$

Osservazione 4.4. *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 $b = 0$.*

Sia $AX = b$ un sistema lineare. Diremo che il sistema omogeneo $AX = 0$ è il sistema lineare omogeneo associato a $AX = b$.

Teorema 4.5 (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 $AX = 0$; Quindi*

$$Sol(A|b) = \{X_o + W,\ W \in Sol(A|0)\}.$$

Dimostrazione. Indichiamo con $E = \{X_o + W,\ X \in Sol(A|0)\}$. Vogliamo dimostrare che $E = Sol(A|b)$. Sia $Y \in Sol(A|b)$. Allora $A(Y - X_o) = AX - AX_o = b - b = 0$, ovvero $Y - X_o$ è soluzione del sistema lineare omogeneo associato, da cui segue che $Y - X_o \in Sol(A|0)$, cioé $Y = X_o + W$ per un certo $W \in Sol(A|0)$. Quindi $Sol(A|b) \subset E$. Viceversa, ogni elemento di E è della forma $X_o + W$, dove W è una soluzione del sistema lineare omogeneo associato. Allora

$$A(X_o + W) = AX_o + AW = b + 0 = b,$$

da cui segue che $E \subset Sol(A|b)$. Quindi $Sol(A|b) = \{X_o + X,\ X \in Sol(A|0)\}$. $\square$

Definizione 4.6. *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 4.7. *Sia $AX = b$ un sistema lineare. Se scambio due righe alla matrice $(A|b)$ ottengo una matrice $(A'|b')$ ed il sistema $A'X = b'$ è equivalente al sistema $AX = b$.*

Definizione 4.8. *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 4.9. *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 delle 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 del sistema lineare $AX = d$, ovvero

$$(A|d) = \left[\begin{array}{ccc|c} a_1 & \cdots & a_n & a \\ b_1 & \cdots & b_n & b \\ \vdots & \vdots & \vdots & \vdots \\ \cdots & \cdots & \cdots & \cdots \end{array}\right]$$

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 $\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ soddisfa sicuramente tutte le equazioni di $AX = d$ tranne la seconda. 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 $b_1x_1 + \cdots + b_nx_n = b$. $\square$

Come conseguenza otteniamo il seguente risultato.

Corollario 4.10. *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 sulle riche della matrice completa.*

4.10.1 Sistemi ridotti a scala

Definizione 4.11. *Un sistema $SX = c$ si dice ridotto a scala se la matrice S è ridotta a scala.*

Esempio 4.12.

$$\begin{cases} x_1 + x_2 - x_3 = 2 \\ x_2 - x_4 = 0 \\ x_3 + x_4 = 8 \end{cases}$$

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 4.13. *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 4.14. *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 4.15 (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 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. $\square$

Osservazione 4.16. $\mathrm{rg}(A) \leq \mathrm{rg}(A|b)$. *Inoltre*

$$\mathrm{rg}(A|b) = \begin{cases} \mathrm{rg}(A) \\ \text{oppure} \\ \mathrm{rg}(A) + 1 \end{cases}$$

Corollario 4.17. *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. $\square$

Corollario 4.18 (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*

$$x_i = \frac{\det(A^1, \dots, A^{i-1}, b, A^{i+1}, \dots, A^n)}{\det(A)}$$

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

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

$$x_1 A^1 + \ldots + x_n A^n = b,$$

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. □

4.18.1 Metodi di Calcolo

Siano $X_1, \dots, X_s \in \mathbb{K}^n$. 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

$$\mathrm{rg}(A) = \mathrm{rg}(A|Z).$$

In particolare, i vettori $X_1, \dots, X_s$ sono:

- linearmente dipendenti se e solamente se il sistema $AX = 0$ ammette soluzioni non banali, se e solamente se $rg(A) < s$;
- linearmente indipendenti se e solamente se il sistema $AX = 0$ ammette una ed una sola soluzione, se e solamente se $rg(A) = s$;

Se $s = n$, allora $X_1, \dots X_n$ sono linearmente indipendenti se e solamente se $\det(A) \neq 0$. Infine, se $s > n$, allora $X_1, \dots X_s$ sono linearmente dipendenti.

4.18.2 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 si intersecanose 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 4.19. *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 4.20. *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, ovvero 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 4.21. *$r \subset \pi$ 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. $\square$

Corollario 4.22. *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 $\det(A) \neq 0$. $\square$

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 $\mathbb{K}^n$;

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 $\{1, t, \ldots, t^n\}$ formano una base di $\mathbb{K}_n[t]$;

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 $\mathbb{R}^3$;

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 $\mathrm{M}_{2\times 2}(\mathbb{R})$.

Proposizione 5.3. *Sia V uno spazio vettoriale su $\mathbb{K}$ e sia $\mathcal{B} = \{v_1, \ldots, v_n\}$ una base di V. Allora ogni vettore $v \in V$ si può scrivere in maniera unica come combinazione lineare dei vettori $v_1, \ldots v_n$. Viceversa, se ogni elemento si può scrive in maniera unica come combinazione lineare di $v_1, \ldots v_n$, allora $\mathcal{B} = \{v_1, \ldots, v_n\}$ è una base di V.*

Dimostrazione. Sia $v \in V$. Poiché $\mathcal{B} = \{v_1, \ldots, v_n\}$ è un sistema di generatori, allora esistono $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$ 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 $\mathcal{B} = \{v_1, \ldots, v_n\}$ un insieme di vettori linearmente indipendenti, ne segue che

$$\alpha_1 = \lambda_1, \ldots, \alpha_k = \lambda_k.$$

Viceversa, se ogni elemento si scrive come combinazione lineare di $v_1, \ldots v_n$, allora

$$\mathcal{L}(v_1, \ldots, v_n) = V,$$

ovvero $v_1, \ldots, v_n$ sono un sistema di generatori. Dall'unicità, se

$$\alpha_1 v_1 + \cdots + \alpha_n v_n = 0,$$

allora necessariamente i coefficienti $\alpha_1 = \cdots = \alpha_n = 0$, poiché il vettore nullo si può scrivere come combinazione lineare dei vettori $v_1, \ldots, v_n$ con tutti i coefficienti nulli. Quindi i vettori $v_1, \ldots v_n$ sono linearmente indipendenti. $\square$

Definizione 5.4. *Sia $v \in V$ e sia $\mathcal{B} = \{v_1, \ldots, v_n\}$ una base di V. Le coordinate di v rispetto a $\mathcal{B}$ sono gli unici scalari $x_1, \ldots, x_n \in \mathbb{K}$ 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 $\mathcal{B}$.*

Sia $\mathcal{B} = \{v_1, \ldots, v_n\}$ 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. Inoltre

- $[0_V]_{\mathcal{B}} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$;
- $[v + w]_{\mathcal{B}} = [v]_{\mathcal{B}} + [w]_{\mathcal{B}}$;
- $[\lambda v]_{\mathcal{B}} = \lambda [v]_{\mathcal{B}}$.

Dimostriamo la seconda proprietà lasciando per esercizio le altre.

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

$$[v + w]_{\mathcal{B}} = \begin{bmatrix} (x_1 + y_1) \\ \vdots \\ (x_n + y_n) \end{bmatrix} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = [v]_{\mathcal{B}} + [w]_{\mathcal{B}}.$$

In particolare l'applicazione $\mathcal{F}_{\mathcal{B}}$ trasforma combinazioni lineari in combinazioni lineari. Infatti, siano $w_1, \ldots, w_k \in V$ e sia $\alpha_1 w_1 + \cdots + \alpha_k w_k$ una combinazione lineare dei vettori $w_1, \ldots, w_k$. Quindi

$$
\begin{aligned}
\mathcal{F}_{\mathcal{B}}(\alpha_1 w_1 + \cdots + \alpha_k w_k) &= F_{\mathcal{B}}(\alpha_1 w_1) + \cdots + F_{\mathcal{B}}(\alpha_k w_k) \\
&= \alpha_1 F_{\mathcal{B}}(w_1) + \cdots + \alpha_k F_{\mathcal{B}}(w_k) \\
&= \alpha_1 [w_1]_{\mathcal{B}} + \cdots + \alpha_k [w_k]_{\mathcal{B}}.
\end{aligned}
$$

Proposizione 5.5. *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 indipendenti, se e solamente se $[w_1]_{\mathcal{B}}, \ldots, [w_k]_{\mathcal{B}} \in \mathbb{K}^n$ sono linearmente dipendenti;

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) $\mathcal{F}_{\mathcal{B}}\left(\mathcal{L}(w_1, \ldots, w_k)\right) = \mathcal{L}([w_1]_{\mathcal{B}}, \ldots, [w_k]_{\mathcal{B}})$.

Dimostrazione. Siano $w_1, \ldots, w_k$ linearmente dipendenti. Allora esistono $\alpha_1, \ldots, \alpha_k \in \mathbb{K}$ non tutti nulli tali che $\alpha_1 w_1 + \cdots + \alpha_k w_k = 0$. Quindi

$$
\begin{aligned}
0_{\mathbb{K}^n} &= \mathcal{F}_{\mathcal{B}}(0_V) \\
&= \mathcal{F}_{\mathcal{B}}(\alpha_1 w_1 + \cdots + \alpha_k w_k) \\
&= \alpha_1 [w_1]_{\mathcal{B}} + \cdots + \alpha_k [w_k]_{\mathcal{B}},
\end{aligned}
$$

ovvero i vettori $[w_1]_{\mathcal{B}}, \ldots, [w_k]_{\mathcal{B}}$ sono linearmente dipendenti. Viceversa, se i vettori $[w_1]_{\mathcal{B}}, \ldots, [w_k]_{\mathcal{B}}$ sono linearmente dipendenti, allora esistono $\alpha_1, \ldots, \alpha_k \in \mathbb{K}$ 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 $w_1, \ldots, w_k \in V$ 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 di $w_1, \dots, w_k \in V$ se e solamente 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, la condizione anteriore è equivalente alla condizione

$$\mathcal{F}_{\mathcal{B}}(v) = \mathcal{F}_{\mathcal{B}}(\beta_1 w_1 + \cdots + \beta_k w_k) = \beta_1 [w_1]_+ \cdots + \beta_k [w_k],$$

ovvero $v \in V$ è combinazione lineare dei vettori $w_1, \dots, w_k$ se e solamente se $[v]_{\mathcal{B}}$ è combinazione lineare dei vettori $[w_1]_{\mathcal{B}}, \dots, [w_k]_{\mathcal{B}}$.

L'ultima affermazione è lasciata per esercizio. □

5.5.1 Dimensione di uno spazio vettoriale

L'obiettivo di questa sezione è quello di dimostrare che uno spazio vettoriale generato da un numero finito di vettori ammette una base e che due basi hanno lo stesso numero di elementi. Cominciamo con il seguente Lemma.

Lemma 5.6. *Siano $v_1, \dots, v_n$ vettori di V. I vettori $v_1, \dots v_n$ 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*

$$v_j \in \mathcal{L}(v_1, \dots, v_{j-1}, v_{j+1}, \dots, v_n).$$

Inoltre

$$\mathcal{L}(v_1, \dots, v_{j-1}, v_{j+1}, \dots, v_n) = \mathcal{L}(v_1, \dots, v_n).$$

Dimostrazione. Sia $\lambda_1 v_1 + \cdots + \lambda_k v_k = 0$ con $\lambda_1, \dots, \lambda_k$ non tutti nulli. Supponiamo che $\lambda_j \neq 0$. Allora

$$v_j = -\lambda_j^{-1}(\lambda_1 v_1 + \cdots + \lambda_{j-1} v_{j-1} + \lambda_{j+1} v_{j+1} + \cdots + \lambda_n v_n).$$

Viceversa se

$$v_j = \lambda_1 v_1 + \cdots + \lambda_{j-1} v_{j-1} + \lambda_{j+1} v_{j+1} + \cdots + \lambda_k v_n,$$

Allora

$$\lambda_1 v_1 + \cdots + \lambda_{j-1} v_{j-1} - v_j + \lambda_{j+1} v_{j+1} + \cdots + \lambda_k v_k = 0.$$

Adesso, proviamo che $\mathcal{L}(v_1, \dots, v_{j-1}, v_{j+1}, \dots, v_n) = \mathcal{L}(v_1, \dots, v_n)$.

Poiché $\mathcal{L}(v_1, \dots, v_n)$ è un sottospazio vettoriale di V e i vettori

$$v_1, \dots, v_{j-1}, v_{j+1}, \dots, v_n \in \mathcal{L}(v_1, \dots, v_n),$$

si ha $\mathcal{L}(v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_n) \subset \mathcal{L}(v_1, \ldots, v_n)$.

