a.a. 2019/2020

Algebra lineare

UNIVERSITÀ DEGLI STUDI DI MILANO BICOCCA

Indice

Capitolo 1

Spazi vettoriali

1.1 Lo spazio $\mathbb{R}^n$

Per $n \geq 1$ indichiamo con $\mathbb{R}^n$ l'insieme delle n-uple ordinate di numeri reali, ovvero

$$\mathbb{R}^n = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} : x_i \in \mathbb{R} \right\}.$$

Un elemento $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ in $\mathbb{R}^n$ si dice vettore, i numeri reali x_i, $i = 1, ..., n$ si dicono componenti del vettore.

Somma e prodotto per uno scalare in $\mathbb{R}^n$.

Definiamo in $\mathbb{R}^n$ due operazioni, la somma (di vettori) e la moltiplicazione per uno scalare (tra un numero reale ed un vettore) come segue:

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{pmatrix}$$

$$\alpha \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \alpha x_1 \\ \alpha x_2 \\ \vdots \\ \alpha x_n \end{pmatrix}$$

per $\alpha \in \mathbb{R}$, $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{R}^n$.

Per brevità scriveremo $\mathbf{x}$ per il vettore $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$, $\mathbf{y}$ per il vettore $\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$, $\mathbf{z}$ per il vettore $\begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix}$. La somma e il prodotto per uno scalare in $\mathbb{R}^n$ soddisfano le seguenti proprietà.

1. $\forall\, \mathbf{x}, \mathbf{y} \in \mathbb{R}^n, \quad \mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$
2. $\forall\, \mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n, \quad (\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{x} + (\mathbf{y} + \mathbf{z})$
3. esiste un vettore (che si dice vettore nullo e si indica con $\mathbf{0}_{\mathbb{R}^n}$) tale che, $\forall\, \mathbf{x} \in \mathbb{R}^n$, $\mathbf{x} + \mathbf{0}_{\mathbb{R}^n} = \mathbf{x}$. Basta porre $\mathbf{0}_{\mathbb{R}^n} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$.
4. $\forall\, \mathbf{x} \in \mathbb{R}^n$ esiste $-\mathbf{x} \in \mathbb{R}^n$ tale che $\mathbf{x} + (-\mathbf{x}) = \mathbf{0}_{\mathbb{R}^n}$. Basta porre
$$-\mathbf{x} = \begin{pmatrix} -x_1 \\ -x_2 \\ \vdots \\ -x_n \end{pmatrix}.$$
5. $\forall\, \alpha,\ \beta \in \mathbb{R},\ \mathbf{x} \in \mathbb{R}^n, \quad \alpha(\beta\mathbf{x}) = (\alpha\beta)\mathbf{x}$
6. $\forall\, \mathbf{x} \in \mathbb{R}^n, \quad 1\mathbf{x} = \mathbf{x}$
7. $\forall \alpha \in \mathbb{R},\ \mathbf{x}, \mathbf{y} \in \mathbb{R}^n, \quad \alpha(\mathbf{x} + \mathbf{y}) = \alpha\mathbf{x} + \alpha\mathbf{y}$
8. $\forall \alpha, \beta \in \mathbb{R},\ \mathbf{x} \in \mathbb{R}^n, \quad (\alpha + \beta)\mathbf{x} = \alpha\mathbf{x} + \beta\mathbf{x}$

1.2 Spazi vettoriali su $\mathbb{R}$

Definizione 1.1. *Si dice spazio vettoriale su $\mathbb{R}$ un insieme non vuoto V, i cui elementi si dicono vettori, con le due operazioni*

$$\begin{aligned} + : V \times V &\to V \quad \text{somma di vettori} \\ (v, u) &\mapsto v + u \\ \cdot : \mathbb{R} \times V &\to V \quad \text{moltiplicazione per uno scalare} \\ (\alpha, v) &\mapsto \alpha v \end{aligned}$$

tali che

1. $\forall\, u, v \in V, \quad u + v = v + u$
2. $\forall\, u, v, w \in V, \quad (u + v) + w = u + (v + w)$
3. *esiste un vettore (che si dice vettore nullo e si indica con 0_V) tale che, $\forall\, u \in V$, $u + 0_V = u$.*
4. $\forall\, u \in V$ *esiste* $-u \in V$ *tale che* $u + (-u) = 0_V$
5. $\forall\, \alpha,\, \beta \in \mathbb{R},\, u \in V, \quad \alpha(\beta u) = (\alpha\beta)u$
6. $\forall\, u \in V, \quad 1u = u$
7. $\forall \alpha \in \mathbb{R},\, u, v \in V, \quad \alpha(u + v) = \alpha u + \alpha v$
8. $\forall \alpha, \beta \in \mathbb{R},\, u \in V, \quad (\alpha + \beta)u = \alpha u + \beta u$

Esempi di spazi vettoriali su $\mathbb{R}$.

1. Per ogni $n \geq 1$, $\mathbb{R}^n$ è uno spazio vettoriale su $\mathbb{R}$ con le operazioni di somma e prodotto per uno scalare definite nella sezione 1.1. In particolare, per $n = 1$, $\mathbb{R}^1 = \mathbb{R}$ è uno spazio vettoriale su $\mathbb{R}$.

2. Sia $\mathbb{P}$ l'insieme dei polinomi a coefficienti reali nell'indeterminata t:
$$\mathbb{P} = \{p(t) = \; a_0 + a_1 t + a_2 t^2 + \cdots + a_n t^n |\; a_i \in \mathbb{R},\, i = 0, ..., n,\, n \in \mathbb{N}\}.$$
L'insieme $\mathbb{P}$ è uno spazio vettoriale su $\mathbb{R}$ con le operazioni di somma di polinomi e prodotto di un polinomio per un numero reale:
$$p(t) + q(t) = \sum_{k=0}^{n} (a_k + b_k) t^k, \quad (b_k = 0 \text{ per } k > m)$$
$$\alpha p(t) = \sum_{k=0}^{n} (\alpha a_k) t^k,$$
per $\alpha \in \mathbb{R}$, $p(t) = \sum_{k=0}^{n} a_k t^k$, $q(t) = \sum_{k=0}^{m} b_k t^k$, $n \geq m$.

3. Sia $\mathcal{F}(\mathbb{R})$ l'insieme delle funzioni reali di variabile reale
$$\mathcal{F}(\mathbb{R}) = \{f : \mathbb{R} \to \mathbb{R} \,|\; f \text{ funzione}\}.$$
Per $f, g \in \mathcal{F}(\mathbb{R})$, $\alpha \in \mathbb{R}$, definiamo
$$\begin{aligned} f + g \; &: \mathbb{R} \to \mathbb{R} \\ &(f + g)(x) = f(x) + g(x), \quad x \in \mathbb{R} \\ \alpha f \; &: \mathbb{R} \to \mathbb{R} \\ &(\alpha f)(x) = \alpha f(x), \quad x \in \mathbb{R} \end{aligned}$$
L'insieme $\mathcal{F}(\mathbb{R})$ è uno spazio vettoriale su $\mathbb{R}$ con le operazioni definite sopra. Il vettore nullo è la funzione identicamente nulla.

Osservazione 1.2. *Sia V uno spazio vettoriale su $\mathbb{R}$. Per ogni $\alpha \in \mathbb{R}$, $v \in V$, si ha*

(i) $\alpha\, 0_V = 0_V$

(ii) $0\, v = 0_V$

(iii) $(-\alpha)\, v = -(\alpha)\, v = \alpha\, (-v)$

(iv) *se* $\alpha\, v = 0_V$ *allora* $\alpha = 0$ *oppure* $v = 0_V$

Sottospazi

Definizione 1.3. *Sia V uno spazio vettoriale su $\mathbb{R}$. Un sottoinsieme W di V si dice un sottospazio di V se*

1. $W \neq \emptyset$
2. $\forall\, w_1, w_2 \in W,\ \forall\, \alpha, \beta \in \mathbb{R}, \quad \alpha w_1 + \beta w_2 \in W$

Osservazione 1.4. *In modo equivalente, abbiamo che un sottoinsieme W di V si dice un sottospazio di V se*

(i) $0_V \in W$

(ii) $\forall\, w_1, w_2 \in W, \quad w_1 + w_2 \in W$

(iii) $\forall\, w \in W,\ \forall\, \alpha \in \mathbb{R}, \quad \alpha w \in W.$

Esempi di sottospazi.

1. In $\mathbb{R}^3$ il sottoinsieme

$$W = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} : x_1 + x_2 + x_3 = 0 \right\}$$

è un sottospazio. Infatti

(i) $\mathbf{0}_{\mathbb{R}^3} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$, $0 + 0 + 0 = 0$

(ii) $\forall\, \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \in W$, $\mathbf{x} + \mathbf{y} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ x_3 + y_3 \end{pmatrix} \in W$. Infatti

$$(x_1 + y_1) + (x_2 + y_2) + (x_3 + y_3) = (x_1 + x_2 + x_3) + (y_1 + y_2 + y_3) = 0.$$

(iii) $\forall\, \alpha \in \mathbb{R}, \forall\, \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$, $\alpha\mathbf{x} = \begin{pmatrix} \alpha x_1 \\ \alpha x_2 \\ \alpha x_3 \end{pmatrix} \in W$. Infatti

$$(\alpha x_1) + (\alpha x_2) + (\alpha x_3) = \alpha(x_1 + x_2 + x_3) = 0.$$

2. Sia V uno spazio vettoriale su $\mathbb{R}$. Sono sottospazi di V l'insieme $\{0_V\}$ e V stesso.

3. Fissato $m \in \mathbb{N}$, sia $\mathbb{P}_m$ l'insieme dei polinomi a coefficienti reali, nell'indeterminata t, di grado minore o uguale a m:

$$\mathbb{P}_m = \{p(t) \in \mathbb{P} : \text{ grado}\{p(t)\} \leq m\}.$$

   $\mathbb{P}_m$ è un sottospazio di $\mathbb{P}$. Infatti

   (i) il polinomio nullo è in $\mathbb{P}_m$;

   (ii) $\forall\, p(t),\, q(t) \in \mathbb{P}_m$, $p(t)+q(t) \in \mathbb{P}_m$ perché la somma di polinomi di grado al più m è un polinomio di grado al più m;

   (iii) $\forall\, \alpha \in \mathbb{R}$, $\forall\, p(t) \in \mathbb{P}_m$, $\alpha\, p(t) \in \mathbb{P}_m$ perché $(\alpha p)(t)$ ha lo stesso grado di $p(t)$ se $\alpha \neq 0$ ed è il polinomio nullo se $\alpha = 0$.