Viceversa, poiché v_j è combinazione lineare di $v_1, \ldots, v_{j-1}, v_{j+1}, \ldots v_n$, si ha

$$v_j = \lambda_1 v_1 + \cdots + \lambda_{j-1} v_{j-1} + \lambda_{j+1} v_{j+1} + \cdots + \lambda_k v_k$$

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 $\mathcal{L}(v_1, \ldots, v_n) \subset \mathcal{L}(v_1, \ldots, v_{j-1}, v_{j+1}, \ldots, v_n)$ concludendo la dimostrazione. $\square$

Proposizione 5.7. *Sia V uno spazio vettoriale che ammette un numero finito di generatori. Allora V ammette una base.*

Dimostrazione. Siano $v_1, \ldots, v_n$ un sistema di generatori. Quindi $V = \mathcal{L}(v_1, \ldots, v_n)$. 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 5.6, esiste $1 \leq j_1 \leq n$ tale che v_{j_1} è combinazione lineare degli altri ed

$$V = \mathcal{L}(v_1, \ldots, v_n) = \mathcal{L}(v_1, \ldots, v_{j_1-1}, v_{j_1+1}, \ldots, v_n),$$

ovvero $v_1, \ldots, v_{j_1-1}, v_{j_1+1}, \ldots, v_n$ formano un sistema di generatori. Adesso, se i vettori $v_1, \ldots, v_{j_1-1}, v_{j_1+1}, \ldots, v_n$ non fossero linearmente indipendenti 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 5.8 (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 $\alpha_1 w_1 + \cdots + \alpha_m w_m = 0$. Poiché $V = \mathcal{L}(v_1, \ldots, v_n)$, ogni vettore w_j, per $j = 1, \ldots, m$, si scrive come combinazione lineare di $v_1, \ldots, v_n$:

$$w_j = \sum_{k=1}^{n} a_{kj} v_k.$$

Quindi

$$0 = \sum_{j=1}^{m} \alpha_j w_j = \sum_{j=1}^{m} \sum_{k=1}^{n} \alpha_j a_{kj} v_k = \sum_{k=1}^{n} \left(\sum_{j=1}^{m} a_{kj} \alpha_j \right) v_k.$$

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 4.17 (del Teorema di Rouché-Capelli) il sistema ammette soluzioni non banali. Quindi i vettori $\mathcal{C} = \{w_1, \dots, w_m\}$ sono linearmente dipendenti. □

Corollario 5.9. *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 $m \leq n$. □

Corollario 5.10. *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 $n = m$. □

Definizione 5.11. *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 5.12. *Sia V uno spazio vettoriale su $\mathbb{K}$ di dimensione n. Allora:*

- *n è il massimo numero di vettori di V linearmente indipendenti;*
- *se $v_1, \dots, v_m$ formano un sistema di generatori di V, allora $n \leq m$.*

Esempio 5.13.

a) i vettori $e_1 = \begin{bmatrix} 1 \\ \vdots \\ 0 \end{bmatrix}, \dots, e_n = \begin{bmatrix} 0 \\ \vdots \\ 1 \end{bmatrix}$ *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* $\{1, t, \ldots, t^n\}$ *formano una base di* $\mathbb{K}_n[t]$*. Quindi* $\dim \mathbb{K}_n[t] = n + 1$*;*

e) *Sia* $V = \left\{A \in M_{2\times 2}(\mathbb{R}) : A = A^T\right\}$*. Le matrici*

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix},$$

formano una base di V*. Quindi* $\dim V = 3$*.*

f) *Sia* $W = \left\{A \in M_{3\times 3}(\mathbb{R}) : A = -A^T\right\}$*. 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 W*. Quindi* $\dim W = 3$*.*

Proposizione 5.14. *Sia* $W = \mathcal{L}(v_1, \ldots, v_n) \subset V$*. Sia* k *il massimo numero dei vettori* $\{v_1, \ldots, v_n\}$ *linearmente indipendenti. Allora* $\dim W = k$

Dimostrazione. Siano $v_{j_1}, \ldots, v_{j_k}$ vettori linearmente indipendenti. Vogliamo dimostrare che $V = \mathcal{L}(v_{j_1}, \ldots, v_{j_k})$. È sufficiente dimostrare che $v_j \in \mathcal{L}(v_{j_1}, \ldots, v_{j_k})$ per $j = 1, \ldots, n$. Se $j = j_i$ per un certo $1 \leq i \leq k$, allora è vero. Supponiamo che $j \neq j_i$. I vettori $v_j, v_{j_1}, \ldots, v_{j_k}$ sono linearmente dipendenti. Quindi esistono $\alpha_j, \alpha_{j_1}, \ldots, \alpha_{j_k}$ non tutti nulli tali che

$$\alpha_j v_j + \alpha_{j_1} v_{j_1} + \cdots + \alpha_{j_k} v_{j_k} = 0.$$

Se $\alpha_j = 0$, allora avrei una combinazione lineare non banale dei vettori $v_{j_1}, \ldots, v_{j_k}$ uguale al vettore nullo. Assurdo perché i vettori $v_{j_1}, \ldots, v_{j_k}$ sono linearmente indipendenti. Quindi $\alpha_j \neq 0$ da cui segue che

$$v_j = (-\alpha_{j_1}/\alpha_j) v_{j_1} + \cdots + (-\alpha_{j_k}/\alpha_j) v_{j_k} \in \mathcal{L}(v_{j_1}, \ldots, v_{j_k}).$$

Quindi abbiamo dimostrato che

$$v_1, \ldots, v_n \in \mathcal{L}(v_{j_1}, \ldots, v_{j_k}),$$

ovvero

$$W = \mathcal{L}(v_1, \ldots, v_n) = \mathcal{L}(v_{j_1}, \ldots, v_{j_k}).$$

$\square$

Corollario 5.15. *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. La prima parte della dimostraziome segue dalla Proposizione anteriore e dalla definizione di rango di una matrice. 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 $rg(A) = \dim(A^1, \ldots, A^n)$, è semplice dimostrare che 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

$$\mathcal{L}(A^1, \ldots, A_i + \lambda A_j, \ldots, A_n) \subset \mathcal{L}(A^1, \ldots, A^n).$$

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

$$\mathcal{L}(A^1, \ldots, A^n) \subset \mathcal{L}(A^1, \ldots, A^i + \lambda A^j, \ldots, A^n).$$

$\square$

Proposizione 5.16. *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 $v \in V$. 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

$$\alpha v + \lambda_1 v_1 + \cdots + \lambda_n v_n = 0.$$

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 $v_j = \sum_{k=1, k\neq j}^{n} \alpha_j v_j$. Quindi $V = \mathcal{L}(v_1, \ldots, v_n) = \mathcal{L}(v_1, \ldots v_{j-1}, v_{j+1}, \ldots, v_n)$. Applicando il Lemma di Steinitz si ha che una base di V è 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 $\dim V = n$ ed $\mathcal{C} = \{w_1, \ldots, w_m\}$ 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 5.17. *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 5.6 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 5.18 (Teorema di completamento a base)**.** *Sia V uno spazio vettoriale su $\mathbb{K}$. 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.*

Dimostrazione. 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

$$\mathcal{L}(w_1, \ldots, w_m) \subset \mathcal{L}(v_1, \ldots, v_n).$$

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 $\mathcal{L}(v_1, \ldots, v_n) \subset \mathcal{L}(w_1, \ldots, w_m)$ 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 5.17, 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 $\mathcal{L}(v_{j_1}, w_1 \ldots, w_m)$ 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. □

Corollario 5.19. *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. Il Lemma di Steinitz implica che $m \leq n$, ovvero $\dim W \leq \dim V$. Il Teorema del completamente a base garantisce che esistono $n - m$ vettori di $\mathcal{B} = \{v_1, \ldots, v_n\}$ che aggiunti a $\mathcal{C} = \{w_1, \ldots, w_m\}$ formano una base di W. Quindi se $\dim W = \dim V$, allora $\mathcal{C} = \{w_1, \ldots, w_m\}$ è anche una base di V, da cui segue che $V = \mathcal{L}(w_1, \ldots, w_m) = W$. □

Diamo una prova alternativa della I parte del Teorema di Rouché-Capelli utilizzando i risultati anteriori.

Corollario 5.20. *Sia* $A \in \mathrm{M}_{m\times n}(\mathbb{K})$*. Il sistema* $AX = b$ *è compatibile se e solamente se* $\mathrm{rg}(A) = \mathrm{rg}(A|b)$*.*

Dimostrazione. Il sistema $AX = b$ è compatibile se e solamente se $b \in \mathcal{L}(A^1, \ldots, A^n)$ se e solamente se $\mathcal{L}(A^1, \ldots, A^n) = \mathcal{L}(A^1, \ldots, A^n, b)$. Poiché in generale vale l'inclusione,

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

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 5.15 si ha che il sistema $AX = b$ è compatibile se e solamente se $\mathrm{rg}(A) = \mathrm{rg}(A|b)$. $\square$

5.20.1 Metodi di Calcolo

Consideriamo $V = \mathbb{K}^n$ dove $\mathbb{K} = \mathbb{C}$ oppure $\mathbb{K} = \mathbb{R}$. Siano $X_1, \ldots, X_k$ vettori di $\mathbb{K}^n$. Allora:

a) se $n = k$, allora i vettori $X_1, \ldots X_n$ formano una base di $\mathbb{K}^n$ se e solamente se $\det(X_1, \ldots, X_n) \neq 0$;

b) Sia $W = \mathcal{L}(X_1, \ldots, X_k)$ dove $X_1, \ldots X_k \in \mathbb{K}^n$. Allora $\dim W = rg((X_1, \ldots, X_k))$. 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 $X_1, \ldots, X_k \in \mathbb{K}^n$ lineramente indipendenti. Possiamo completarli a base di $\mathbb{K}^n$. Una possibilità è quella di aggiungere $n - k$ vettori $Y_{k+1}, \ldots, Y_n$ tali che

$$\det(X_1, \ldots, X_k, Y_{k+1}, \ldots, Y_n) \neq 0.$$

   Un altra possiblità potrebbe essere la seguente.

Sia $\mathcal{C} = \{e_1, \dots e_n\}$ la base canonica di $\mathbb{K}^n$ e sia

$$A = (X_1, \dots X_k, e_1, \dots, e_n) \in M_{n\times(n+k)}(\mathbb{K}).$$

Si osservi che $rg(A) = n$. Infatti $rg(A) \leq n$ poiché la matrice A ha n righe. Dall'altro lato, i vettori $e_1, \dots, e_n$ sono linearmente indipendenti e quindi esistono almento n colonne linearmente indipendenti, ovvero $rg(A) \geq n$. Quindi $rg(A) = n$. Possiamo ridurre la matrice A a scala attraverso l'algoritmo di Gauss. Sia $S = (S^1, \dots, S^k, B^1, \dots, B^n)$ la matrice ottenuta. Poiché i vettori $X_1, \dots, X_k$ sono linearmente indipendenti, si può dimostrare che le colonne corrispondenti ai perni sono: $S^1, \dots, S^k, B^{j_1}, \dots, B^{j_{n-k}}$. Allora i vettori $X_1, \dots X_k, e_{j_1}, \dots e_{j_{n-k}}$ formano un base di V.

d) sia $\mathcal{B} = \{v_1, \dots, v_n\}$ 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 $x_1, \dots, x_n \in \mathbb{K}$ tale che

$$v = x_1 v_1 + \cdots + x_n v_n.$$

Se indichiamo con $A = (v_1, \dots, v_n) \in M_{n\times n}(\mathbb{K})$, allora le coordinate di v rispetto a $\mathcal{B}$ è l'unica soluzione del sistema lineare

$$AX = v.$$

e) Sia $W \subset \mathbb{K}^n$ un sottospazio vettoriale generato dai vettori $w_1, \dots, w_k$, ovvero $W = \mathcal{L}(w_1, \dots, w_k)$. Il vetore $w = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in W$ se e solamente se esistono degli scalari $\alpha_1, \dots, \alpha_k$ 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 $A = (w_1, \dots, w_k) \in M_{n\times k}(\mathbb{K})$. Possiamo ridurre a scala il sistema lineare mediante il metodo di Gauss. Indichiamo con $(S|c)$ la matrice completa del sistema lineare ridotta a scala e equivalente al sistema $AX = w$. Noi sappiamo che $(S|c)$ è compatibile se e solamente se le ultime $n - rg(A)$ coordinate di c sono nulle, i.e.,

$$c_{k+1} = 0, \dots, c_n = 0,$$

le quali definiscono delle equazioni lineari in $x_1, \ldots, x_n$. L'insieme delle soluzione è esattamente il sottospazio $W = \mathcal{L}(w_1, \ldots, w_k)$. Queste equazioni lineari sono chiamate le *equazioni cartesiane* di W.

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 $\mathcal{B}$. 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 $[w_1]_{\mathcal{B}}, \ldots, [w_k]_{\mathcal{B}} \in \mathbb{K}^n$. Possiamo completarli a base in $\mathbb{K}^n$, ovvero esistono $X_{j_1}, \ldots X_{j_{n-k}} \in \mathbb{K}^n$ tali che i vettori $[w_1], \ldots, [w_k]_{\mathcal{B}}, X_{j_1}, \ldots X_{j_{n-k}}$ formano una base di $\mathbb{K}^n$. I vettori

$$w_1, \ldots, w_k, \mathcal{F}^{-1}(X_{j_1}), \ldots, \mathcal{F}^{-1}(X_{j_{n-k}}),$$

formano una base di V. Il Teorema di completamento a base afferma che i vettori $X_{j_1}, \ldots X_{j_{n-k}}$ li posso scegliere fra i vettori della base canonica $\{e_1, \ldots, e_n\}$. Quindi utilizzando la stessa notazione, 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 $\mathbb{K}^n$. Quindi, tenendo in mente che $\mathcal{F}^{-1}(e_j) = v_j$ per $j = 1, \dots, n$, i vettori

$$w_1, \dots, w_k, v_{j_1}, \dots, v_{j_{n-k}},$$

formano una base di V.

Sia $W = \mathcal{L}(w_1, \dots, w_k) \subset V$. Per calcolare una base possiamo seguire il seguente procedimento.

Sia $W' = \mathcal{L}([w_1]_\mathcal{B}, \dots, [w_k]_\mathcal{B})$. $\dim W = \dim W' = \mathrm{rg}([w_1]_\mathcal{B}, \dots, [w_k]_\mathcal{B})$. Siano $1 \le j_1 < \cdots < j_k \le n$ tale che i vettori $[w_{j_1}]_\mathcal{B}, \dots, [w_{j_k}]_\mathcal{B})$ formano una base di W'. Allora i vettori $w_{j_1}, \dots, w_{j_k}$ formano una base di W.

Sia $\mathcal{C} = \{w_1, \dots, w_n\}$ una base di V. Le coordinate di un vettore v rispetto alla base $\mathcal{C}$ sono gli unici scalari $y_1, \dots, y_n \in \mathbb{K}$ tali che

$$v = y_1 w_1 + \cdots + y_n w_n.$$

Quindi

$$[v]_\mathcal{B} = y_1 [w_1]_\mathcal{B} + \cdots + y_n [w_n]_\mathcal{B},$$

ovvero le coordinate del vettore $[v]_\mathcal{B}$ rispetto alla base $\mathcal{C}' = ([w_1]_\mathcal{B}, \dots, [w_n]_\mathcal{B})$.

5.21 Formula di Grassman

Sia V uno spazio vettoriale su $\mathbb{K}$ e siano $U, W \subset V$ sottospazi vettoriali di V. $U \cap W$ è un sottospazio vettoriale (esercizio). In generale $U \cup W$ non è un sottospazio vettoriale. Definiamo

$$U + W := \{u + w : \, u \in U \; w \in W\}.$$

Proposizione 5.22. *$U + W$ è un sottospazio vettoriale di V.*

Dimostrazione. Proviamo che $U + W$ è chiuso rispetto alla moltiplicazione per scalare lasciando per esercizio che è chiuso rispetto alla somma.

Sia $\lambda \in \mathbb{K}$ e sia $z \in U + W$. Poiché $z = u + w$, si ha

$$\lambda z = \lambda u + \lambda w \in U + W.$$

□

È facile verificare che $U + W$ contiene sia U, sia W ed è il più piccolo sottospazio vettoriale di V che contiene $U \cup W$.

Lemma 5.23. *Sia $\mathcal{B} = \{v_1, \dots, v_n\}$ un sistema di generatori di U e $\mathcal{C} = \{w_1, \dots, w_m\}$ un sistema di generatori di W. Allora $\mathcal{B} \cup \mathcal{C}$ è un sistema di generatori di $U + W$.*

Dimostrazione. I vettori $v_1, \dots, v_n, w_1, \dots w_m \in U + W$ da cui segue

$$\mathcal{L}(v_1, \dots, v_n, w_1, \dots, w_m) \subset U + W.$$

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 $\mathcal{B} = \{v_1, \dots, v_n\}$, rispettivamente $\mathcal{C} = \{w_1, \dots, w_m\}$, è un sistema di generatori di U, rispettivamente W, esistono $\alpha_1, \dots, \alpha_n$, rispettivamente $\beta_1, \dots, \beta_m$, tale che

$$u = \alpha_1 v_1 + \cdots + \alpha_n v_n,$$

rispettivamente

$$w = \beta_1 w_1 + \cdots + \beta_m w_m.$$

Quindi

$$u + w = \alpha_1 v_1 + \cdots + \alpha_n v_n + \beta_1 w_1 + \cdots + \beta_m w_m \in \mathcal{L}(v_1, \dots, v_n, w_1, \dots, w_m),$$

ovvero $U + W \subset \mathcal{L}(v_1, \dots, v_n, w_1, \dots, w_m)$. Quindi

$$U + W \subset \mathcal{L}(v_1, \dots, v_n, w_1, \dots, w_m).$$

□

Teorema 5.24 (Formula di Grassman)**.**

$$\dim(U + W) + \dim(U \cap W) = \dim U + \dim W$$

Dimostrazione. Sia $\{s_1, \dots, s_k\}$ una base di $U \cap W$. Per il Teorema di completamento a base, possiamo completarla a base di U, $\{s_1, \dots, s_k, u_1, \dots, u_p\}$, rispettivamente a una base di W, $\{s_1, \dots, s_k, w_1, \dots, w_q\}$. Quindi $\dim(U \cap W) = k$, $\dim U = k + p$ ed infine $\dim W = k + q$. Per il Lemma anteriore, i vettori

$$\{s_1, \dots, s_k, u_1, \dots, u_p, w_1, \dots, w_q\},$$

formano un sistema di generatori di $U + W$. Vogliamo dimostrare che sono linearmente indipendenti e quindi formano una base di $U + W$. Consideriamo una combinazione lineare:

$$\underbrace{\alpha_1 s_1 + \cdots + \alpha_k s_k}_{s} + \underbrace{\beta_1 u_1 + \cdots + \beta_p u_p}_{u} + \underbrace{\gamma_1 w_1 + \cdots + \gamma_q w_q}_{w} = 0,$$

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 $\lambda_1, \dots \lambda_k$ rispettivamente $\mu_1, \dots, \mu_k$ 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 $\{s_1, \dots, s_k, u_1, \dots, u_p\}$, rispettivamente $\{s_1, \dots, s_k, w_1, \dots, w_q\}$, è una base di U, rispettivamente 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 $s_1, \dots, s_k$ sono linearmente indipendenti si ha $\alpha_1 = \cdots = \alpha_k = 0$, ovvero i vettori $s_1, \dots, s_k, u_1, \dots, u_p, w_1, \dots, w_q$ sono linearmente indipendenti. Inoltre

$$\dim(U+W) = k+p+q = (k+p)+(k+q)-k = \dim U + \dim W - \dim(U \cap W).$$

$\square$

Definizione 5.25. *Sia V uno spazio vettoriale su $\mathbb{K}$. Siano U e W due 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 $V = U \oplus W$.

Applicando la formula di Grassman si ha il seguente risultato.

Corollario 5.26. *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* $U \cap W = \{0\}$ *e* $\dim V = \dim U + \dim W$*.*

Dimostrazione. Tenendo in mente che $U \cap W = \{0\}$ è equivalente a $\dim(U \cap W) = 0$, ed 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 $V = U + W$. Applicando la formula di Grassmann si ha $\dim V = \dim U + \dim W$. Viceversa, supponiamo che $U \cap W = \{0\}$ e $\dim V = \dim U + \dim W$. Devo dimostrare che $V = U + W$. Applicando la formula di Grassmann si ha

$$\dim(U+W) = \dim U + \dim W,$$

cioé $\dim(U+W) = \dim V$. Per il Corollario 5.19 si ha $V = U + W$, concludento la dimostrazione. $\square$

Il seguente risultato descrive in maniere equivalente cosa significa che V è somma diretta di U e W.

Corollario 5.27. *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 almento due maniere distinte. Assurdo. Quindi $U \cap W = \{0\}$. $\square$

Esempio 5.28. *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 5.29. *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 $z = 0$. □

5.29.1 Metodi di Calcolo

Siano $U = \mathcal{L}(X_1, \ldots, X_k)$ e $W = \mathcal{L}(Y_1, \ldots, Y_s)$ sottospazi vettoriale di $\mathbb{K}^n$. Sia $A = (X_1, \ldots, X_k, Y_1, \ldots, Y_s) \in M_{n\times(k+s)}(\mathbb{K})$. Allora $\dim(U+W) = \mathrm{rg}(A)$.

Se considero le equazioni cartesiane di U e W, rispettivamente, allora l'unione di queste equazioni sono le equazioni cartesiane di $U \cap W$.

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:

- $\dim U = rg(([u_1]_{\mathcal{B}}, \dots, [u_p]_{\mathcal{B}})$;
- $\dim W = rg(([w_1]_{\mathcal{B}}, \dots, [w_q]_{\mathcal{B}})$;
- $\dim U + W = rg(([u_1]_{\mathcal{B}}, \dots, [u_p]_{\mathcal{B}}, [w_1]_{\mathcal{B}}, \dots, [w_q]_{\mathcal{B}}))$.
- 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*

- $T(0) = 0$*;*
- $T(-v) = -T(v)$*.*

Esempio 6.4.

a) $T : \mathbb{R} \longrightarrow \mathbb{R}$, $T(x) = x + 1$ *non è una applicazione lineare;*

b) $T : \mathbb{R} \longrightarrow \mathbb{R}$, $T(x) = x^2$ *non è una applicazione lineare;*

c) $T : V \longrightarrow W$, $v \mapsto 0$ *è lineare;*

d) $\mathrm{Id}_V : V \longrightarrow V$, $v \mapsto v$ *è una applicazione lineare;*

e) L'applicazione $A \in M_{m\times n}(\mathbb{R}) \mapsto A^T \in M_{n\times m}(\mathbb{R})$ *è una applicazione lineare;*

f) L'applicazione $A \in M_{m\times n}(\mathbb{C}) \mapsto A^* \in M_{n\times m}(\mathbb{C})$ *non è una applicazione lineare: è additiva ma non è omogenea;*

g) 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;

h) 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.

i) 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.

j) L'applicazione

$$M_{n\times n}(\mathbb{K}) \longrightarrow \mathbb{K} \qquad A \mapsto \mathrm{Tr}(A)$$

è una applicazione lineare.

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 $T(v) = L(v)$. Definiamo

$$(T + L)(v) := T(v) + L(v), \qquad (\lambda T)(v) = \lambda T(v).$$

Proposizione 6.5. *Se V, W sono spazi vettoriali su $\mathbb{K}$, allora $L(V, W)$ con le operazioni di somma e moltiplicazione per scalare definite come sopra è uno spazio vettoriale su $\mathbb{K}$.*

Proposizione 6.6. *Siano $f \in L(V, W)$ e $g \in L(W, Z)$. Allora $g \circ f \in L(V, Z)$ 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. $\square$

Il prossimo risultato garantisce che una applicazione lineare è completamente determinate dai valori che assume su una base di V.

Proposizione 6.7. *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 $\mathcal{B} = \{v_1, \dots, v_n\}$ è una base di V esistono, e sono unici, $x_1, \dots, x_n \in \mathbb{K}$, tale che $v = x_1 v_1 + \cdots + x_n v_n$. Quindi

$$\begin{aligned} T(v) &= T(x_1 v_1 + \cdots + x_n v_n) \\ &= x_1 T(v_1) + \cdots + x_n T(v_n) \\ &= x_1 L(v_1) + \cdots + x_n L(v_n) \\ &= L(x_1 v_1 + \cdots + x_n v_n) \\ &= L(v). \end{aligned}$$

Adesso dimostriamo che esiste una applicazione lineare $T : V \longrightarrow W$ tale che $T(v_1) = w_1, \dots, T(v_n) = w_n$.

Sia $v \in V$. Allora esistono, e sono unici, $x_1, \dots, x_n \in \mathbb{K}$ 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 $T(v_1) = w_1, \dots, T(v_n) = w_n$. Rimane 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

$$\begin{aligned} T(v + w) &= (x_1 + y_1) w_1 + \cdots + (x_n + y_n) w_n \\ &= x_1 w_1 + \cdots + x_n w_n + y_1 w_1 + \cdots + y_n w_n \\ &= T(v) + T(w). \end{aligned}$$

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 6.8. *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. 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

$$\begin{aligned} T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) &= (-x - y + 2z)\begin{bmatrix} 1 \\ 1 \\ 2 \\ -1 \end{bmatrix} + (-x + z)\begin{bmatrix} 0 \\ -3 \\ 1 \\ 1 \end{bmatrix} \\ &\quad + (2x + y - 2z)\begin{bmatrix} 2 \\ -1 \\ 5 \\ -1 \end{bmatrix} \\ &= \begin{bmatrix} 3x + y - 2z \\ -2y + z \\ 7x + 3y - 5z \\ -2x + z \end{bmatrix}. \end{aligned}$$

Corollario 6.9. *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 $T(v_1) = L(v_1), \ldots, T(v_n) = L(v_n)$.*

Definizione 6.10. *Sia $T : V \longrightarrow W$ una applicazione lineare. Si pone*

- $Ker\, T = \{v \in V : T(v) = 0\}$ *(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 6.11. *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.