Osservazione 1.5. *Se W è un sottospazio di V allora W è, a sua volta, uno spazio vettoriale rispetto alle stesse operazioni di V.*

Intersezione e somma di sottospazi.

Proposizione 1.6. *Sia V uno spazio vettoriale su R, e siano W e U due sottospazi di V. Allora l'insieme*

$$W \cap U = \{v \in V \mid v \in W \text{ e } v \in U\},$$

è un sottospazio di V.

Dim. Verifichiamo che $W \cap U$ è un sottospazio di V. Inoltre per ipotesi W e U sono sottospazi di V, ovvero

$$\begin{array}{lll} 0_V \in W, & w_1, w_2 \in W \Rightarrow w_1 + w_2 \in W, & \alpha \in \mathbb{R}, w \in W \Rightarrow \alpha w \in W, \\ 0_V \in U, & u_1, u_2 \in U \Rightarrow u_1 + u_2 \in U, & \alpha \in \mathbb{R}, u \in U \Rightarrow \alpha u \in U. \end{array}$$

Si ha

(i) $0_V \in W$, $0_V \in U$ dunque $0_V \in W \cap U$.

(ii) $\forall\, v_1, v_2 \in W \cap U$, $v_1 + v_2 \in W \cap U$. Infatti

$$v_1, v_2 \in W \cap U \Rightarrow v_1, v_2 \in W \Rightarrow v_1 + v_2 \in W$$
$$v_1, v_2 \in W \cap U \Rightarrow v_1, v_2 \in U \Rightarrow v_1 + v_2 \in U$$

(iii) $\forall\, \alpha \in \mathbb{R}$, $\forall\, v \in W \cap U$, $\alpha\, v \in W \cap U$. Infatti

$$\alpha \in \mathbb{R},\, v \in W \cap U \Rightarrow \alpha \in \mathbb{R},\, v \in W \Rightarrow \alpha v \in W$$
$$\alpha \in \mathbb{R},\, v \in W \cap U \Rightarrow \alpha \in \mathbb{R},\, v \in U \Rightarrow \alpha v \in U$$

$\square$

Osservazione 1.7. *La Proposizione 1.6 si generalizza immediatamente: se* V *è uno spazio vettoriale e* $\{W_i\}_{i=1,...,n}$ *è una famiglia di sottospazi di* V*, allora*

$$\cap_{i=1}^{n} W_i = \{w \in V \mid w \in W_i,\ i = 1, ..., n\}$$

è un sottospazio di V*.*

Proposizione 1.8. *Sia* V *uno spazio vettoriale su* $\mathbb{R}$ *e siano* W *e* U *due sottospazi di* V*. L'insieme*

$$U + W = \{v \in V \mid v = u + w \text{ con } u \in U,\, w \in W\}$$

è un sottospazio di V*.*

Dim. Verifichiamo che $U + W$ è un sottospazio di V.

(i) $0_V \in W$, $0_V \in U$ dunque $0_V \in U + W$. Infatti

$$0_V = 0_V + 0_V, \quad 0_V \in U,\ 0_V \in W.$$

(ii) $\forall\, v_1, v_2 \in U + W$, $v_1 + v_2 \in U + W$. Infatti

$$v_1 \in U + W \Rightarrow v_1 = u_1 + w_1,\ u_1 \in U,\, w_1 \in W$$
$$v_2 \in U + W \Rightarrow v_2 = u_2 + w_2,\ u_2 \in U,\, w_2 \in W$$

perciò

$$v_1 + v_2 = (u_1 + w_1) + (u_2 + w_2) = (u_1 + u_2) + (w_1 + w_2)$$

dove $u_1 + u_2 \in U$, $w_1 + w_2 \in W$.

(iii) $\forall\, \alpha \in \mathbb{R}$, $\forall\, v \in U + W$, $\alpha\, v \in U + W$. Infatti

$$v \in W + U \Rightarrow v = u + w,\ u \in U,\, w \in W$$

perciò

$$\alpha v = \alpha(u + w) = (\alpha u) + (\alpha w)$$

dove $\alpha u \in U$, $\alpha w \in W$.

□

Osservazione 1.9. *La Proposizione ?? si generalizza immediatamente: se* V *è uno spazio vettoriale e* $\{W_i\}_{i=1,...,n}$ *è una famiglia di sottospazi di* V*, allora*

$$\sum_{i=1}^{n} W_i = \{v \in V \mid v = w_1 + w_2 + \cdots + w_n,\ w_i \in W_i,\ i = 1, ..., n\}$$

è un sottospazio di V*.*

1.3 Combinazioni lineari ed indipendenza

Definizione 1.10. *Sia V uno spazio vettoriale su $\mathbb{R}$ e siano $v_1, v_2, ..., v_m$ vettori, $m \geq 1$. Si dice combinazione lineare di $v_1, v_2, ..., v_m$ un qualunque vettore della forma*

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_m v_m$$

con $\alpha_1, ..., \alpha_m \in \mathbb{R}$.

Osservazione 1.11. *(i) Se $V \neq \{0_V\}$ esistono infinite combinazioni lineari dei vettori $v_1, v_2, ..., v_m$.*

(ii) Se $\alpha_1 = \alpha_2 = \cdots = \alpha_m = 0$ otteniamo

$$0\, v_1 + 0\, v_2 + \cdots + 0\, v_m = 0_V$$

qualunque siano $v_1, v_2, ..., v_m$.

Esempio 1.12. *In $\mathbb{R}^3$ siano $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$, $\mathbf{v}_2 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$, $\mathbf{v}_3 = \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix}$. Il vettore $\mathbf{u} = \begin{pmatrix} 3 \\ -2 \\ -3 \end{pmatrix}$ è combinazione lineare di $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$. Infatti*

$$\mathbf{u} = \begin{pmatrix} 3 \\ -2 \\ -3 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} + \begin{pmatrix} 2 \\ -2 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ -2 \end{pmatrix} = \mathbf{v}_1 + 2\mathbf{v}_2 - \mathbf{v}_3.$$

Definizione 1.13. *Sia V uno spazio vettoriale su $\mathbb{R}$ e siano $v_1, v_2, ..., v_m$ vettori in V, $m \geq 1$. I vettori $v_1, v_2, ..., v_m$ si dicono linearmente dipendenti (l.d.) se almeno uno di essi si può scrivere come combinazione lineare dei rimanenti.*

Esempio 1.14. *In $\mathbb{R}^3$ i vettori $\mathbf{v}_1 = \begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix}$, $\mathbf{v}_2 = \begin{pmatrix} -3 \\ 0 \\ 1 \end{pmatrix}$, $\mathbf{v}_3 = \begin{pmatrix} -1 \\ -4 \\ 1 \end{pmatrix}$ sono linearmente dipendenti, perché*

$$\mathbf{v}_3 = 2\mathbf{v}_1 + \mathbf{v}_2 = \begin{pmatrix} 2 \\ -4 \\ 0 \end{pmatrix} + \begin{pmatrix} -3 \\ 0 \\ 1 \end{pmatrix}.$$

Esempio 1.15. *In $\mathbb{P}_2$ i polinomi $p_1(t) = 1 - t$, $p_2(t) = t(1 - t)$ e $p_3(t) = 1 - t^2$ sono linearmente dipendenti, perché*

$$p_1(t) + p_2(t) = 1 - t + t - t^2 = 1 - t^2 = p_3(t).$$

Definizione 1.16. *Sia V uno spazio vettoriale su $\mathbb{R}$, e siano $v_1, v_2, ..., v_m$ vettori in V, $m \geq 1$. I vettori $v_1, v_2, ..., v_m$ si dicono linearmente indipendenti (l.i.) se nessuno di essi si può scrivere come combinazione lineare dei rimanenti.*

Esempio 1.17. *In $\mathbb{R}^3$ i vettori* $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$, $\mathbf{v}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$, $\mathbf{v}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ *sono linearmente indipendenti. Infatti, se fosse* $\mathbf{v}_1 = \alpha_2\mathbf{v}_2 + \alpha_3\mathbf{v}_3$ *per* $\alpha_2, \alpha_3 \in \mathbb{R}$*, sarebbe*

$$\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \mathbf{v}_1 = \alpha_2\mathbf{v}_2 + \alpha_3\mathbf{v}_3 = \begin{pmatrix} 0 \\ \alpha_2 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \alpha_3 \end{pmatrix} = \begin{pmatrix} 0 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} \quad \text{impossibile}$$

Se fosse $\mathbf{v}_2 = \alpha_1\mathbf{v}_1 + \alpha_3\mathbf{v}_3$ *per* $\alpha_1, \alpha_3 \in \mathbb{R}$*, sarebbe*

$$\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \mathbf{v}_2 = \alpha_1\mathbf{v}_1 + \alpha_3\mathbf{v}_3 = \begin{pmatrix} \alpha_1 \\ 0 \\ \alpha_1 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \alpha_3 \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ 0 \\ \alpha_1 + \alpha_3 \end{pmatrix} \quad \text{impossibile}$$