- Siano $v, w \in Ker\, T$. Allora $T(v) = T(w) = 0$. Dobbiamo dimostrare che $v + w \in Ker\, T$ e $\lambda v \in Ker\, T$ per ogni $\lambda \in \mathbb{K}$, cioè

  $$T(v_1 + v_2) = 0, \ T(\lambda v) = 0.$$

  Poiché T è lineare, si ha

  $$T(v_1 + v_2) = T(v_1) + T(v_2) = 0,$$

  rispettivamente

  $$T(\lambda v) = \lambda T(v) = 0.$$

- Siano $w_1, w_2 \in Im\, T$ e $\lambda \in \mathbb{K}$. 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 $\mathcal{B} = \{v_1, \ldots, v_n\}$ una base di V e sia $w \in Im\, T$. Allora $w = T(v)$ per un certo $v \in V$. Poiché $\mathcal{B} = \{v_1, \ldots, v_n\}$ è una base di V si ha $v = x_1 v_1 + \cdots + x_n v_n$. Per la linearità di T, si ha

$$w = T(v) = x_1 T(v_1) + \cdots + x_n T(v_n) \in \mathcal{L}(T(v_1), \ldots, T(v_n)).$$

Quindi $Im\, T \subset \mathcal{L}(T(v_1), \ldots, T(v_n))$. Poiché $T(v_1), \ldots, T(v_n) \in Im\, T$, tenendo in mente che $Im\, T$ è un sottospazio vettoriale di W, si ha

$$\mathcal{L}(T(v_1), \ldots, T(v_n)) \subset Im\, T.$$

Quindi

$$Im\, T = \mathcal{L}(T(v_1), \ldots, T(v_n))$$

e $\dim Im\, T \leq \dim V$.

$\square$

Esempio 6.12. *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 6.11, 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

$$\begin{aligned} T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) &= (-x - y + 2z) \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} + (-x + z) \begin{bmatrix} 3 \\ -2 \\ 1 \end{bmatrix} \\ &+ (2x + y - 2z) \begin{bmatrix} 2 \\ 3 \\ 5 \end{bmatrix} \\ &= \begin{bmatrix} y + z \\ 7x + 2y - 6z \\ 7x + 3y - 5z \end{bmatrix}. \end{aligned}$$

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* $\left\{ \begin{bmatrix} 8/7 \\ -1 \\ 1 \end{bmatrix} \right\}$.

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 6.11, 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*

$$\left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix} \right\}.$$

Osservazione 6.14. *Sia* $T : V \longrightarrow W$ *una applicazione lineare. Ricordiamo che se* $Z \subset W$*, allora la controimmagine di* Z *rispetto a* T*, che indicheremo con* $T^{-1}(Z)$*, 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 $L \subset V$, 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 6.15. *Sia $T : V \longrightarrow W$ una applicazione lineare. Allora:*

- *T è iniettiva se e solamente se $Ker\, T = \{0\}$;*
- *T è suriettiva se e solamente se $Im\, T = W$ 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 $T(0) = 0$. Quindi se $v \neq 0$, allora $T(v) \neq T(0) = 0$, ovvero $\operatorname{Ker} T = \{0\}$.

Viceversa, supponiamo che $Ker\, T = \{0\}$. 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 $v - w \in Ker\, T = \{0\}$ da cui segue $v = w$.

Un'applicazione T è suriettiva se $\operatorname{Im} T = W$. Poiché, in generale vale che $\operatorname{Im} T \subset W$, per il corollario 5.19, la condizione $\operatorname{Im} T = W$ è equivalente a $\dim W = \dim \operatorname{Im} T$. □

Teorema 6.16 (Teorema della dimensione)**.** *Sia $T : V \longrightarrow W$ una applicazione lineare. Allora*

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

Dimostrazione. Sia $(v_1, \ldots, v_k)$ una base di $Ker\, T$. Per il Teorema del complemento della base, esistono $v_{k+1}, \ldots, v_n$ tale che $(v_1, \ldots, v_n)$ è una base di V. Poiché

$$Im\, T = \mathcal{L}(T(v_1), \ldots, T(v_n)) = \mathcal{L}(T(v_{k+1}), \ldots, T(v_n)),$$

è sufficiente dimostrare 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 $\alpha_{k+1}, \ldots, \alpha_n \in \mathbb{K}$ tali che

$$\alpha_{k+1}T(v_{k+1}) + \cdots + \alpha_n T(v_n) = 0.$$

Per la linearità segue che

$$T(\alpha_{k+1}v_{k+1} + \cdots + \alpha_n v_n) = 0,$$

ovvero $\alpha_{k+1}v_{k+1} + \cdots + \alpha_n v_n \in Ker\, T$. Allora esistono $\alpha_1, \ldots, \alpha_k \in \mathbb{K}$ tali che

$$\alpha_1 v_1 + \cdots + \alpha_k v_k = \alpha_{k+1}v_{k+1} + \cdots + \alpha_n v_n,$$

ovvero

$$\alpha_1 v_1 + \cdots + \alpha_k v_k - \alpha_{k+1}v_{k+1} - \cdots - \alpha_n v_n = 0.$$

Poiché i vettori $v_1, \ldots, v_n$ formano una base di V, si ha $\alpha_1 = \cdots = \alpha_n = 0$, ed in particolare $\alpha_{k+1} = \cdots = \alpha_n = 0$, ovvero i vettori $T(v_{r+1}), \ldots, T(v_n)$ sono linearmenti indipendenti. □

Vediamo alcune conseguenze del Teorema della dimensione.

Corollario 6.17. *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 6.15, T è iniettiva se e solamente se $\operatorname{Ker} T = \{0\}$. Applicando il Teorema della dimensione, si ha che T è iniettiva se e solamente se $\dim Im\, T = \dim V$. □

Corollario 6.18. *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 6.15, 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 6.19. *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. $\square$

Definizione 6.20. *Due spazi vettoriali* V, W *sono isomorfi se esiste una applicazione lineare* $T : V \longrightarrow W$ *iniettiva e suriettiva, ovvero biunivoca.*

Proposizione 6.21. *Due spazi vettoriali sono isomorfi se e solamente se* $\dim V = \dim W$.

Dimostrazione. Applicando il Corollario 6.18 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 6.7). 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: $L_A(X) = AX$. Allora

a) $Ker\, L_A = \{X \in \mathbb{K}^n : AX = 0\} = Sol(A|0)$;

b) $Im\, L_A = \{b \in \mathbb{K}^n : \exists X \in \mathbb{K}^n$ 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. In particolare $b \in Im\, L_A$ se e solamente se $\mathrm{rg}(A) = \mathrm{rg}(A|b)$;

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

$$\mathrm{Im}\, L_A = \mathcal{L}(L_A(e_1), \ldots, L_A(e_n)) = \mathcal{L}(A^1, \ldots, A^n),$$

ovvero $\dim \mathrm{Im}\, L = \mathrm{rg}(A)$;

d) sia $b \in \mathbb{K}^m$. Allora $L_A^{-1}(b) = \{X \in \mathbb{K}^n : AX = b\} = \mathrm{Sol}(A|b)$.

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*

$$n = \dim Ker\, L_A + rg(A)$$

.

Una conseguenza del risultato anteriore è la dimensione dello spazio delle soluzioni di un sistema lineare omogeneo.

Corollario 7.3. *Sia* $A \in \mathrm{M}_{m\times n}(\mathbb{K})$. *Allora* $\dim \mathrm{Sol}(A|0) = n - \mathrm{rg}(A)$.

Applicando corollari del Teorema dell dimensione 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* $\det(A) \neq 0$.

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 invertibile se e solamente se A è una matrice quadrata e $\dim \mathrm{Im}\, L_A = \dim \mathbb{K}^n$ 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. □

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 $ImL_A = \{0\}$ se e solamente se

$$\mathcal{L}(A^1, \ldots, A^n) = \{0\},$$

quindi se e solamente se $A = 0$.

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. Vogliamo dimostrare che $L_A \circ L_B = L_{AB}$. Infatti

$$(L_A \circ L_B)(X) = L_A(L_B(X)) = L_A(BX) = ABX = (AB)(X) = L_{AB}(X).$$

Infine, se $A \in \mathrm{M}_{n\times n}(\mathbb{K})$ è invertibile allora L_A è invertibile e viceversa. Inoltre, tenendo in mente che $L_{\mathrm{Id}} = \mathrm{Id}_{\mathbb{K}^n}$, si ha $L_A \circ L_{A^{-1}} = L_{AA^{-1}} = L_{\mathrm{Id}} = \mathrm{Id}_{\mathbb{K}^n}$ e $L_{A^{-1}} \circ L_A = L_{A^{-1}A} = L_{\mathrm{Id}} = \mathrm{Id}_{\mathbb{K}^n}$. Riassumendo, L_A è invertibile se e solamente se $A \in \mathrm{M}_{n\times n}(\mathbb{K})$ è invertibile. Inoltre

$$(L_A)^{-1} = L_{A^{-1}}$$

Vogliamo dimostrare che ogni applicazione lineare è della forma L_A, ovvero che $\mathcal{K}$ è suriettiva e quindi un isomorfismo.

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

$$T(X) = x_1T(e_1) + \cdots + x_nT(e_n) = M_T \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

dove $M_T = (T(e_1), \ldots, T(e_n)) \in M_{m\times n}(\mathbb{K})$. Quindi

$$T(X) = M_TX = L_{M_T}(X).$$

Riassumento abbiamo dimostrato il seguente risultato.

Corollario 7.5. *Sia $T : \mathbb{K}^n \longrightarrow \mathbb{K}^m$ una applicazione lineare. Allora $T = L_{M_T}$ dove*

$$M_T = (T(e_1), \ldots, T(e_n)) \in M_{m\times n}(\mathbb{K})$$

e valgono le seguenti proprietà:

- $KerT = Sol(M_T|0)$;
- $ImT = \mathcal{L}(M_T^1, \ldots, M_T^n)$ *e* $\dim Im\, T = rg(M_T)$.

Inoltre:

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 invertibile se e solamente se* M_T *è invertibile;*

Il prossimo risultato riguarda la corrispondenza $T \mapsto M_T$.

Proposizione 7.6. *L'applicazione*

$$\mathcal{H} : \mathrm{Lin}(\mathbb{K}^n, \mathbb{K}^m) \mapsto \mathrm{M}_{m\times n}(\mathbb{K}) \qquad T \mapsto M_T,$$

è un isomorfismo di spazi vettoriali, ovvero valgono le seguenti proprietà:

a) $M_{T+L} = M_T + M_L$;

b) $M_{\lambda T} = \lambda M_T$.

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* $M_{T^{-1}} = M_T^{-1}$.

Facoltativa. Abbiamo dimostrato che

$$\mathcal{K} : \mathrm{M}_{m\times n}(\mathbb{K}) \longrightarrow \mathrm{Lin}(\mathbb{K}^n, \mathbb{K}^m) \qquad A \mapsto L_A.$$

è un isomorfismo. E' facile verifivare che l'applicazione

$$\mathcal{H} : \mathrm{Lin}(\mathbb{K}^n, \mathbb{K}^m) \longrightarrow \mathrm{M}_{m\times n}(\mathbb{K}) \qquad T \mapsto M_T,$$

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 $T^{-1} = L_{M_T^{-1}}$. □

7.7 Matrice associata ad una applicazione lineare

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 $v = x_1v_1 + \cdots + x_nv_n$. Come usuale indichiamo con $[v]_{\mathcal{B}} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ le coordinate di v rispetto a $\mathcal{B}$. Poiché $\mathcal{F}_{\mathcal{B}}; V \longrightarrow \mathbb{K}^n$ è lineare, si ha

$$[T(v)]_{\mathcal{C}} = x_1[T(v_1)]_{\mathcal{C}} + \cdots + x_n[T(v_n)]_{\mathcal{C}} = \mathcal{M}_{\mathcal{C},\mathcal{B}}(T)[v]_{\mathcal{B}},$$

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* $\mathcal{B}$ *e* $\mathcal{C}$. 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., $\mathcal{F}_{\mathcal{C}} \circ T = L_{\mathcal{M}_{\mathcal{C},\mathcal{B}}(T)} \circ \mathcal{F}_{\mathcal{B}}$.

Proposizione 7.8. *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*

- $\mathcal{F}_{\mathcal{C}}(Im\, T) = Im\, L_{\mathcal{M}_{\mathcal{C},\mathcal{B}}(T)}$;
- $\mathcal{F}_{\mathcal{B}}(Ker\, T) = Ker\, L_{\mathcal{M}_{\mathcal{C},\mathcal{B}}(T)}$;

Dimostrazione. 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

$$\mathcal{F}_{\mathcal{B}}(Ker\, T) = \mathrm{Sol}(\mathcal{M}_{\mathcal{C},\mathcal{B}}(T)|0),$$

ovvero $\dim \mathrm{Ker}\, T = \dim V - rg(\mathcal{M}_{\mathcal{C},\mathcal{B}}(T))$. 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 alcuni corollari del Teorema della dimensione si ha il seguente corollario..

Corollario 7.9. *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 invertibile se e solamente se $\mathcal{M}_{\mathcal{C},\mathcal{B}}(T)$ è invertibile.*

Esempio 7.10. *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:*

$$T(X) = X - \operatorname{Tr}(X)A.$$

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

$$\mathcal{F}_{\mathcal{B}} : M_{2\times 2}(\mathbb{R}) \longrightarrow \mathbb{R}^4 \qquad \begin{bmatrix} a & b \\ c & d \end{bmatrix} \mapsto \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix},$$

si ha

$$\mathcal{M}_{\mathcal{B},\mathcal{B}}(T) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & -1 \\ -1 & 0 & 1 & -1 \\ -1 & 0 & 0 & 0 \end{bmatrix}.$$

Poiché $\mathrm{rg}(\mathcal{M}_{\mathcal{B},\mathcal{B}}(T)) = 3$, *allora* $\dim \mathrm{Im}\, T = 3$ *e* $\dim \mathrm{Ker}\, T = 1$

Esempio 7.11. *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 $X \in \mathrm{Ker}\, T$ *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 $\mathrm{Ker}\, T$ *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 7.12. *Siano V e W spazi vettoriali su $\mathbb{K}$ e siano $\mathcal{B} = \{v_1, \dots, v_n\}$ e $\mathcal{C} = \{w_1, \dots, w_m\}$ 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.

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

$$T(v_1) = \sum_{i=1}^{m} a_{i1} w_i, \dots, T(v_n) = \sum_{i=1}^{m} a_{in} w_i.$$

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. L'iniettività segue dal fatto che $T = 0$ se e solamente se $T(v) = 0$ per ogni $v \in V$ e quindi se e sòlamente 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 $\mathrm{Lin}(\mathbb{K}^n, \mathbb{K}^m)$, la corrispondenza biunivoca sopra definita verifica altre proprietà.

Proposizione 7.13. *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 $\mathcal{M}_{\mathcal{B},\mathcal{B}}(Id_V) = \mathrm{Id}_{\dim V}$;

e) se $T : V \longrightarrow W$ è invertibile, allora $\mathcal{M}_{\mathcal{B},\mathcal{C}}(T^{-1}) = \mathcal{M}_{\mathcal{C},\mathcal{B}}(T)^{-1}$.

Dimostrazione. 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

$$\mathcal{M}_{\mathcal{C},\mathcal{X}}(T \circ G) = \mathcal{M}_{\mathcal{C},\mathcal{B}}(T)\mathcal{M}_{\mathcal{B},\mathcal{X}}(G).$$

□

7.14 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 $v = x_1 v_1 + \cdots + x_n v_n$. 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

$$[v]_{\mathcal{B}} = [x'_1 w_1 + \cdots + x'_n w_n]_{\mathcal{B}} = x'_1[w_1]_{\mathcal{B}} + \cdots + x'_n[w_n]_{\mathcal{B}} = \mathcal{M}(\mathcal{B},\mathcal{C})[v]_{\mathcal{C}},$$

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 $\mathcal{M}(\mathcal{B},\mathcal{C}) = \mathcal{M}_{\mathcal{B},\mathcal{C}}(Id)$ da cui segue che la matrice del cambiamento di base è invertibile.

Proposizione 7.15. *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* $(\mathcal{M}(\mathcal{B},\mathcal{C}))^{-1} = \mathcal{M}(\mathcal{C},\mathcal{B})$.

Dimostrazione. Dimostreremo solamente la seconda proprietà. 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 $\mathcal{M}(\mathcal{B},\mathcal{D}) = \mathcal{M}(\mathcal{B},\mathcal{C})\mathcal{M}(\mathcal{C},\mathcal{D})$. □

Esempio 7.16. *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

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix}_{\mathcal{B}} = \begin{bmatrix} -x-y+2z \\ -x+z \\ 2x+y-2z \end{bmatrix},$$

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

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix}_{\mathcal{B}'} = \begin{bmatrix} x - y \\ y - z \\ z \end{bmatrix},$$

si ha

$$\mathcal{M}(\mathcal{B}', \mathcal{B}) = \begin{bmatrix} 1 & -2 & 0 \\ -1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}.$$

Esempio 7.17. *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

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix}_{\mathcal{B}} = \begin{bmatrix} -x - y + 2z \\ -x + z \\ 2x + y - 2z \end{bmatrix},$$

si ha

$$\mathcal{M}(\mathcal{B}, \mathcal{C}) = \begin{bmatrix} -1 & -1 & 2 \\ -1 & 0 & 1 \\ 2 & 1 & -2 \end{bmatrix}.$$

Teorema 7.18. *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.

$$\begin{aligned}[T(v)]_{\mathcal{C}} &= \mathcal{M}_{\mathcal{C},\mathcal{B}}(T)[v]_{\mathcal{B}} \\ &= \mathcal{M}_{\mathcal{C},\mathcal{B}}(T)\mathcal{M}(\mathcal{B},\mathcal{B}')[v]_{\mathcal{B}'}.\end{aligned}$$

e

$$[T(v)]_{\mathcal{C}} = \mathcal{M}(\mathcal{C},\mathcal{C}')[T(v)]_{\mathcal{C}'},$$

ovvero, ricordando che $\mathcal{M}(\mathcal{C},\mathcal{C}') = (\mathcal{M}(\mathcal{C}',\mathcal{C}))^{-1}$,

$$[T(v)]_{\mathcal{C}'} = \mathcal{M}(\mathcal{C}',\mathcal{C})\mathcal{M}_{\mathcal{C},\mathcal{B}}(T)\mathcal{M}(\mathcal{B},\mathcal{B}')[v]_{\mathcal{B}'}.$$

□

Corollario 7.19. *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 7.20. *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 7.17 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 7.21. *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 7.18, 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

$$\mathcal{M}(\mathcal{B},\mathcal{C}) = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \qquad \mathcal{M}(\mathcal{C},\mathcal{B}) = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix},$$

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 7.22. *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. □

Il risultato anteriore suggerisce la seguente definizione.

Definizione 7.23. *Siano $A, B \in \mathrm{M}_{n\times n}(\mathbb{K})$. Diremo che A e B sono* simili *se esiste una matrice $P \in \mathrm{M}_{n\times n}(\mathbb{K})$ invertibile, tale che $A = P^{-1}BP$.*

La relazione $A \sim B$, i.e., A e B sono matrici simili, è una relazione di equivalenza. Quindi $A \sim A$; $A \sim B$ 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.

Il corollario anteriore dimostra che se $T : V \longrightarrow V$ è una applicazione lineare, 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* $X = 0$ *(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 $X = 0$. $\square$

Sia X un vettore. Chiameremo *norma* o *lunghezza* di X il numero reale $\| X \| = \sqrt{\langle X, X\rangle}$. Dunque se $X \neq 0$, 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 il vettore

$$\mathrm{pr}_w(v) = \frac{\langle v, w\rangle}{\| w \|^2} w.$$

Proposizione 8.3. *(Disuguaglianza di Cauchy-Schwartz)*

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

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

Dimostrazione. Se $Y = 0$, la disuguaglianza è vera. Supponiamo che $Y \neq 0$. Allora

$$0 \leq \langle X + tY, X + tY\rangle = \langle X, X\rangle + 2\langle X, Y\rangle t + \langle Y, Y\rangle t^2.$$

L'equazione anteriore definisce una parabola il cui disciminante è non positivo. Quindi

$$0 \geq \Delta = b^2 - 4ac = 4\langle X, Y\rangle^2 - 4\langle X, X\rangle\langle Y, Y\rangle,$$

da cui segue $|\langle X, Y\rangle| \leq \| X \| \| Y \|$.

Se $|\langle X, Y\rangle| = \| X \| \| Y \|$, allora $\Delta = 0$. Quindi esiste $t_o \in \mathbb{R}$ tale che $0 = \langle X + t_oY, X + t_oY\rangle$. Per le proprietà del prodotto scalare si ha $X + t_oY = 0$, ovvero X e Y sono linearmente dipendenti. $\square$

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 $v_1, \ldots, v_k \in \mathbb{R}^n$ non nulli e ortogonali a due a due. Allora $v_1, \ldots, v_k$ sono linearmente indipendenti.*

Dimostrazione. Sia $\alpha_1 v_1 + \cdots + \alpha_k v_k = 0$ una combinazione lineare uguale al vettore nullo. Sia $1 \leq j \leq k$. Allora

$$\begin{aligned} 0 &= \langle 0, v_j \rangle \\ &= \langle \alpha_1 v_1 + \cdots + \alpha_k v_k, v_j \rangle \\ &= \alpha_1 \langle v_1, v_j \rangle + \cdots + \alpha_k \langle v_k, v_j \rangle \\ &= \alpha_i \langle v_j, v_j \rangle \text{ (essendo } v_1, \ldots, v_k \text{ ortogonali a due a due)} \end{aligned}$$

Poiché $v_j \neq 0$, $\langle v_j, v_j \rangle > 0$ da cui segue $\alpha_j = 0$ per $j = 1, \ldots, k$, ovvero i vettori $v_1, \ldots, v_k$ sono linearmente indipendenti. □

Definizione 8.9. *Una base $\mathcal{B} = \{v_1, \ldots, v_n\}$ si dice ortogonale se i vettori $v_1, \ldots, v_n$ sono a due a due ortogonali. Diremo che $\mathcal{B} = \{v_1, \ldots, v_n\}$ è una base ortonormale se $\mathcal{B} = \{v_1, \ldots, v_n\}$ 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* $[v]_{\mathcal{B}} = \begin{bmatrix} \langle v, v_1 \rangle \\ \vdots \\ \langle v, v_n \rangle \end{bmatrix}$;
- $\| v \| = \sqrt{\langle v, v_1 \rangle^2 + \cdots + \langle v, v_n \rangle^2}$.
- $\langle v, w \rangle = [v]_{\mathcal{B}}^T [w]_{\mathcal{B}}$.

Dimostrazione. Se $v = x_1 v_1 + \cdots + x_n v_n$. Sia $1 \leq j \leq n$. Allora
$$\langle v, v_j \rangle = \langle \sum_{m=1}^{n} x_m v_m, v_j \rangle = \sum_{m=1} x_m \langle v_m, v_j \rangle = x_j \langle v_j, v_j \rangle,$$
ovvero $x_j = \frac{\langle v, v_j \rangle}{\langle v_j, v_j \rangle}$ 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 $\| v \| = \sqrt{\langle v, v_1 \rangle^2 + \cdots + \langle v, v_n \rangle^2}$. $\square$

Il prossimo risultato fornisce un algoritmo per costruire basi ortogonali, e ortonormali, di sottospazi vettoriali di $\mathbb{R}^n$.

Proposizione 8.11 (Procedimento di ortogonalizzazione di Gram-Schmidt)**.** *Siano $v_1, \dots, v_k \in \mathbb{R}^n$ 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 $\mathcal{L}(w_1, \dots, w_j) = \mathcal{L}(v_1, \dots, v_j)$ per $j = 1, \dots, k$.

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 $v_1, \dots, v_{k-1}$ sono lineaременте indipendenti. Per ipotesi induttiva i vettori $w_1, \dots, w_{k-1}$ 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 $w_k = 0$ 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 $w_k \neq 0$. Inoltre

$$\begin{aligned} \langle w_k, w_s \rangle &= \langle v_k - \sum_{j=1}^{k-1} \frac{\langle v_k, w_j \rangle}{\langle w_j, w_j \rangle} w_j, w_s \rangle \\ &= \langle v_k, w_s \rangle - \sum_{j=1}^{k-1} \Big\langle \frac{\langle v_k, w_j \rangle}{\langle w_j, w_j \rangle} w_j, w_s \Big\rangle \\ &= \langle v_k, w_s \rangle - \sum_{j=1}^{k-1} \frac{\langle v_k, w_j \rangle}{\langle w_j, w_j \rangle} \langle w_j, w_s \rangle \\ &= \langle v_k, w_s \rangle - \frac{\langle w_s, v_k \rangle}{\langle w_s, w_s \rangle} \langle w_s, w_s \rangle \\ &= \langle v_k, w_s \rangle - \langle v_k, w_s \rangle \\ &= 0. \end{aligned}$$

Infine, i vettori $w_1, \dots, w_k$ sono a due a due ortogonali e quindi linearmente indipendenti. Dimostriamo che $\mathcal{L}(v_1, \dots, v_k) = \mathcal{L}(w_1, \dots, w_k)$. Per ipotesi induttiva $w_1, \dots, w_{k-1} \in \mathcal{L}(v_1, \dots, v_{k-1}) \subset \mathcal{L}(v_1, \dots, v_k)$. 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 $\mathcal{L}(w_1, \dots, w_k) \subset \mathcal{L}(v_1, \dots, v_k)$. Poiché $\dim L(w_1, \dots, w_k) = \dim \mathcal{L}(v_1, \dots, v_k) = k$ si ha $\mathcal{L}(w_1, \dots, w_k) = \mathcal{L}(v_1, \dots, v_k)$ $\square$

Corollario 8.12. *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. $\square$

Corollario 8.13. *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?). $\square$

Vediamo infine un legame fra matrici ortogonali e basi ortonormali.

Proposizione 8.14. *Sia $A = (A^1, \dots, A^n) \in \mathrm{M}_{n \times n}(\mathbb{R})$. A è una matrice ortogonale se e solamente se $A^1, \dots, A^n$ formano una base ortonormale.*

Dimostrazione. La matrice A è ortogonale se e solamente se $AA^T = A^T A = \mathrm{Id}$. Si può dimostrare che $A^T A = \mathrm{Id}$, rispettivamente $AA^T = \mathrm{Id}$, implica $AA^T = \mathrm{Id}$, rispettivamente $A^T A = \mathrm{Id}$.