Se fosse $\mathbf{v}_3 = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2$ *per* $\alpha_1, \alpha_2 \in \mathbb{R}$*, sarebbe*

$$\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \mathbf{v}_3 = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 = \begin{pmatrix} \alpha_1 \\ 0 \\ \alpha_1 \end{pmatrix} + \begin{pmatrix} 0 \\ \alpha_2 \\ 0 \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_1 \end{pmatrix} \quad \text{impossibile}$$

Esempio 1.18. *In $\mathbb{P}_2$ i polinomi* $p_1(t) = 1-t$, $p_2(t) = 1+t$) *e* $p_3(t) = t^2$ *sono linearmente indipendenti. Infatti, se fosse* $p_1(t) = \alpha_2 p_2(t) + \alpha_3 p_3(t)$*, sarebbe*

$$1 - t = \alpha_3 t^2 + \alpha_2 t + \alpha_2 \Leftrightarrow \alpha_3 = 0,\ \alpha_2 = -1,\ \alpha_2 = 1 \quad \text{impossibile}$$

Se fosse $p_2(t) = \alpha_1 p_1(t) + \alpha_3 p_3(t)$*, sarebbe*

$$1 + t = \alpha_3 t^2 - \alpha_1 t + \alpha_1 \Leftrightarrow \alpha_3 = 0,\ \alpha_1 = -1,\ \alpha_1 = 1 \quad \text{impossibile}$$

Se fosse $p_3(t) = \alpha_1 p_1(t) + \alpha_2 p_2(t)$*, sarebbe*

$$t^2 = (\alpha_2 - \alpha_1)t + (\alpha_1 + \alpha_2) \quad \text{impossibile}$$

Osservazione 1.19. *(i) Siano* $v_1, v_2, ..., v_m$ *vettori in* V, $m \geq 1$. *Se uno di essi è il vettore nullo, allora* $v_1, v_2, ..., v_m$ *sono linearmente dipendenti. Infatti, possiamo supporre che sia* $v_1 = 0_V$*, e dunque*

$$v_1 = 0_V = 0v_2 + \cdots + 0v_m.$$

(ii) Sia $T = \{v_1, v_2, ..., v_m\}$ *un insieme di vettori linearmente indipendenti. Ogni sottoinsieme* S *di* T *è costituito da vettori linearmente indipendenti.*

(iii) Sia $D = \{v_1, v_2, ..., v_m\}$ *un insieme di vettori linearmente dipendenti. Ogni insieme* C *che contiene* D *è costituito da vettori linearmente dipendenti.*

Teorema 1.20. *Sia* V *uno spazio vettoriale su* $\mathbb{R}$.

1. *I vettori* $v_1, v_2, ..., v_m$ *di* V *sono linearmente dipendenti se e solo se esiste una loro combinazione lineare a coefficienti non tutti nulli uguale a* 0_V.

2. *I vettori* $v_1, v_2, ..., v_m$ *di* V *linearmente indipendenti se e solo se l'unica combinazione lineare di* $v_1, v_2, ..., v_m$ *uguale a* 0_V *è quella i cui coefficienti sono tutti nulli.*

Dim. (1) Mostriamo l'implicazione ($\Rightarrow$). Siano $v_1, v_2, ..., v_m$ linearmente dipendenti. Uno di essi è combinazione lineare degli altri. Possiamo supporre che sia

$$v_1 = \alpha_2 v_2 + \cdots + \alpha_m v_m.$$

Allora

$$v_1 - \alpha_2 v_2 - \cdots - \alpha_m v_m = 0_V.$$

Abbiamo trovato una combinazione lineare di $v_1, v_2, ..., v_m$ i cui coefficienti non sono tutti nulli (il coeffciente di v_1 è 1) uguale a 0_V.
(1) Mostriamo l'implicazione ($\Leftarrow$). Supponiamo che esistano $\beta_1, \beta_2, ..., \beta_m \in \mathbb{R}$, non tutti nulli, tali che

$$\beta_1 v_1 + \beta_2 v_2 + \cdots + \beta_m v_m = 0_V.$$

Possiamo supporre che $\beta_1 \neq 0$. Allora

$$v_1 = -\frac{\beta_2}{\beta_1} v_2 - \cdots - \frac{\beta_m}{\beta_1} v_m.$$

(2) Segue dalla parte (1). Infatti, se $v_1, ..., v_m$ sono linearmente indipendenti, non possiamo avere che una loro combinazione lineare a coefficienti non tutti nulli è uguale a 0_V, perché altrimenti, per la parte 1), $v_1, ..., v_m$ sarebbero linearmente dipendenti. Il viceversa si argomenta in modo analogo. □

Esempio 1.21. *Consideriamo i vettori dell'Esempio 1.17. Per mostrare che sono linearmente indipendenti ora usiamo il Teorema 1.20-(2). Basta mostrare che*

$$\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \alpha_3 \mathbf{v}_3 = \mathbf{0}_{\mathbb{R}^3}$$

ammette l'unica soluzione $\alpha_1 = \alpha_2 = \alpha_3 = 0$. Ora,

$$\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \alpha_3 \mathbf{v}_3 = \alpha_1 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + \alpha_2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + \alpha_3 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_1 + \alpha_3 \end{pmatrix}.$$

Si ha

$$\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_1 + \alpha_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{cases} \alpha_1 = 0 \\ \alpha_2 = 0 \\ \alpha_3 = 0 \end{cases}$$

Esempio 1.22. *Consideriamo i vettori dell'Esempio 1.14. Per mostrare che sono linearmente dipendenti ora usiamo il Teorema 1.20-(1). Basta mostrare che*

$$\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \alpha_3 \mathbf{v}_3 = \mathbf{0}_{\mathbb{R}^3}$$

ammette almeno una soluzione $\alpha_1, \alpha_2, \alpha_3$ non tutti nulli. Ora,

$$\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \alpha_3 \mathbf{v}_3 = \alpha_1 \begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix} + \alpha_2 \begin{pmatrix} -3 \\ 0 \\ 1 \end{pmatrix} + \alpha_3 \begin{pmatrix} -1 \\ -4 \\ 1 \end{pmatrix} = \begin{pmatrix} \alpha_1 - 3\alpha_2 - \alpha_3 \\ -2\alpha_1 - 4\alpha_3 \\ \alpha_2 + \alpha_3 \end{pmatrix}.$$

Si ha

$$\begin{pmatrix} \alpha_1 - 3\alpha_2 - \alpha_3 \\ -2\alpha_1 - 4\alpha_3 \\ \alpha_2 + \alpha_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{cases} \alpha_1 - 3\alpha_2 - \alpha_3 = 0 \\ -2\alpha_1 - 4\alpha_3 = 0 \\ \alpha_2 + \alpha_3 = 0 \end{cases} \Leftrightarrow \begin{cases} \alpha_1 = -2\alpha_3 \\ \alpha_2 = -\alpha_3 \end{cases}$$

Esistono pertanto infinite soluzioni, che dipendono da α_3*. Per esempio, se* $\alpha_3 = 1$, $2\mathbf{v}_1 + \mathbf{v}_2 - \mathbf{v}_3 = \mathbf{0}_{\mathbb{R}^3}$.

1.4 Basi e dimensione di uno spazio vettoriale

Definizione 1.23. *Sia V uno spazio vettoriale su $\mathbb{R}$. I vettori $v_1, v_2, ..., v_m$ in V si dicono generatori di V se ogni vettore di V si può scrivere come combinazione lineare di $v_1, v_2, ..., v_m$, ovvero*

$$\forall\, v,\ \exists\, \alpha_1, ..., \alpha_m \in \mathbb{R} \text{ tali che } v = \alpha_1 v_1 + \cdots + \alpha_m v_m.$$

Esempio 1.24. *I vettori* $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$, $\mathbf{v}_2 = \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix}$, $\mathbf{v}_3 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$ *sono generatori di* $\mathbb{R}^3$*. Per verificarlo, bisogna mostrare che, per ogni* $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3$*, esistono* $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{R}$ *tali che*

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \alpha_3 \mathbf{v}_3 = \alpha_1 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + \alpha_2 \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix} + \alpha_3 \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

ovvero

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \alpha_1 + \alpha_3 \\ -\alpha_2 + \alpha_3 \\ \alpha_1 \end{pmatrix}.$$

Si ottiene il sistema

$$\begin{cases} \alpha_1 + \alpha_3 = x_1 \\ -\alpha_2 + \alpha_3 = x_2 \\ \alpha_1 = x_3 \end{cases}$$

che ha la soluzione

$$\begin{cases} \alpha_1 = x_3 \\ \alpha_2 = x_1 - x_2 - x_3 \\ \alpha_3 = x_1 - x_3 \end{cases}$$

Abbiamo provato che il generico vettore $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3$ *si scrive come*

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_3 \mathbf{v}_1 + (x_1 - x_2 - x_3)\mathbf{v}_2 + (x_1 - x_3)\mathbf{v}_3.$$

Definizione 1.25. *Una base di uno spazio vettoriale V è un insieme di vettori di V che*

1. *generano V;*
2. *sono linearmente indipendenti.*

Base canonica di $\mathbb{R}^n$.

In $\mathbb{R}^n$ consideriamo i vettori

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \cdots, \mathbf{e}_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

L'insieme di vettori $\{\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n\}$ è una base di $\mathbb{R}^n$, e si indica con base canonica.

(1) Proviamo che $\{\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n\}$ generano $\mathbb{R}^n$. Per ogni $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$,

$$\begin{aligned} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} &= x_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \cdots + x_n \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \\ &= x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots + x_n\mathbf{e}_n. \end{aligned}$$

(2) Proviamo che $\{\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n\}$ sono linearmente indipendenti. Se fosse $\alpha_1\mathbf{e}_1 + \alpha_2\mathbf{e}_2 + \cdots + \alpha_n\mathbf{e}_n = \mathbf{0}_{\mathbb{R}^n}$, avremmo

$$\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

da cui $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$.