Sia $C = A^T A$. Allora

$$c_{ij} = \sum_{m=1}^{n} a^T_{im} a_{mj} = \sum_{m=1}^{n} a_{mi} a_{mj} = \langle A^i, A^j \rangle.$$

La matrice A è ortogonale se e solamente se $C = \mathrm{Id}$, 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 $\{A^1, \dots, A^n\}$ formano una base ortonormale. $\square$

8.15 Proiezione ortogonali

Sia $S \subset \mathbb{R}^n$. Porremo

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

ovvero l'insieme dei vettori di $\mathbb{R}^n$ che sono ortogonali a ogni elemento di S.

Proposizione 8.16. *L'insieme $S^\perp$ è un sottospazio vettoriale di $\mathbb{R}^n$. Inoltre, se $S = \mathcal{L}(v_1, \ldots, v_k)$, allora*

$$S^\perp = \{v \in \mathbb{R}^n : \langle v, v_1 \rangle = \cdots = \langle v, v_k \rangle = 0\}$$

Dimostrazione. Lasciamo per esercizio la dimostrazione che $S^\perp$ è un sottospazio vettoriale.

Supponiamo che $S = \mathcal{L}(v_1, \ldots, v_k)$. Sia $S' = \{v \in \mathbb{R}^n : \langle v, v_1 \rangle = \cdots = \langle v, v_k \rangle = 0\}$. Vogliamo dimostrare che $S^\perp = S'$.

L'inclusione $S^\perp \subset S'$ è immediata poiché se $v \in S^\perp$, allora v in particolare è ortogonale ai vettori $v_1, \ldots, v_k \in S$ e quindi $v \in S'$.

Sia $w \in S'$. Vogliamo dimostrare che

$$\langle w, s \rangle = 0, \ \forall s \in S.$$

Sia $s \in \mathcal{L}(v_1, \ldots, v_k)$. Esistono $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$ tali che

$$s = \alpha_1 v_1 + \cdots + \alpha_k v_k.$$

Quindi

$$\begin{aligned} \langle w, s \rangle &= \langle w, \alpha_1 v_1 + \cdots + \alpha_k v_k \rangle \\ &= \alpha_1 \langle w, v_1 \rangle + \cdots + \alpha_k \langle w, v_k \rangle \\ &= 0 \text{ (essendo } w \in S') \end{aligned}$$

da cui segue che $w \in S^\perp$. □

Sia $W \subset \mathbb{R}^n$ un sottospazio vettoriale.

Proposizione 8.17. *$\mathbb{R}^n$ 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 $u = 0$. Sia $\mathcal{C} = \{w_1, \ldots, w_k\}$ una base di W. Allora

$$W^\perp = \{X \in \mathbb{R}^n : \langle X, w_1 \rangle = \cdots = \langle X, w_k \rangle = 0\}.$$

Sia $A = (w_1, \dots, w_k)$. Allora $W^\perp = \mathrm{Sol}(A^T|0)$ e quindi $\dim W^\perp = n - \mathrm{rg}(A^T) = n - \mathrm{rg}(A) = n - \dim W$. Per la formula di Grassmann si ha

$$\dim(W + W^\perp) = \dim W + \dim W^\perp = n.$$

Quindi $\mathbb{R}^n = W + W^\perp$. Poiché $W \cap W^\perp = \{0\}$ si ha che $\mathbb{R}^n$ è in somma diretta di W e $W^\perp$.

Un'altra dimostrazione è la seguente.

Sia $\{w_1, \dots, w_k\}$ una base di W. Per il Teorema di completamente a base posso completarla a base $\mathcal{C} = \{w_1, \dots, w_n\}$ di $\mathbb{R}^n$. Applicando il procedimento di Gram-Schimdt alla base $\mathcal{C} = \{w_1, \dots, w_n\}$, ottengo una base ortogonale di $\mathbb{R}^n$ che indicheremo con $\mathcal{B} = \{v_1, \dots, v_n\}$. Poiché

$$\mathcal{L}(v_1, \dots, v_k) = \mathcal{L}(w_1, \dots, w_k),$$

i vettori $v_1, \dots, v_k$ formano una base ortogonale di W. I vettori $v_{k+1}, \dots, v_n$ essendo ortogonali a una base di W appartengono a $W^\perp$. Adesso dimostriamo che i vettori $v_{k+1}, \dots, v_n$ formano una base ortogonale di $W^\perp$.

Sia $z \in W^\perp$. Poiché $\mathcal{B} = \{v_1, \dots, v_n\}$ è una base ortogonale di $\mathbb{R}^n$, si ha

$$z = \frac{\langle z, v_1\rangle}{\langle v_1, v_1\rangle} v_1 + \cdots + \frac{\langle z, v_n\rangle}{\langle v_n, v_n\rangle} v_n.$$

Tuttavia $z \in W^\perp$ e $W = \mathcal{L}(v_1, \dots, v_k)$, per cui si ha

$$\frac{\langle z, v_j\rangle}{\langle v_j, v_j\rangle} = 0, \text{ per } j = 1, \dots, k,$$

ovvero $z \in \mathcal{L}(v_{k+1}, \dots, v_n)$. Abbiamo quindi dimostrato che $W^\perp = \mathcal{L}(v_{k+1}, \dots, v_n)$ e $\dim W^\perp = n - k$. In particolare

$$\mathbb{R}^n = \mathcal{L}(v_1, \dots, v_n) = W + W^\perp,$$

e

$$n = \dim W + \dim W^\perp.$$

Applicando la formula di Grassmann, si ha $\dim(W \cap W^\perp) = 0$, i.e., $\mathbb{R}^n$ è in somma diretta di W e $W^\perp$. □

Definizione 8.18. *Sia W un sottospazio vettoriale dello spazio vettoriale $\mathbb{R}^n$. Sia $\mathcal{C} = \{w_1, \dots, w_k\}$ una base ortonormale di W. Si dice proiezione ortogonale di $\mathbb{R}^n$ su W l'applicazione*

$$P_W(v) = \langle v, w_1\rangle w_1 + \cdots + \langle v, w_k\rangle w_k.$$

Vogliamo dimostrare che $\operatorname{Ker} P_W = W^\perp$ e $\operatorname{Im} P_W = W$.

$v \in \operatorname{Ker} P_W$ se e solamente se $P_W(v) = \langle v, w_1\rangle w_1 + \cdots + \langle v, w_k\rangle w_k = 0$. 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 $\langle v, w_1\rangle = \cdots = \langle v, w_k\rangle = 0$ se e solamente se $w \in W^\perp$.

Poiché $\operatorname{Im} P_W \subset W$ (perché?) e $\dim \operatorname{Im} P_W = n - \dim W^\perp = \dim W$ si ha $\operatorname{Im} P_W = W$.

Il prossimo risultato garantisce che $P_W(v)$ non dipende dalla base ortonormale scelta di W.

Proposizione 8.19. *Sia $v \in \mathbb{R}^n$. Allora esiste un unico $w \in W$ tale che $v - w \in W^\perp$. Inoltre, se $\mathcal{C} = \{w_1, \ldots, w_k\}$ 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 $\mathcal{C} = \{w_1, \ldots, w_k\}$ 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 $j = 1, \ldots, k$, 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. $\square$

8.20 Prodotto Hermitiano canonica 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 8.21. *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* $Z = 0$;

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. $\square$

Sia X un vettore. Chiameremo *norma* o *lunghezza* di X il numero reale $\| X \| = \sqrt{\langle X, X\rangle}$. Dunque se $X \neq 0$, 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 8.22 (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 linearemente dipendenti.*

Facoltativa. Siano $\alpha, \beta \in \mathbb{C}$. Se $Z = 0$, allora la disuguaglianza è banalmente verificata. Supponiamo che $Z \neq 0$. 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 $\alpha Z + \beta W = 0$. $\square$

Definizione 8.23. *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 vale il seguente risultato.

Proposizione 8.24. *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 $\mathcal{B} = \{v_1, \ldots, v_n\}$ si dice ortogonale, rispettivamente ortonormale, se i vettori $v_1, \ldots, v_n$ 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 8.25 (Procedimento di ortogonalizzazione di Gram-Schmidt)**.** *Siano $v_1, \ldots, v_k \in V$ 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 $\mathcal{L}(w_1, \ldots, w_j) = \mathcal{L}(v_1, \ldots, v_j)$ per $j = 1, \ldots, k$.

Proposizione 8.26. *Sia $A = (A^1, \ldots, A^n) \in \mathrm{M}_{n \times n}(\mathbb{C})$. La matrice A è unitaria se e solamente se $\{A^1, \ldots, A^n\}$ 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 8.27. *$S^\perp$ è un sottospazio vettoriale di $\mathbb{C}^n$. Se $S = \mathcal{L}(v_1, \ldots, v_k)$, allora*

$$S^\perp = \{v \in \mathbb{C}^n : \langle v, v_1 \rangle = \cdots = \langle v, v_k \rangle = 0\}$$

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

$$V_\lambda = \{v \in V : T(v) = \lambda v\} = \{v \in V : (T - \lambda \mathrm{Id}_V)(v) = 0\}.$$

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 $V_\lambda \neq \{0\}$ se e solamente se λ è un autovalore. In tal caso

$$V_\lambda = \big\{v \in V : v \text{ è un autovettore di } T \text{ relativo a } \lambda\big\} \cup \{0\}.$$

che chiameremo *autospazio relativo a* λ.

Proposizione 9.4. *Uno scalare $\lambda \in \mathbb{K}$ è un autovalore di T se e solamente se $T - \lambda Id$ non è invertibile. 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 $V_\lambda \neq \{0\}$. Poiché $V_\lambda = \operatorname{Ker}(T - \lambda \mathrm{Id}_V)$, si ha che $\lambda \in \mathbb{K}$ è un autovalore se e solamente se $T - \lambda \mathrm{Id}_V$ non è iniettiva e quindi non è invertibile. □

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 $T(v_1) = \lambda_1 v_1, \ldots, T(v_n) = \lambda_n v_n$. Tenendo in mente che $[v_i]_= e_i$ dove $e_1, \ldots, e_n$ è la base canonica di $\mathbb{K}^n$, si ha che

$$[T(v_i)]_{\mathcal{B}} = \lambda_i [v_i]_{\mathcal{B}} = \lambda_i e_i,$$

ovvero 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 la matrice

$$\mathcal{M}_{\mathcal{B},\mathcal{B}}(T) = \begin{pmatrix} \lambda_1 & & & \\ & \ddots & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix},$$

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 $T : V \longrightarrow V$. 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.*

Dimostrazione. 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.

Supponiamo che la proposizione sia vera per m autovettori di T corrispondeti a m autovalori distinti e dimostriamolo per $m + 1$.

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. Allora v_{m+1} sarebbe combinazione lineare di $v_1, \ldots, v_m$ (perché?), ovvero

$$v_{m+1} = \alpha_1 v_1 + \cdots + \alpha_m v_m.$$

Applicando T si ha

$$T(v_{m+1}) = \lambda_{m+1} v_{m+1} = T(\alpha_1 v_1 + \cdots + \alpha_m v_m) = \lambda_1 \alpha_1 v_1 + \cdots + \lambda_m \alpha_m v_m.$$

Dall'altro lato

$$\lambda_m v_{m+1} = \lambda_m \alpha_1 v_1 + \cdots + \lambda_m \alpha_m v_m.$$

Quindi

$$\lambda_m\alpha_1 v_1 + \cdots + \lambda_m\alpha_m v_m = \lambda_1\alpha_1 v_1 + \cdots + \lambda_m\alpha_m v_m,$$

ovvero

$$(\lambda_{m+1} - \lambda_1)\alpha_1 v_1 + \cdots + (\lambda_{m+1} - \lambda_m)\alpha_m v_m = 0.$$

Poiché i vettori $v_1, \ldots, v_m$ sono linearmente indipendenti, ne segue che i coefficienti sono tutti nulli:

$$\alpha_1(\lambda_{m+1} - \lambda_1) = \ldots = \alpha_m(\lambda_{m+1} - \lambda_m) = 0$$

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 $v_1, \ldots, v_{m+1}$ sono linearmente indipendenti. $\square$

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 $v_1, \ldots, v_n$ corrispondenti ad autovalori distinti. Per il risultato anteriore $v_1, \ldots, v_n$ sono linearmente indipendenti e quindi formano una base di V concludendo la dimostrazione. $\square$

9.10 Polinomio caratteristico

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 $T - \lambda Id$ non è invertibile ovvero se e solamente se la matrice $\mathcal{M}_{\mathcal{B},\mathcal{B}}(T - \lambda Id)$ non è invertibile. Poiché $\mathcal{M}_{\mathcal{B},\mathcal{B}}(T - \lambda Id) = \mathcal{M}_{\mathcal{B},\mathcal{B}}(T) - \lambda Id$, si ha che $\lambda \in \mathbb{K}$ è un autovalore di T se e solamente se

$$det(\mathcal{M}_{\mathcal{B},\mathcal{B}}(T) - \lambda Id_n) = 0.$$

Quindi gli autovalori di un endomorfismo T sono le radice del polinomio $p_{\mathcal{B}}(t) = det(\mathcal{M}_{\mathcal{B},\mathcal{B}}(T) - tId_n)$.

Proposizione 9.11. *Il polinomio $p_{\mathcal{B}}(t)$ ha grado n e non dipende dalla scelta della base $\mathcal{B}$. Si indicherà con $p_T(t)$ e verrà chiamato* il polinomio caratteristico di T. *Inoltre*

$$p_T(t) = (-1)^n t^n + (-1)^{n-1}(\mathrm{Tr}(\mathcal{M}_{\mathcal{B},\mathcal{B}}(T))t^{n-1} + \cdots + det(\mathcal{M}_{\mathcal{B},\mathcal{B}}(T)).$$

Dimostrazione. Sia $\mathcal{C} = \{w_1, \dots, w_n\}$ 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

$$\begin{aligned}
det(\mathcal{M}_{\mathcal{C},\mathcal{C}}(T) - tId_n) &= det\left(\mathcal{M}(\mathcal{B},\mathcal{C})^{-1}\mathcal{M}_{\mathcal{B},\mathcal{B}}(T)\mathcal{M}(\mathcal{B},\mathcal{C}) - tId\right)\\
&= det\left(\mathcal{M}(\mathcal{B},\mathcal{C})^{-1}\mathcal{M}_{\mathcal{B},\mathcal{B}}(T)\mathcal{M}(\mathcal{B},\mathcal{C}) - t\mathcal{M}(\mathcal{B},\mathcal{C})^{-1}\mathcal{M}(\mathcal{B},\mathcal{C})\right)\\
&= det\left((\mathcal{M}(\mathcal{B},\mathcal{C})^{-1}(\mathcal{M}_{\mathcal{B},\mathcal{B}}(T) - tId)\mathcal{M}(\mathcal{B},\mathcal{C})\right)\\
&= det((\mathcal{M}(\mathcal{B},\mathcal{C})^{-1})det(\mathcal{M}_{\mathcal{B},\mathcal{B}}(T) - tId)det(\mathcal{M}(\mathcal{B},\mathcal{C}))\\
&= det(\mathcal{M}_{\mathcal{B},\mathcal{B}}(T) - tId).
\end{aligned}$$

La seconda parte della dimostrazione è omessa. □

Poiché gli autovalori di T sono le radici del polinomio caratteristico che ha grado $n = \dim V$, si ha il seguente risultato.

Corollario 9.12. *Sia $T : V \longrightarrow V$ un endomorfismo di V. Allora T ha al massimo $n = \dim V$ autovalori distinti.*

9.13 Molteplicità geometrica e algebrica

Sia $T : V \longrightarrow V$ un endomorfismo.

Definizione 9.14. *Sia $\lambda \in sp(T)$. 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 9.15. *Sia $T : V \longrightarrow V$ un endomorfismo e sia $\lambda \in \mathbb{K}$ un autovalore. Sia $n = \dim V$ sia $\mathcal{B}$ una base di V. Allora*

$$m_g(\lambda) = n - \mathrm{rg}(\mathcal{M}_{\mathcal{B},\mathcal{B}}(T) - \lambda Id_n).$$

Dimostrazione. Poiché $V_\lambda = \operatorname{Ker}(T - \lambda Id)$, applicando il teorema della dimensione si ha

$$\begin{aligned} m_g(\lambda) &= n - \dim \operatorname{Im}(T - \lambda Id) \\ &= n - \operatorname{rg}(\mathcal{M}_{\mathcal{B},\mathcal{B}}(T - \lambda Id)) \\ &= n - \operatorname{rg}(\mathcal{M}_{\mathcal{B},\mathcal{B}}(T) - \lambda \mathrm{Id}_n). \end{aligned}$$

□

Il prossimo risultato garantisce che la maolteplicità algebrica è maggiore oppure uguale della molteplicità geometrica.

Proposizione 9.16. *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

$$\begin{aligned} p_T(t) &= \det(\mathcal{M}_{\mathcal{B},\mathcal{B}}(T) - t\mathrm{Id}_n) \\ &= (\lambda - t)^{m_g(\lambda)} \det(D - tId_{(n-m_g(\lambda)) \times (n - m_g(\lambda))}), \end{aligned}$$

da cui segue che $m_g(\lambda) \leq m_a(\lambda)$. □

Corollario 9.17. *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. □

Il prossimo Teorema fornisce un criterio necessario e sufficiente affinché un endomorfismo sia diagonalizzabile.

Teorema 9.18. *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 $\mathcal{B} = \{v_1, \ldots, v_n\}$ una base formata da autovettori di T. Siano $\lambda_1, \ldots, \lambda_r$ gli autovalori distinti di T. La matrice associata a T rispetto alla base $\mathcal{B} = \{v_1, \ldots, v_n\}$ in partenza ed in arrivo ha la forma

$$\mathcal{M}_{\mathcal{B},\mathcal{B}}(T) = \begin{pmatrix} \lambda_1 Id_{m_g(\lambda_1)\times m_g(\lambda_1)} & & \\ & \ddots & \\ & & \lambda_k Id_{m_g(\lambda_k)\times m_g(\lambda_k)} \end{pmatrix}.$$

e quindi

$$\begin{aligned} p_T(t) &= \det(\mathcal{M}_{\mathcal{B},\mathcal{B}}(T) - tId) \\ &= \det \begin{pmatrix} (\lambda_1 - t) Id_{m_g(\lambda_1)\times m_g(\lambda_1)} & & \\ & \ddots & \\ & & (\lambda_k - t) Id_{m_g(\lambda_k)\times m_g(\lambda_k)} \end{pmatrix} \\ &= (\lambda_1 - t)^{m_g(\lambda_1)} \cdots (\lambda_k - t)^{m_g(\lambda_k)}. \end{aligned}$$

da cui segue che tutti gli autovalori stanno in $\mathbb{K}$ ed $m_a(\lambda_j) = m_g(\lambda_j)$ per $j = 1, \ldots, k$.

Viceversa, poiché T ha tutti gli autovalori in $\mathbb{K}$, T ha n autovalori non necessariamente distinti essendo tutte le radici di un polinomio di grado n.

Siano $\lambda_1, \ldots, \lambda_k \in \mathbb{K}$ gli autovalori distinti di T. Sia $\mathcal{B}_{V_{\lambda_j}}$ una base di V_{λ_j} per $j = 1, \ldots, k$. L'insieme $\mathcal{B} = \mathcal{B}_{V_{\lambda_1}} \bigcup \cdots \bigcup \mathcal{B}_{V_{\lambda_k}}$ è formato da $n = \dim V$ autovettori. Infatti, poiché il polinomio caratteristico ha n autovalori allora $n = \sum_{i=m}^k m_a(\lambda_m)$. Infine, la molteplicitità algebrica di ogni autovalore è uguale alla molteplicità geometrica, da cui segue

$$n = \sum_{m=1}^k m_a(\lambda_m) = \sum_{m=1}^k m_g(\lambda_m),$$

cioé $\mathcal{B}$ ha esattamente $n = \dim V$ autovettori. Per concludere la dimostrazione, tenendo in mente che se $n = \dim V$ vettori formano una base di V se e solamente se sono linearmente indipendenti, è suffciente dimostrare che i vettori di $\mathcal{B}$ sono linearmente indipendenti. A meno di riordianre i vettori di $\mathcal{B}$, esistono $1 \leq j_1 < j_2 < \cdots < j_k \leq n =$ tale che

$$\begin{array}{ll} v_1, \dots, v_{j_1} & \text{formano una base di } V_{\lambda_1} \\ v_{j_1+1}, \dots, v_{j_2} & \text{formano una base di } V_{\lambda_2} \\ \vdots & \\ v_{j_k+1}, \dots, v_n & \text{formano una base di } V_{\lambda_k} \end{array}$$

Siano $\alpha_1, \dots, \alpha_{j_1}, \dots, \alpha_{j_k}, \dots, \alpha_n \in \mathbb{K}$ e sia $\alpha_1 v_1 + \cdots + \alpha_n v_n = 0$ 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

$$w_1 + \cdots + w_k = 0$$

Per la Proposizione 9.8, necessariamente si ha $w_1 = \cdots = w_k = 0$ e quindi

$$\begin{array}{l} 0 = \alpha_1 v_1 + \cdots + \alpha_{j_1} v_{j_1} \\ 0 = \alpha_{j_1+1} v_{j_1+1} + \cdots + \alpha_{j_2} v_{j_2} \\ \vdots \\ 0 = \alpha_{j_k+1} v_{j_k+1} + \cdots + \alpha_n v_n. \end{array}$$

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 $\alpha_1 = \cdots = \alpha_n = 0$. □

Osservazione 9.19. *Sia T un endomorfismo diagonalizzabile e siano $\lambda_1, \ldots, \lambda_k$ gli autovalori distinti di T. Sia $\mathcal{B}_{V_{\lambda_j}}$ una base di V_{λ_j} per $j = 1, \ldots, k$. Nella proposizione anteriore abbiamo dimostrato che $\mathcal{B} = \mathcal{B}_{V_{\lambda_1}} \bigcup \cdots \bigcup \mathcal{B}_{V_{\lambda_k}}$ è una base di V formata da autovettori di T. Inoltre*

$$\mathcal{M}_{\mathcal{B},\mathcal{B}}(T) = \begin{pmatrix} \lambda_1 \mathrm{Id}_{m_g(\lambda_1)\times m_g(\lambda_1)} & & \\ & \ddots & \\ & & \lambda_k \mathrm{Id}_{m_g(\lambda_k)\times m_g(\lambda_k)} \end{pmatrix}$$

9.20 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. Poiché $L_A(X) = AX$, definiamo autovettori, rispettivamente autovalori di A, gli autovettori, rispettivamente 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 $Av = \lambda v$. Quindi $\lambda \in \mathbb{K}$ è una autovalore di A se e solamente se $\det(A - \lambda \mathrm{Id}_n) = 0$. Chiamaremo $p_A(t) = \det(A - tId)$ il *polinomio caratteristico* di A. Poiché $\mathcal{M}_{\mathcal{C},\mathcal{C}}(L_A) = A$, dove $\mathcal{C}$ è la base canonica di $\mathbb{K}^n$, il polinomio caratteristico di A è il polinomio caratteristico di L_A. Analogamento possiamo definire l'autospazio relativo all'autovalore λ essendo

$$V_\lambda = \{v \in \mathbb{K}^n : \; L_A(v) = Av = \lambda v\} = \mathrm{Sol}(A - \lambda Id|0).$$

La dimensione di V_λ è la molteplicità geometrica di λ. Inoltre $m_g(\lambda) = n - \mathrm{rg}(A - \lambda Id)$. 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 $p_A(t)$.

Definizione 9.21. *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.*

Dalla definizione segue che una matrice quadrata A è diagonalizzabile se e solamente se A è simile a una matrice diagonale.

Proposizione 9.22. *A è diagonalizzabile se e solamente se l'endomorfismo $L_A : \mathbb{K}^n \longrightarrow \mathbb{K}^n$ è diagonalizzabile.*

Dimostrazione. $\mathcal{M}_{\mathcal{C},\mathcal{C}}(L_A) = A$, dove $\mathcal{C}$ è la base canonica. Quindi L_A è diagonalizzabile se e solamente se A è simile a una matrice diagonale.

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 gli autovettori e gli autovalori dell'endomorfismo L_A.

Supponiamo che A sia diagonalizzabile. Quindi esiste una matrice invertibile P e una matrice diagonale D tale che $P^{-1}AP = D$. In particolare $AP = PD$. Le colonne di P, ovvero i vettori $\{P^1, \ldots, P^n\}$, formano una base di $\mathbb{K}^n$. Inoltre

$$L_A(P^i) = AP^i = (AP)^i = (PD)^i = d_{ii}P^i.$$

Quindi P^i è una autovettore di L_A relativo all'autovalore d_{ii}, ovvero $P^1, \ldots, P^n$ formano una base di $\mathbb{R}^n$ formata da autovettori di L_A.

Viceversa supponiamo che L_A sia diagonalizzabile. Sia $\mathcal{B} = \{v_1, \ldots, v_n\}$ una base di $\mathbb{K}^n$ formata da autovettori di L_A. La matrice $P = (v_1, \ldots, v_n)$ è invertibile poiché $\mathcal{B} = \{v_1, \ldots, v_n\}$ è una base di $\mathbb{K}^n$. Vogliamo dimostrare che $P^{-1}AP = D$ è una matrice diagonale. Infatti, se indichiamo con $\mathcal{C}$ la base canonica di $\mathbb{K}^n$, si ha

$$P^{-1}AP = \mathcal{M}(\mathcal{B}, \mathcal{C})\mathcal{M}_{\mathcal{C},\mathcal{C}}(L_A)\mathcal{M}(\mathcal{C}, \mathcal{B}) = \mathcal{M}_{\mathcal{B},\mathcal{B}}(L_A).$$

Poiché $\mathcal{B} = \{v_1, \ldots, v_n\}$ è una base di $\mathbb{K}^n$ formata da autovettori di L_A, la matrice $\mathcal{M}_{\mathcal{B},\mathcal{B}}(L_A)$ è diagonale. Quindi A è simile a una matrice diagonale. Infine, se indichiamo con $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$ gli autovalori di L_A corrispondenti agli autovettori $\mathcal{B} = \{v_1, \ldots, v_n\}$, i.e., $L_A(v_i) = Av_i = \lambda_i v_i$ per $i = 1, \ldots, n$, allora

$$\mathcal{M}_{\mathcal{B},\mathcal{B}}(L_A) = \begin{pmatrix} \lambda_1 & & & \\ & \ddots & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}.$$

□

Applicando il Teorema 9.18 e la Proposizione 9.22 si ha il seguente criterio di diagonalizzazione di matrici.