Osservazione 1.26. *La base canonica di $\mathbb{R}^3$ è $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$, dove*

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \mathbf{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

La base canonica di $\mathbb{R}^4$ è $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4\}$ dove

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \mathbf{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \mathbf{e}_4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

Esempio 1.27. *I vettori dell'Esempio 1.24 formano una base di $\mathbb{R}^3$. Abbiamo già visto che generano $\mathbb{R}^3$, resta da mostare che sono linearmente indipendenti. Siano $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{R}$ con $\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \alpha_3 \mathbf{v}_2 = \mathbf{0}_{\mathbb{R}^3}$. Allora*

$$\begin{cases} \alpha_1 + \alpha_3 = 0 \\ -\alpha_2 + \alpha_3 = 0 \\ \alpha_1 = 0 \end{cases}$$

da cui $\alpha_1 = \alpha_2 = \alpha_3 = 0$.

Esempio 1.28. *I polinomi $p_1(t) = 1$, $p_2(t) = t$, $p_3(t) = t^2$ formano una base di $\mathbb{P}_2$. È chiaro che generano $\mathbb{P}_2$, poiché ogni polinomio di grado minore o uguale a 2 si scrive come combinazione lineare di $1, t, t^2$. Inoltre sono anche linearmente indipendenti, perché se $\alpha_1 + \alpha_2 t + \alpha_3 t^2$ è usuale al polinomio nullo, allora $\alpha_1 = \alpha_2 = \alpha_3 = 0$.*

Esempio 1.29. *I polinomi $p_1(t) = 1$, $p_2(t) = t$, $p_3(t) = t^2$,..., $p_m(t) = t^m$ formano una base di $\mathbb{P}_m$. Si ragiona come nell'Esempio 1.28.*

Si può dimostrare che ogni spazio vettoriale V ammette almeno una base. L'importanza delle basi è dato dal seguente teorema.

Teorema 1.30. *Siano V uno spazio vettoriale e $B = \{v_1, ..., v_m\}$ una base di V. Allora ogni vettore di V si può scrivere in un modo unico come combinazione lineare dei vettori di B.*

Dim. Sia $v \in V$ e supponiamo che

$$v = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_m v_m$$
$$v = \beta_1 v_1 + \beta_2 v_2 + \cdots + \beta_m v_m$$

per $\alpha_1, \alpha_2, ..., \alpha_m \in \mathbb{R}$, $\beta_1, \beta_2, ..., \beta_m \in \mathbb{R}$. Allora

$$0_V = (\alpha_1 - \beta_1)v_1 + (\alpha_2 - \beta_2)v_2 + \cdots + (\alpha_m - \beta_m)v_m.$$

Poiché i vettori di B sono linearmente indipendenti, risulta $\alpha_1 = \beta_1$, ..., $\alpha_m = \beta_m$. □

Definizione 1.31. *Siano V uno spazio vettoriale e $B = \{v_1, ..., v_m\}$ una base di V. Si dicono coordinate di un vettore $v \in V$ rispetto a B i coefficienti $\alpha_1, ..., \alpha_m \in \mathbb{R}$ tali che $v = \alpha_1 v_1 + \cdots + \alpha_m v_m$.*

Un altro risultato fondamentale sulle basi, che non dimostriamo, è il seguente.

Teorema 1.32. *Siano V uno spazio vettoriale, B una base di V costituita da n vettori e B' una base di V costituita da m vettori. Allora $n = m$.*

Abbiamo quindi che tutte le basi di V hanno la stesso numero di vettori. Questo giustifica la seguente definizione.

Definizione 1.33. *La dimensione di uno spazio vettoriale V è il numero di elementi di una base di V.*

Notazione. Dato uno spazio vettoriale V, indicheremo con $\dim(V)$ la sua dimensione.

Osservazione 1.34. *(i)* $\mathbb{R}^n$ *ha dimensione* n.

(ii) $\mathbb{P}_m$ *ha dimensione* $m+1$.

Enunciamo senza dimostrare il seguente risultato, che è particolarmente utile per trovare una base di uno spazio vettoriale di cui si conosce la dimensione.

Teorema 1.35. *Sia V uno spazio vettoriale di dimensione n. Allora*

1. *m vettori in V con $m > n$ sono linearmente dipendenti;*
2. *n vettori di V formano una base di V se e solo se essi sono linearmente indipendenti;*
3. *n vettori di V formano una base di V se e solo se essi generano V.*

Esempio 1.36. *I vettori*

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}, \mathbf{v}_3 = \begin{pmatrix} 0 \\ 2 \\ -1 \end{pmatrix}$$

sono linearmente indipendenti (si lascia la verifica per eserciziot). Per il Teorema 1.35, $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ è una base di $\mathbb{R}^3$.

Esempio 1.37. *I vettori*

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \mathbf{v}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \mathbf{v}_4 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}$$

sono linearmente indipendenti (si lascia la verifica per esercizio). Per il Teorema 1.35, $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ è una base di $\mathbb{R}^4$.

Esempio 1.38. *I polinomi $p_1(t) = 1$, $p_2(t) = t+1$, $p_3(t) = (t+1)^2$ sono linearmente indipendenti (si lascia la verifica per esercizio). Per il Teorema 1.35, $\{p_1(t), p_2(t), p_3(t)\}$ è una base di $\mathbb{P}_2$.*

Sottospazio generato da un sottoinsieme.

Definizione 1.39. *Sia V uno spazio vettoriale su $\mathbb{R}$, e sia X un sottoinsieme di V. Indichiamo con $\langle X \rangle$ il sottospazio di V generato da X, dato dall'intersezione di tutti i sottospazi di V che contengono X:*

$$\langle X \rangle = \bigcap_{\substack{W \supseteq X \\ W \text{ sottospazio di } V}} W.$$

Nel caso in cui $X = \{v_1, ..., v_m\}$ è un insieme di $m \geq 1$ vettori, indichiamo $\langle X \rangle$ con

$$\operatorname{span}\{v_1, ..., v_m\}.$$

Osservazione 1.40. *Se X è già un sottospazio, allora $\langle X \rangle = X$.*

Nella prossima proposizione descriviamo esplicitamente $\text{span}\{v_1, ..., v_m\}$.

Proposizione 1.41. *Sia V uno spazio vettoriale su $\mathbb{R}$ e siano $v_1, ..., v_m$ vettori in V, $m \geq 1$. Allora $\text{span}\{v_1, ..., v_m\}$ è l'insieme di tutte le combinazioni lineari di $v_1, ..., v_m$ a coefficienti reali:*

$$\text{span}\{v_1, ..., v_m\} = \{\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_m v_m \mid \alpha_1, ..., \alpha_m \in \mathbb{R}\}.$$

Dim. Poniamo

$$U := \{\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_m v_m \mid \alpha_1, ..., \alpha_m \in \mathbb{R}\}$$

e mostriamo che U è il sottospazio di V generato da $\{v_1, ..., v_m\}$, dato cioè dall'intersezione di tutti i sottospazi di V che contengono $\{v_1, ..., v_m\}$:

$$U = \bigcap_{\substack{W \supseteq \{v_1, ..., v_m\} \\ W \text{ sottospazio di } V}} W.$$

In altre parole, dobbiamo mostrare che

1. U è un sottospazio di V;
2. se W è un sottospazio di V che contiene i vettori $v_1, ..., v_m$, allora $U \subseteq W$.

(1) Abbiamo

(i) $0_V \in U$ perché $0_V = 0v_1 + 0v_2 + \cdots 0v_m$.

(ii) $\forall\, (\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_m v_m), (\beta_1 v_1 + \beta_2 v_2 + \cdots + \beta_m v_m) \in U$ si ha

$$\begin{aligned}&(\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_m v_m) + (\beta_1 v_1 + \beta_2 v_2 + \cdots + \beta_m v_m)\\ &= (\alpha_1 + \beta_1) v_1 + (\alpha_2 + \beta_2) v_2 + \cdots + (\alpha_m + \beta_m) v_m \in U\end{aligned}$$

(iii) $\forall\, \alpha \in \mathbb{R}, \forall\, (\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_m v_m) \in U$ si ha

$$\alpha(\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_m v_m) = (\alpha\alpha_1) v_1 + (\alpha\alpha_2) v_2 + \cdots + (\alpha\alpha_m) v_m \in U$$

(2) Sia W un sottospazio di V che contiene $v_1, ..., v_m$. Per definizione di sottospazio, per ogni $\alpha_1, \alpha_2 \in \mathbb{R}$ e $w_1, w_2 \in W$ si ha $\alpha_1 w_1 + \alpha_2 w_2 \in W$. In particolare, W contiene ogni combinazione lineare dei suoi vettori, cioè

$$\forall\, w_1, ..., w_r \in W, \forall\, \alpha_1, ..., \alpha_r \in \mathbb{R}, \quad \alpha_1 w_1 + \cdots + \alpha_r w_r \in W, \quad r \in \mathbb{N}.$$

È chiaro allora che $U \subseteq W$ perché, in particolare, $\forall\, \alpha_1, ..., \alpha_m \in \mathbb{R}$ si ha $\alpha_1 v_1 + \cdots + \alpha_m v_m \in W$. $\square$

Possiamo riformulare la definizione di base di uno spazio vettoriale come segue.

Definizione 1.42. *Sia V uno spazio vettoriale su $\mathbb{R}$. I vettori $v_1, ..., v_n$ di V formano una base di V se*

1. $\operatorname{span}\{v_1, ..., v_n\} = V$;
2. *i vettori $v_1, ..., v_n$ sono linearmente indipendenti.*

Proposizione 1.43. *Siano $v_1, ..., v_m$ vettori di uno spazio vettoriale V e sia*

$$u = \beta_1 v_1 + \cdots + \beta_m v_m.$$

Allora

$$\operatorname{span}\{v_1, ..., v_m\} = \operatorname{span}\{v_1, ..., v_m, u\}.$$

Dim. È sempre vero che

$$\operatorname{span}\{v_1, ..., v_m\} \subseteq \operatorname{span}\{v_1, ..., v_m, u\}.$$

Viceversa, sia $w \in \operatorname{span}\{v_1, ..., v_m, u\}$ tale che

$$w = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_m v_m + \alpha_{m+1} u.$$

Allora

$$\begin{aligned} w &= \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_m v_m + \alpha_{m+1}(\beta_1 v_1 + \cdots + \beta_m v_m) \\ &= (\alpha_1 + \alpha_{m+1}\beta_1) v_1 + (\alpha_2 + \alpha_{m+1}\beta_2) v_2 + \cdots + (\alpha_m + \alpha_{m+1}\beta_m) v_m \end{aligned}$$

cioè $w \in \operatorname{span}\{v_1, ..., v_m\}$. □

In particolare, se $v_1, ..., v_m, u$ generano V (cioè $\operatorname{span}\{v_1, ..., v_m, u\} = V$) e u è combinazione lineare di $v_1, ..., v_m$, allora anche $v_1, ..., v_m$ generano V (cioè $\operatorname{span}\{v_1, ..., v_m\} = V$).

Dimensione di un sottospazio.

Se W è un sottospazio dello spazio vettoriale V, sappiamo che W è uno spazio vettoriale rispetto alle stesse operazioni di V, si veda l'Osservazione 1.5. Possiamo allora ripetere per un sottospazio la nozione di base.

Definizione 1.44. *Siano V uno spazio vettoriale su $\mathbb{R}$ e W un sottospazio di V. I vettori $w_1, ..., w_r$ di W formano una base di W se*

1. $\operatorname{span}\{w_1, ..., w_r\} = W$;
2. *i vettori $w_1, ..., w_r$ sono linearmente indipendenti.*

Osservazione 1.45. *Siano $v_1, .., v_m$ vettori di uno spazio vettoriale V. La dimensione di $\operatorname{span}\{v_1, ..., v_m\}$ (il sottospazio di V generato da $v_1, ..., v_m$) è pari al massimo numero di vettori tra $v_1, ..., v_m$ linearmente indipendenti.*

Infatti, dalla Proposizione 1.43, se un vettore $v_1, ..., v_m$ è combinazione lineare degli altri possiamo "eliminarlo" ottenendo sempre lo stesso sottospazio.

Osservazione 1.46. *La dimensione di un sottospazio W di V è sempre minore o uguale alla dimensione di V. L'uguaglianza vale se e solo se $W = V$.*

Capitolo 2

Matrici

2.1 Introduzione all'algebra delle matrici

Definizione 2.1. *Una matrice $m \times n$ a coefficienti in $\mathbb{R}$ è una tabella di $m \times n$ elementi disposti su m righe e n colonne. Se $m \neq n$ la matrice di dice rettangolare, se $m = n$ la matrice si dice quadrata di ordine n. Nel caso di matrici rettangolari,*

- *se $m = 1$ la matrice si dice vettore riga;*
- *se $n = 1$ la matrice si dice vettore colonna.*

Notazioni. L'insieme delle matrici $m \times n$ a coefficienti in $\mathbb{R}$ si denota con $\mathrm{Mat}_{m,n}(\mathbb{R})$. Nel caso di matrici quadrate, la scrittura $\mathrm{Mat}_{m,n}(\mathbb{R})$ è sostituita con $\mathrm{Mat}_n(\mathbb{R})$.

Indicheremo indistintamente $i = 1, , , m$, $j = 1, ...n$, oppure $i = 1 \div m$, $j = 1 \div n$, oppure $i \in I_m$, $j \in I_n$, con $I_m := 1, ..., m$, $I_n := 1, ..., n$.

Data una matrice $A \in \mathrm{Mat}_{m,n}(\mathbb{R})$, scriveremo esplicitamente $A = [a_{ij}]$: per ogni $i \in I_m := 1, ..., m$, $j \in I_n := 1, ..., n$, a_{ij} indica il coefficiente di A sulla riga i e colonna j, ed è detto l'elemento di A di posto i, j.

Definizione 2.2. *Data una matrice $A = [a_{ij}] \in \mathrm{Mat}_n(\mathbb{R})$, si possono individuare due sottoinsiemi*

- *diagonale principale: $\{a_{ii} \ : \ i \in I_n\}$;*
- *diagonale secondaria: $\{a_{i\ (n+1)-i} \ : \ i \in I_n\}$.*

Definiamo alcuni sottoinsiemi di $\mathrm{Mat}_n(\mathbb{R})$.

Definizione 2.3. *Sia $A = [a_{ij}] \in \mathrm{Mat}_n(\mathbb{R})$. A è detta*

- *triangolare superiore se $a_{ij} = 0$ per ogni $i, j \in I_n$ tali che $j < i$;*
- *triangolare inferiore se $a_{ij} = 0$ per ogni $i, j \in I_n$ tali che $j > i$;*
- *diagonale se $a_{ij} = 0$ per ogni $i, j \in I_n$ con $i \neq j$;*
- *scalare se A è diagonale e $a_{ii} = \lambda \in \mathbb{R}$ per ogni $i \in I_n$.*

Notazione. Indichiamo con $\mathrm{Diag}_n(\mathbb{R})$ il sottoinsieme di $\mathrm{Mat}_n(\mathbb{R})$ delle matrici diagonali, con $\mathrm{Triang}_n(\mathbb{R})$ il sottoinsieme di $\mathrm{Mat}_n(\mathbb{R})$ delle matrici triangolari.

Definizione 2.4. *Due matrici $A = [a_{ij}]$ e $B = [b_{ij}]$ si dicono uguali se*

$$A, B \in \mathrm{Mat}_{m,n}(\mathbb{R}) \quad \text{e} \quad a_{ij} = b_{ij} \quad \forall (i,j) \in I_m \times I_n.$$

Definizione 2.5. *Sia $A = [a_{ij}] \in \mathrm{Mat}_{m,n}(\mathbb{R})$. Si dice trasposta di A la matrice A^T i cui elementi a'_{ij} sono*

$$a'_{ij} = a_{ji} \quad \forall (i,j) \in I_m \times I_n.$$

Proposizione 2.6. *Sia $A = [a_{ij}] \in \mathrm{Mat}_{m,n}(\mathbb{R})$. Si ha $(A^T)^T = A$.*

Proposizione 2.7. *Una matrice $A = [a_{ij}] \in \mathrm{Mat}_{m,n}(\mathbb{R})$ è triangolare superiore se e solo se A^T è triangolare inferiore.*

Definizione 2.8. *Una matrice $A = [a_{ij}] \in \mathrm{Mat}_{m,n}(\mathbb{R})$ si dice simmetrica se e solo se $A = A^T$.*

Osservazione 2.9. *(i) Se una matrice è simmetrica allora è quadrata;*

(ii) ogni matrice diagonale è simmetrica.

Definizione 2.10. *Siano $A = [a_{ij}]$, $B = [b_{ij}] \in \mathrm{Mat}_{m,n}(\mathbb{R})$. Definiamo*

- *matrice somma di A con B, la matrice $A + B \in \mathrm{Mat}_{m,n}(\mathbb{R})$ il cui elemento di posto i, j è*
$$a_{ij} + b_{ij} \quad \forall (i,j) \in I_m \times I_n.$$
- *matrice prodotto dello scalare $\alpha \in \mathbb{R}$ per la matrice A, la matrice $(\alpha A) \in \mathrm{Mat}_{m,n}(\mathbb{R})$ il cui elemento di posto i, j è*
$$\alpha\, a_{ij} \quad \forall (i,j) \in I_m \times I_n.$$

Esercizio 2.11. *Siano $\alpha = -2 \in \mathbb{R}$ e*

$$A = \begin{bmatrix} 3 & 4 & \sqrt{5} \\ 1 & 0 & 8 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & -1 & -\sqrt{5} \\ -2 & 7 & 4 \end{bmatrix} \in \mathrm{Mat}_{2,3}(\mathbb{R}).$$

Calcolare $A + B$ e αB.

Definizione 2.12. *Sia $A \in \mathrm{Mat}_n(\mathbb{R})$. L'operazione somma degli elementi sulla diagonale di A è detta traccia di A ed indicata con $Tr(\cdot)$:*

$$Tr(A) = \sum_{i=1}^{n} a_{ii}.$$

2.2 Lo spazio vettoriale delle matrici

Proposizione 2.13. *L'insieme* $\text{Mat}_{m,n}(\mathbb{R})$ *con le operazioni di somma di matrici e prodotto di uno scalare per una matrice è uno spazio vettoriale su* $\mathbb{R}$.