Teorema 9.23. *Sia $A \in M_{n\times n}(\mathbb{K})$. Le seguenti condizioni sono equivalenti:*

a) A è diagonalizzabile;

b) tutti gli autovalori di A sono in $\mathbb{K}$ e per ogni autovalore $\lambda \in sp(T)$, si ha $m_a(\lambda) = m_g(\lambda)$;

Osservazione 9.24.

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 $\lambda_1, \ldots, \lambda_k \in \mathbb{C}$ 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 radici di $p_A(t)$. Infatti, poiché $p_A(t)$ è un polinomio a coefficienti reali, si ha $\overline{p_A(t)} = p_A(\bar{t})$ 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*

$$\det(A) = \lambda_1 \cdots \lambda_n \qquad \mathrm{Tr}(A) = \lambda_1 + \cdots + \lambda_n.$$

9.25 Metodi di calcolo

Vediamo quali sono i passi per dimostrare che una matrice quadrata $A \in \mathrm{M}_{n\times n}(\mathbb{K})$ è diagonalizzabile oppure ed in caso affermatico come determinare una matrice invertibile P e una matrice diagonale D tale che $P^{-1}AP = D$.

a) Gli autovalori della matrice A sono le radici del polinomio caratteristico $p_A(t) = \det(A - t\mathrm{Id})$. Se le radici del polinomio caratteristico appartengono a $\mathbb{K}$, allora continuo. Altrimenti la matrice A non è diagonalizzabile.

b) Sia $\lambda \in \mathbb{K}$ un autovalore. Allora $V_\lambda = \mathrm{Sol}(A - \lambda Id|0)$ e $m_g(\lambda) = \dim V_\lambda = n - rg(A - \lambda\mathrm{Id})$. La molteplicità algebrica di λ è il numero di volte che λ è radice. Se $m_a(\lambda) = m_g(\lambda)$ per ogni autovalore λ allora la matrice è diagonalizzabile. Altrimenti no.

Supponiamo che A sia diagonalizzabile. Se $\lambda \in \mathbb{K}$ è un autovalore di A, allora indicheremo con $\mathcal{B}_\lambda$ una base di $V_\lambda = \mathrm{Sol}(A - \lambda Id|0)$.

Siano $\lambda_1, \ldots, \lambda_s \in \mathbb{K}$ autovalori distinti di A. Allora

$$\mathcal{B} = \mathcal{B}_{\lambda_1} \cup \cdots \cup \mathcal{B}_{\lambda_s} = (v_1, \ldots, v_n)$$

è una base di $\mathbb{K}^n$ formata da autovettori di A. Sia $P = (v_1, \ldots, v_n)$ la matrice le cui colonne sono gli autovettori $v_1, \ldots, v_n$. Allora

$$P^{-1}AP = D = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

è una matrice diagonale e gli scalari $\lambda_1, \ldots, \lambda_n$ sono gli autovalori di A corrispondenti agli autovettori $v_1, \ldots, v_n$. Dimostreremo che le due matrici hanno lo stesse colonne.

Siano $e_1, \ldots, e_n$ la base canonica di $\mathbb{K}^n$. Allora $Pe_i = v_i$. Poiché P è invertibile si ha $P^{-1}v_i = e_i$. Quindi

$$P^{-1}APe_i = P^{-1}Av_i = \lambda_i P^{-1}v_i = \lambda_i e_i,$$

ovvero $P^{-1}AP = D$.

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

$$\left\{\mathcal{F}_{\mathcal{B}}^{-1}(w_1) =, \ldots, \mathcal{F}_{\mathcal{B}}^{-1}(w_n)\right\} = \{a_{11}v_1 + \cdots + a_{n1}v_n, \ldots, a_{1n}v_1 + \cdots + a_{nn}v_n\}.$$

9.26 Teorema spettrale

Vogliamo dimostrare che una matrice a coefficienti reali simmetrica è sempre diagonalizzabile. Non solo, possiamo trovare una base ortonormale formata da autovettori di A. Il primo risultato che dimostreremo è che gli autovalori di una matrice simmetrica sono numeri reali.

Lemma 9.27. *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 $\mathbb{C}^n$. Per ogni $X, Y \in \mathbb{C}^n$, si ha

$$\langle L_A(X), Y\rangle = \langle X, L_A(Y)\rangle.$$

Infatti

$$\begin{aligned} \langle L_A(X), Y\rangle &= (AX)^T\overline{Y} \\ &= X^T A^T \overline{Y} \\ &= X^T \overline{AY} \\ &= \langle X, AY\rangle \\ &= \langle X, L_A(Y)\rangle. \end{aligned}$$

Sia X un autovettore di L_A relativo all'autovalore λ. Allora

$$\langle L_A X, X\rangle = \langle AX, X\rangle = \lambda\langle X, X\rangle$$

Per la formula anteriore otteniamo che

$$\langle L_A(X), X\rangle = \langle X, L_A(X)\rangle = \langle X, AX\rangle = \overline{\lambda}\langle X, X\rangle.$$

Quindi

$$\lambda\langle X, X\rangle = \overline{\lambda}\langle X, X\rangle.$$

Poiché $\langle X, X\rangle \neq 0$ si ha $\lambda = \overline{\lambda}$, ovvero $\lambda \in \mathbb{R}$. $\square$

La seguente proprosizione garatisce che autospazi corrispondenti a autovettori distinti sono ortogonali.

Proposizione 9.28. *Sia A una matrice simmetrica di ordine n e siano v e w autovettori corrispondenti 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. Se A è simmetrica vale la seguente formula:

$$\langle AX, Y\rangle = \langle X, AY\rangle,$$

dove $\langle\cdot,\cdot\rangle$ è il prodotto scalare canonico di $\mathbb{R}^n$. 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 relativi agli autovalori λ e μ rispettivamente. Allora

$$\begin{aligned}\langle Av, w\rangle &= \lambda\langle v, w\rangle \\ &= \langle v, Aw\rangle \text{ (formula anteriore)} \\ &= \mu\langle v, w\rangle.\end{aligned}$$

Quindi

$$\lambda\langle v, w\rangle = \mu\langle v, w\rangle \Rightarrow (\mu - \lambda)\langle v, w\rangle = 0,$$

da cui segue, essendo λ e μ distinti, $\langle v, w\rangle = 0$. $\square$

Teorema 9.29 (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 $\mathcal{B} = \{v_1, \ldots, v_n\}$ 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 9.30. *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 9.31. *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.*

9.32 Metodi di calcolo

Sia $A \in \mathrm{A}_{n\times n}(\mathbb{R})$ una matrice simmetrica. Il teorema spettrale garantisce che A è diagonalizzabile e che gli autospazi sono a due a due ortogonali fra loro. Per calcolare una matrice P ortogonale tale che $P^T AP = D$ si può procedere come segue.

Sia $\mathcal{B} = \{v_1, \ldots, v_n\}$ una base di $\mathbb{R}^n$ formata da autovettori di A che può essere calcolata come nella sezione 9.25. Se indichiamo con $Q = (v_1, \ldots, v_n)$, allora $Q^{-1}AQ = D$ è diagonale. Applicando il procedimento di Gram-Schimdt alla base $\mathcal{B} = \{v_1, \ldots, v_n\}$, e poi dividendo ciascun vettore per la sua norma, ottengo una base ortonormale $\mathcal{C} = \{w_1, \ldots, w_n\}$ di $\mathbb{R}^n$. Poiché gli autospazi della matrice A sono a due a due ortogonali, la base $\mathcal{C} = \{w_1, \ldots, w_n\}$ è ancora una base formata da autovettori di A (SE A NON È SIMMETRICA TALE AFFERMAZIONE È FALSA!!!!!!!). La matrice $P = (w_1, \ldots, w_n)$ è una matrice ortogonale le cui colonne formano una base ortonormale di autovettori di A. Quindi $P^{-1}AP = P^T AP = D$ (D è la stessa matrice diagonale ottenuta in precedenza. (Perché?))

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. Piché $Ae_i = A^i$, 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. $\square$

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 $\langle \cdot, \cdot \rangle$ il prodotto Hermitiano canonico. Dimostriamo che per ogni $Z, W \in \mathbb{C}^n$ si ha

$$\langle AZ, AW \rangle = \langle Z, W \rangle.$$

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

$$|\lambda|^2 \langle Z, Z \rangle = |\lambda|^2 \langle Z, Z \rangle.$$

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 un isometria 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 $A^T = A^{-1}$ e quindi W è A^T invariante. Quindi $W^\perp$ 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 $\theta, \psi \in [0, \pi]$ tale che

$$\begin{matrix} a = \cos\theta & c = \sin\theta \\ b = \cos\psi & d = \sin\psi \end{matrix}.$$

Inoltre dalla secondo equazione si deduce che $\cos\theta\cos\psi + \sin\theta\sin\psi = \cos(\theta - \psi) = 0$, 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

$$A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \qquad (\det(A) = 1);$$

se k è dispari, allora $\cos\psi = \cos(\theta + \frac{3\pi}{2}) = \sin(\theta)$ e $\sin\psi = \sin(\theta + \frac{3\pi}{2}) = -\cos(\theta)$

$$A = \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix} \qquad (\det(A) = -1).$$

Se k è dispari, allora

$$A = \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix}.$$

A è simmetrica e quindi diagonalizzabile. Gli autovalori 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

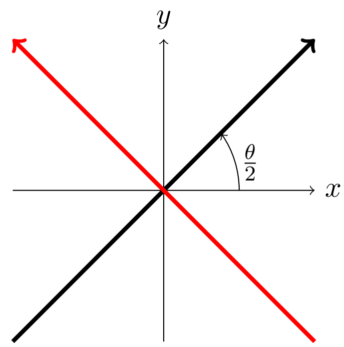
$$V_1 = \mathcal{L}\Big(\begin{bmatrix} \cos\frac{\theta}{2} \\ \sin\frac{\theta}{2} \end{bmatrix}\Big).$$

Poiché $V_{-1} = V_1^{\perp}$, si ha

$$V_{-1} = \mathcal{L}\Big(\begin{bmatrix} -\sin\frac{\theta}{2} \\ \cos\frac{\theta}{2} \end{bmatrix}\Big).$$

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

$$\mathcal{M}_{\mathcal{B},\mathcal{B}}(L_A) = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix},$$



i.e., L_A è la riflessione rispetto alla retta passante per l'origine e che ha $\begin{bmatrix} \cos\frac{\theta}{2} \\ \sin\frac{\theta}{2} \end{bmatrix}$. come vettore direttore.
Se k è pari, allora $p_A(t) = t^2 - 2\cos\theta + 1$. Quindi L_A non è diagonalizzabile se $\theta \neq 0, -\pi$. Osserviamo 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 $\mathbb{R}^2$ da cui segue se $v = \begin{bmatrix} x \\ y \end{bmatrix}$, allora

$$\begin{aligned} \langle Av, v \rangle &= x^2\langle Ae_1, e_1 \rangle + xy\langle Ae_1, e_2 \rangle + xy\langle Ae_2, e_1 \rangle + y^2\langle Ae_2, e_2 \rangle \\ &= (x^2 + y^2)\cos\theta + xy\sin\theta - xy\sin\theta \\ &= \parallel v \parallel^2 \cos\theta. \end{aligned}$$

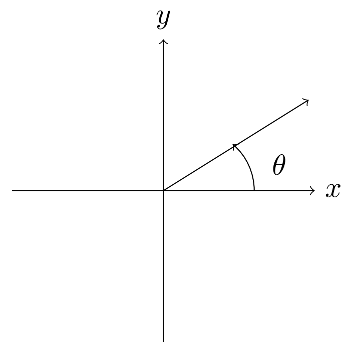
Allora

$$\frac{\langle Av, v \rangle}{\parallel v \parallel \parallel Av \parallel} = \cos\theta,$$

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. $\square$

10.11 Matrici ortogonali di formato 3×3

Sia $A \in \mathrm{M}_{3\times 3}(\mathbb{R})$ ortogonale.

Lemma 10.12. *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 $\det(-A) = (-1)^3 \det(A) = -1$, ovvero $\det(A - \mathrm{Id}_3) = 0$. 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 10.13. *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 10.14. *$T : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ 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 $\tilde{A}$ formana 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 una 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 $\det(A) = \det(\mathcal{M}_{\mathcal{B},\mathcal{B}}(T)) = \det(\tilde{A}) = 1$. □