Dim. L'insieme $\text{Mat}_{m,n}(\mathbb{R})$ con le operazioni di somma di matrici e prodotto di uno scalare per una matrice verifica le proprietà 1-8 della Definizione 1.1:

1. $\forall A, B \in \text{Mat}_{m,n}(\mathbb{R}), \quad A + B = B + A$
2. $\forall A, B, C \in \text{Mat}_{m,n}(\mathbb{R}), \quad (A + B) + C = A + (B + C)$
3. $\forall A \in \text{Mat}_{m,n}(\mathbb{R})$, $A + O = A$, per $O = [\bar{a}_{ij}]$ con $\bar{a}_{ij} = 0$ per ogni $(i, j) \in I_m \times I_n$; O si dice matrice nulla
4. $\forall A \in \text{Mat}_{m,n}(\mathbb{R})$ $A + (-A) = O$, per $-A = [a'_{i,j}]$ con $a'_{ij} = -a_{ij}$ per ogni $(i, j) \in I_m \times I_n$; $-A$ si dice la matrice opposta di A
5. $\forall \alpha, \beta \in \mathbb{R}, A \in \text{Mat}_{m,n}(\mathbb{R}), \quad \alpha(\beta A) = (\alpha\beta)A$
6. $\forall A \in \text{Mat}_{m,n}(\mathbb{R}), \quad 1A = A$
7. $\forall \alpha \in \mathbb{R}, A, B \in \text{Mat}_{m,n}(\mathbb{R}), \quad \alpha(A + B) = \alpha A + \alpha B$
8. $\forall \alpha, \beta \in \mathbb{R}, A \in \text{Mat}_{m,n}(\mathbb{R}), \quad (\alpha + \beta)A = \alpha A + \beta A$ □

Esercizio 2.14. *Date le matrici*

$$A = \begin{bmatrix} 0 & 0 \\ 1 & 2 \\ 3 & 4 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 1 \\ 4 & 1 \\ 0 & 2 \end{bmatrix}, \quad C = \begin{bmatrix} 4 & 3 \\ 3 & 3 \\ 1 & 0 \end{bmatrix} \in \text{Mat}_{3,2}(\mathbb{R}),$$

verificare che $(A + B) + C = A + (B + C)$ *e che* $A + B = B + A$.

Proposizione 2.15. *Per ogni* $A, B \in \text{Mat}_{m,n}(\mathbb{R})$, $\alpha \in \mathbb{R}$, *si ha*

$$(A + B)^T = A^T + B^T$$
$$(\alpha A)^T = \alpha A^T$$

2.3 Prodotto tra matrici

Notazione. Nel seguito indicheremo con $\mathbb{I}_n$ la matrice scalare con $a_{ii} = 1$ per ogni $i \in I_n$.

www.unidocs.it - Appunti e dispense per superare i tuoi esami universitari

Definizione 2.16. *Dati un vettore riga $A \in \mathrm{Mat}_{1,p}(\mathbb{R})$ e un vettore colonna $B \in \mathrm{Mat}_{p,1}(\mathbb{R})$ con*

$$A = [a_{11} \cdots a_{1p}], \quad B = [b_{11} \cdots b_{p1}]^T,$$

chiamiamo prodotto della riga A per la colonna B lo scalare k definito da

$$k = [a_{11} \cdots a_{1p}] \cdot [b_{11} \cdots b_{p1}]^T = a_{11}b_{11} + a_{12}b_{21} + \cdots + a_{1p}b_{p1} = \sum_{i=1}^{p} a_{1i}b_{i1}.$$

Esercizio 2.17. *Siano $A = \begin{bmatrix} 2 & 1 & 0 & 6 \end{bmatrix}$, $B = \begin{bmatrix} -1 & 2 & \frac{1}{2} & 1 \end{bmatrix}^T$. Calcolare AB.*

Definizione 2.18. *Date le matrici $A = [a_{ih}] \in \mathrm{Mat}_{m,p}(\mathbb{R})$ e $B = [b_{hj}] \in \mathrm{Mat}_{p,n}(\mathbb{R})$, chiamiamo prodotto (righe per colonne) di A con B, la matrice $AB \in \mathrm{Mat}_{m,n}(\mathbb{R})$ avente come elemento di posto (i,j) il prodotto della i-esima riga di A con la j-esima colonna di B:*

$$(AB)_{ij} = [a_{i1} \cdots a_{ip}] \cdot [b_{1j} \cdots b_{pj}]^T, \quad \forall (i,j) \in I_m \times I_n.$$

Esercizio 2.19. *Siano*

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

Calcolare AB.

Osservazione 2.20. *(i) Siano $A \in \mathrm{Mat}_{m,n}(\mathbb{R})$, $B \in \mathrm{Mat}_{p,q}(\mathbb{R})$. Il prodotto AB è definito se $n = p$; in tal caso si ha $AB \in \mathrm{Mat}_{m,q}(\mathbb{R})$. Il prodotto BA è definito se $m = q$; in tal caso si ha $BA \in \mathrm{Mat}_{p,n}(\mathbb{R})$.*

(ii) Il prodotto tra matrici non è commutativo. Nelle notazioni del punto (i), ponendo $n = p$ e $m = q$, si hanno

$$AB \in \mathrm{Mat}_m(\mathbb{R}) \quad \text{e} \quad BA \in \mathrm{Mat}_n(\mathbb{R}).$$

Il prodotto tra matrici gode delle seguenti proprietà.

Proposizione 2.21. *Siano $A, B \in \mathrm{Mat}_{m,n}(\mathbb{R})$, $C, D \in \mathrm{Mat}_{n,p}(\mathbb{R})$, $E \in \mathrm{Mat}_{p,q}(\mathbb{R})$. Allora*

(i) $(AC)E = A(CE)$;

(ii) $(A+B)C = AC + BC$, $B(C+D) = BC + BD$;

(iii) per ogni $\alpha \in \mathbb{R}$, $\alpha(AC) = (\alpha A)C = A(\alpha C)$;

(iv) $A\mathbb{I}_n = A$, $\mathbb{I}_n C = C$;

(v) se è definito il prodotto di matrici XY, allora lo è anche quello $Y^T X^T$, ed equivale a $(XY)^T$.

Definizione 2.22. *Una matrice* $A \in \mathrm{Mat}_n(\mathbb{R})$ *è detta invertibile se esiste una matrice* $B \in \mathrm{Mat}_n(\mathbb{R})$ *tale che*

$$AB = BA = \mathbb{I}_n.$$

In tal caso, la matrice B *è unica, si dice matrice inversa di* A *e si indica con* A^{-1}.

Proposizione 2.23. *Siano* $A, B \in \mathrm{Mat}_n(\mathbb{R})$ *invertibili. Allora*

(i) A^{-1} *è invertibile e* $(A^{-1})^{-1} = A$;

(ii) A^T *è invertibile e* $(A^T)^{-1} = (A^{-1})^T$;

(iii) AB *è invertibile e* $(AB)^{-1} = B^{-1}A^{-1}$.

Definizione 2.24. *Una matrice* $A \in \mathrm{Mat}_n(\mathbb{R})$ *si dice ortogonale se* $A^{-1} = A^T$.

2.4 Il determinante

Notazione. Sia $A \in \mathrm{Mat}_n(\mathbb{R})$. Nel seguito indicheremo con A_{ij} la matrice di $\mathrm{Mat}_{n-1}(\mathbb{R})$ ottenuta eliminando da A la i-esima riga e la j-esima colonna.

Definizione 2.25. *Chiamiamo determinante di* $A = [a_{ij}] \in \mathrm{Mat}_n(\mathbb{R})$ *rispetto alla prima riga lo scalare definito come segue:*

$$\det A := \begin{cases} a_{11} & \text{se } n = 1 \\ \sum_{j \in I_n} (-1)^{1+j} a_{1j} \det A_{1j} & \text{se } n \geq 2. \end{cases}$$

Il determinante si indica anche con $|A|$.

Esempio 2.26. *Calcolare il determinante delle seguenti matrici:*

$$A = \begin{bmatrix} 2 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \quad C = \begin{bmatrix} -2 & 1 & 0 \\ 1 & -2 & 6 \\ -1 & 2 & 3 \end{bmatrix}.$$

Osservazione 2.27 (Calcolo del determinante di matrici particolari)**.**

- $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \in \mathrm{Mat}_2(\mathbb{R})$

  $\det A = a_{11}a_{22} - a_{12}a_{21}$;

- $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \in \mathrm{Mat}_3(\mathbb{R})$

  $\det A = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}$;

- $A \in \mathrm{Diag}_n(\mathbb{R})$ oppure $A \in \mathrm{Triang}_n(\mathbb{R})$

  $\det A = \Pi_{i \in I_n} a_{ii}$.

Definizione 2.28. *Sia $A = [a_{ij}] \in \mathrm{Mat}_n(\mathbb{R})$. Chiamiamo complemento algebrico dell'elemento a_{ij} lo scalare $c_{ij} = (-1)^{i+j} \det A_{ij}$, e matrice dei complementi algebrici la matrice $[c_{ij}] \in \mathrm{Mat}_n(\mathbb{R})$.*

Teorema 2.29 (I Teorema di Laplace)**.** *Per ogni $A = [a_{ij}] \in \mathrm{Mat}_n(\mathbb{R})$, il determinante di A può essere sviluppato rispetto a qualunque riga o colonna:*

- *rispetto alla i-esima riga:* $\det A = \sum_{j \in I_n} a_{ij} c_{ij}$,
- *rispetto alla j-esima colonna:* $\det A = \sum_{i \in I_n} a_{ij} c_{ij}$,

dove c_{ij} è il complemento algebrico dell'elemento a_{ij}.

Esercizio 2.30. *Siano*

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

Calcolare il determinante di A e B, e tutti i complementi algebrici di A.

Il determinante gode delle seguenti proprietà.

Proposizione 2.31. *Siano $A = [a_{ij}] \in \mathrm{Mat}_n(\mathbb{R})$. Allora*

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

(ii) se B è la matrice ottenuta moltiplicando una riga o una colonna di A per $\alpha \in \mathbb{R}$, allora $\det B = \alpha \det A$;

(iii) $A = [A_1|...|A_n]$ (A_j indica la j-esima colonna di A), $C \in \mathrm{Mat}_{n,1}(\mathbb{R})$, si ha

$$\det[A_1|...|A_j + C|...|A_n] = \det[A_1|...|A_j|...|A_n] + \det[A_1|...|C|...|A_n]$$

e analogamente se viene sommato un vettore riga;

(iv) se A ha una riga (o una colonna) nulla, si ha $\det A = 0$;

(v) se B è la matrice ottenuta scambiando di posto due righe (o due colonne) di A, allora $\det B = -\det A$;

(vi) se si sostituisce a una colonna la somma tra essa e un multiplo di un'altra, il determinante non varia (lo stesso vale rispetto alle righe).

Osservazione 2.32. *Se una matrice $A \in \mathrm{Mat}_n(\mathbb{R})$ ha due righe o due colonne proporzionali (linearmente dipendenti), allora* $\det A = 0$.

Infatti, scambiando queste righe o queste colonne, segue dalla Proposizione 2.31-(ii)-(v) che

$$\det A = -\det A \quad \Rightarrow \det A = 0.$$

Teorema 2.33 (Teorema di Binet)**.** *Siano* $A, B \in \mathrm{Mat}_n(\mathbb{R})$*. Allora*

$$\det(AB) = \det A \cdot \det B.$$

Esercizio 2.34. *Siano*

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 2 & 0 & 2 \\ 3 & 0 & 4 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 2 & 0 \\ 5 & 0 & 0 \\ 0 & 6 & 2 \end{bmatrix}.$$

verificare che vale il teorema di Binet.

2.5 Calcolo della matrice inversa

Teorema 2.35 (II Teorema di Laplace)**.** *Per ogni* $A \in \mathrm{Mat}_n(\mathbb{R})$*, sommando i prodotti degli elementi di una certa riga (o colonna) per i complementi algebrici dei corrispondenti elementi di un'altra riga (o colonna) si ottiene sempre zero.*

Definizione 2.36. *Una matrice* $A \in \mathrm{Mat}_n(\mathbb{R})$ *si dice singolare se* $\det A = 0$*, non singolare altrimenti.*

Proposizione 2.37. *Una matrice* $A \in \mathrm{Mat}_n(\mathbb{R})$ *è invertibile se e solo se* $\det A \neq 0$.

Dim. ($\Rightarrow$) Se A è invertibile, allora esiste la matrice inversa A^{-1} ed è tale che $AA^{-1} = \mathbb{I}_n$. Ricordando il Teorema di Binet, si ha

$$1 = \det(\mathbb{I}_n) = \det(AA^{-1}) = \det(A)\det(A^{-1}).$$

Dunque non può essere $\det(A) = 0$.
($\Leftarrow$) Supponiamo che sia $\det(A) \neq 0$. Consideriamo la matrice trasposta della matrice dei complementi algebrici di A, ovvero

$$B = [b_{ij}] = [c_{ji}], \quad \text{dove} \quad c_{ij} := (-1)^{i+j} \det A_{ij} \ \forall i, j \in I_n.$$

Ora consideriamo la matrice prodotto AB. Chiamando $A_1, A_2, ...A_n$ le colonne della matrice A, la j-esima colonna di AB è

$$A_1 b_{1j} + A_2 b_{2j} + \cdots + \cdots A_n b_{nj} = A_1 c_{j1} + A_2 c_{j2} + \cdots + \cdots A_n c_{jn}.$$

L'elemento i-esimo di tale vettore colonna è dunque

$$a_{i1}c_{j1} + a_{i2}c_{j2} + \cdots + \cdots a_{in}c_{jn}. \tag{2.1}$$

Caso $i = j$. L'elemento in (2.1) è

$$a_{i1}c_{i1} + a_{i2}c_{i2} + \cdots + \cdots a_{in}c_{in} = \det A$$

per il I Teorema di Laplace.
Caso $i \neq j$. L'elemento in (2.1) è

$$a_{i1}c_{j1} + a_{i2}c_{j2} + \cdots + \cdots a_{in}c_{jn} = 0$$

per il II Teorema di Laplace (stiamo sommando i prodotti degli elementi della riga i-esima di A per i complementi algebrici dei corrispondenti elementi della riga j-esima di A, con $j \neq j$).

Segue che la j-esima colonna della matrice AB è un vettore che ha un unico elemento non nullo, uguale a $\det A$, al posto j. Indicando con D_j la colonna j-esima di AB, si ha quindi $D_j = \det A\, \mathbf{e}_j$, ovvero

$$AB = \det A\, \mathbb{I}_n.$$

Poiché per ipotesi $\det A \neq 0$, possiamo dividere per $\det A$ e ottenere

$$A \frac{B}{\det A} = \mathbb{I}_n.$$

Si dimostra analogamente che $\frac{B}{\det A} A = \mathbb{I}_n$, e dunque $A^{-1} = \frac{1}{\det A} B$. □

La dimostrazione della Proposizione 2.37 fornisce un metodo per il calcolo della matrice inversa, quando questa esiste.

Teorema 2.38. *Sia* $A \in \mathrm{Mat}_n(\mathbb{R})$ *con* $\det A \neq 0$. *Allora*

$$A^{-1} = \frac{1}{\det A} C^T,$$

dove C *è la matrice dei complementi algebrici di* A *introdotta nella Definizione 2.28.*

Proposizione 2.39. *Se* $A \in \mathrm{Mat}_n(\mathbb{R})$ *è invertibile, allora* $\det(A^{-1}) = (\det A)^{-1}$.

Osservazione 2.40. *Sia*

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \in \mathrm{Mat}_2(\mathbb{R}).$$

Allora

$$A^{-1} = \frac{1}{a_{11}a_{22} - a_{12}a_{21}} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}.$$

Esempio 2.41. *Calcolare, se esiste, l'inversa della matrice*

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

2.6 Rango di una matrice

Definizione 2.42. *Data una matrice* $A \in \mathrm{Mat}_{m,n}(\mathbb{R})$*, si dice minore di* A *di ordine* p *(con* $p \leq \min\{m, n\}$*) una matrice quadrata di ordine* p*, ottenuta da* A *sopprimendo* $m-p$ *righe e* $n-p$ *colonne.*

Esercizio 2.43. *Scrivere i minori di ordine* 3 *della matrice*

$$A = \begin{bmatrix} 1 & 2 & 3 & 0 \\ 4 & 5 & 0 & 1 \\ 2 & 6 & 3 & 4 \end{bmatrix}.$$

Definizione 2.44. *Data una matrice* $A \in \mathrm{Mat}_{m,n}(\mathbb{R})$*, si dice rango di* A*, e si indica con* $\mathrm{rg}(A)$*, il massimo ordine di un minore non singolare di* A*.*

Osservazione 2.45. $\mathrm{rg}(A) = p$ *se e solo*

1. *esiste un minore non singolare di* A *di ordine* p
2. *tutti i minori di* A *di ordine* $k > p$ *sono singolari*

Osservazione 2.46. *Data* $A \in \mathrm{Mat}_n(\mathbb{R})$*,* $\mathrm{rg}(A)$ *coincide con il massimo numero di linee (righe o colonne) linearmente indipendenti di* A*.*

Infatti il rango di A *è per definizione il massimo ordine di un minore non singolare di* A*. D'atra parte, per l'Osservazione 2.32, se una matrice ha due righe o due colonne proporzionali (linearmente dipendenti), allora essa ha determinante nullo.*

Esempio 2.47. *Calcolare il rango della matrice*

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

usando la definizione.

Teorema 2.48 (Teorema degli orlati)**.** *Una matrice* $A \in \mathrm{Mat}_{m,n}(\mathbb{R})$ *ha rango* p *se e solo se*

1. *esiste un minore* M *di* A *di ordine* p *non singolare*
2. *sono singolari tutti i minori di ordine* $p+1$ *che contengono* M*.*

Esempio 2.49. *Calcolare il rango della matrice*

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

usando il Teorema degli orlati.

Capitolo 3

Applicazioni lineari

Definizione 3.1. *Siano V e W spazi vettoriali su $\mathbb{R}$. Una funzione $f : V \to W$ si dice applicazione lineare (o trasformazione lineare) se*

1. $f(v_1 + v_2) = f(v_1) + f(v_2)$ *(proprietà di additività)*
2. $f(\alpha v) = \alpha f(v)$ *(proprietà di omogeneità)*

per ogni v, v_1, $v_2 \in V$ e per ogni $\alpha \in \mathbb{R}$.

Osservazione 3.2. *Si noti che nella definizione di applicazione trasformazione lineare si utilizza la struttura (le operazioni) di spazio vettoriale, sia nel dominio che nel codominio di f.*

Osservazione 3.3. *Come dirette conseguenze della definizione si ha*

(i) $f(0_V) = 0_W$,

(ii) $f(-v) = -f(v)$ *per ogni* $v \in V$.

(iii) Per ogni combinazione lineare $\alpha_1 v_1 + \cdots + \alpha_k v_k$ di k vettori in V si ha

$$f(\alpha_1 v_1 + \cdots + \alpha_k v_k) = \alpha_1 f(v_1) + \cdots + \alpha_k f(v_k).$$

Esempio 3.4. *L'applicazione $f : \mathbb{R}^3 \to \mathbb{R}^3$ definita da*

$$f\left(\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}\right) = \begin{pmatrix} x_1 + 3x_2 + 5x_3 \\ 2x_2 + x_3 \\ x_2 + 2x_3 \end{pmatrix}$$

è lineare. Infatti

1. *per ogni* $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \in \mathbb{R}^3,$

$$f\left(\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}\right) = f\left(\begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ x_3 + y_3 \end{pmatrix}\right) = \begin{pmatrix} x_1 + y_1 + 3x_2 + 3y_2 + 5x_3 + 5y_3 \\ 2x_2 + 2y_2 + x_3 + y_3 \\ x_2 + y_2 + 2x_3 + 2y_3 \end{pmatrix}$$

$$f\left(\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}\right) + f\left(\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}\right) = \begin{pmatrix} x_1 + 3x_2 + 5x_3 \\ 2x_2 + x_3 \\ x_2 + 2x_3 \end{pmatrix} + \begin{pmatrix} y_1 + 3y_2 + 5y_3 \\ 2y_2 + y_3 \\ y_2 + 2y_3 \end{pmatrix}$$
$$= \begin{pmatrix} x_1 + y_1 + 3x_2 + 3y_2 + 5x_3 + 5y_3 \\ 2x_2 + 2y_2 + x_3 + y_3 \\ x_2 + y_2 + 2x_3 + 2y_3 \end{pmatrix}$$

2. *per ogni* $\alpha \in \mathbb{R}$, $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3$,

$$f\left(\alpha \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}\right) = f\left(\begin{pmatrix} \alpha x_1 \\ \alpha x_2 \\ \alpha x_3 \end{pmatrix}\right) = \begin{pmatrix} \alpha x_1 + 3\alpha x_2 + 5\alpha x_3 \\ 2\alpha x_2 + \alpha x_3 \\ \alpha x_2 + 2\alpha x_3 \end{pmatrix}$$

$$\alpha f\left(\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}\right) = \alpha \begin{pmatrix} x_1 + 3x_2 + 5x_3 \\ 2x_2 + x_3 \\ x_2 + 2x_3 \end{pmatrix} = \begin{pmatrix} \alpha x_1 + 3\alpha x_2 + 5\alpha x_3 \\ 2\alpha x_2 + \alpha x_3 \\ \alpha x_2 + 2\alpha x_3 \end{pmatrix}.$$

Esempio 3.5. *L'applicazione* $g : \mathbb{R}^3 \to \mathbb{R}^3$ *definita da*

$$g\left(\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}\right) = \begin{pmatrix} x_1 + x_2 \\ 2x_3 \\ x_1^2 - x_2^2 \end{pmatrix}$$

non è lineare. Infatti, per esempio, non vale (1): se $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}$, $\mathbf{v}_2 = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}$, *si ha*

$$g(\mathbf{v}_1 + \mathbf{v}_2) = g\left(\begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}\right) = \begin{pmatrix} 4 \\ 4 \\ -8 \end{pmatrix}$$

mentre

$$g(\mathbf{v}_1) + g(\mathbf{v}_2) = \begin{pmatrix} 3 \\ 0 \\ -3 \end{pmatrix} + \begin{pmatrix} 1 \\ 4 \\ -1 \end{pmatrix} = \begin{pmatrix} 4 \\ 4 \\ -4 \end{pmatrix}.$$

Esempio 3.6. *Nello spazio vettoriale* $\mathbb{P}$ *dei polinomi nella indeterminata* t *a coefficienti reali, l'applicazione*

$$d : \mathbb{P} \longrightarrow \mathbb{P}$$
$$p(t) \mapsto p'(t)$$

dove $p'(t)$ *è la derivata (formale) di* $p(t)$ *(cioè, se* $p(t) = a_0 + a_1 t + \cdots + a_n t^n$*, si ha* $p'(t) = a_1 + 2a_2 + \cdots + n a_n t^{n-1}$*) è un'applicazione lineare. Infatti*

1. $\forall p(t), q(t) \in \mathbb{P} \Rightarrow (p + q)'(t) = p'(t) + q'(t)$

2. $\forall \alpha \in \mathbb{R}, p(t) \in \mathbb{P} \Rightarrow (\alpha p)'(t) = \alpha p'(t)$.

Definizione 3.7. *Sia $f : V \to W$ un'applicazione lineare. Definiamo*

$$\operatorname{Ker}(f) = \{v \in V : \ f(v) = 0_W\} \quad \text{nucleo di } f$$
$$\operatorname{Im}(f) = \{w \in W : \exists\, v \in V \text{ tale che } w = f(v)\} \quad \text{immagine di } f$$

Osservazione 3.8. *Si ha* $\operatorname{Ker}(f) \subseteq V$ *e* $\operatorname{Im}(f) \subseteq W$.

Proposizione 3.9. *Sia $f : V \to W$ un'applicazione lineare. Allora* $\operatorname{Ker}(f)$ *è un sottospazio di V, e* $\operatorname{Im}(f)$ *è un sottospazio di W.*

Dim. Proviamo dapprima che $\operatorname{Ker}(f)$ è un sottospazio di V.

1. $0_V \in \operatorname{Ker}(f)$ perché $f(0_V) = 0_W$
2. $\forall v_1, v_2 \in \operatorname{Ker}(f) \Rightarrow v_1 + v_2 \in \operatorname{Ker}(f)$ perché
$$f(v_1 + v_2) = f(v_1) + f(v_2) = 0_W + 0_W = 0_W.$$
3. $\forall \alpha \in \mathbb{R}, v \in \operatorname{Ker}(f) \Rightarrow \alpha v \in \operatorname{Ker}(f)$ perché
$$f(\alpha v) = \alpha f(v) = \alpha 0_W = 0_W.$$

Proviamo ora che $\operatorname{Im}(f)$ è un sottospazio di W.

1. $0_W \in \operatorname{Im}(f)$ perché $0_W = f(0_V)$
2. $\forall w_1, w_2 \in \operatorname{Im}(f) \Rightarrow w_1 + w_2 \in \operatorname{Im}(f)$. Infatti, essendo $w_1, w_2 \in \operatorname{Im}(f)$, esistono $v_1, v_2 \in V$ tali che
$$w_1 = f(v_1)$$
$$w_2 = f(v_2).$$
Allora
$$w_1 + w_2 = f(v_1) + f(v_2) = f(v_2) = f(v_1 + v_2), \quad v_1, v_2 \in V.$$
3. $\forall \alpha \in \mathbb{R}, w \in \operatorname{Im}(f) \Rightarrow \alpha w \in \operatorname{Im}(f)$. Infatti, essendo $w \in \operatorname{Im}(f)$, esiste $v \in V$ tale che
$$w = f(v).$$
Allora
$$\alpha w = \alpha f(v) = f(\alpha v), \quad v \in V.$$

□

Applicazioni lineari notevoli.

1. *Applicazione identità*:

$$\begin{aligned} 1_V : V &\to V \\ v &\mapsto v \end{aligned}$$

2. *Applicazione composta*. Siano $f : V \to W$ e $g : W \to U$ due applicazioni lineari. L'applicazione composta di f con g è definita come

$$\begin{aligned} g \circ f : &V \to U \\ &v \mapsto (g \circ f)(v) = g(f(v)) \end{aligned}$$

3. *Applicazione inversa*. Sia $f : V \to V$ un'applicazione lineare. f si dice invertibile, se esiste $g : V \to V$ (anch'essa lineare) tale che

$$f \circ g = g \circ f = 1_V.$$

In tal caso $g = f^{-1}$ è detta l'applicazione inversa di f.

Si ha il seguente risultato, che diamo senza dimostrazione.

Proposizione 3.10. *Un'applicazione lineare $f : \mathbb{R}^n \to \mathbb{R}^m$ è invertibile se vale (almeno) una delle due seguenti proprietà*

(i) $\mathrm{Ker}(f) = \{\mathbf{0}_{\mathbb{R}^n}\}$,

(ii) $\mathrm{Im}(f) = \mathbb{R}^m$.

3.1 Matrice associata ad un'applicazione lineare

Siano V e W spazi vettoriali su $\mathbb{R}$, con $\dim(V) = n$ e $\dim(W) = m$. Fissiamo una base $\mathcal{B}_V = \{v_1, v_2, ..., v_n\}$ in V e una base $\mathcal{B}_W = \{w_1, w_2, ..., w_m\}$ in W.

Sia $f : V \to W$ un'applicazione lineare. Allora $f(v_1)$ è un vettore di W, che si scrive pertanto in modo unico come combinazione lineare dei vettori di $\mathcal{B}_W$ (si veda il Teorema 1.30):

$$f(v_1) = a_{11}w_1 + a_{21}w_2 + \cdots + a_{m1}w_m$$

per un'unica scelta di a_{11}, a_{21}, ..., a_{m1}. Lo stesso vale per $f(v_2)$, ..., $f(v_n)$, cioè

$$\begin{aligned} f(v_2) &= a_{12}w_1 + a_{22}w_2 + \cdots + a_{m2}w_m \\ &\vdots \\ f(v_2) &= a_{1n}w_1 + a_{2n}w_2 + \cdots + a_{mn}w_m. \end{aligned}$$

Definizione 3.11. *Si definisce matrice associata a f rispetto alle basi (ordinate) $\mathcal{B}_V$ e $\mathcal{B}_W$ la matrice*

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \in \mathrm{Mat}_{m,n}(\mathbb{R}).$$

Osservazione 3.12. *Siano V e W spazi vettoriali su $\mathbb{R}$, con $\dim(V) = n$ e $\dim(W) = m$, e siano $\mathcal{B}_V = \{v_1, v_2, ..., v_n\}$ e $\mathcal{B}_W = \{w_1, w_2, ..., w_m\}$ rispettivamente una base in V e una base in W. La matrice di rappresentazione di una applicazione lineare $f : V \to W$ è la matrice le cui colonne sono le immagini dei vettori $\{v_1, ..., v_n\}$ attraverso f, ovvero*

$$A = \left[\ f(v_1)\,|\, f(v_2)\,| \cdots |f(v_n)\ \right] \in \mathrm{Mat}_{m,n}(\mathbb{R}).$$

Esempio 3.13. *Sia $f : \mathbb{R}^3 \to \mathbb{R}^2$ l'applicazione lineare definita da*

$$f\left(\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}\right) = \begin{pmatrix} 2x_1 - x_2 \\ x_2 + x_3 \end{pmatrix}.$$

Fissiamo in $\mathbb{R}^3$ e in $\mathbb{R}^2$ le basi canoniche date rispettivamente da

$$\mathcal{U}_{\mathbb{R}^3} = \left\{ \mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\},$$
$$\mathcal{U}_{\mathbb{R}^2} = \left\{ \mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}.$$

Si ha

$$f(\mathbf{e}_1) = f\left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}\right) = \begin{pmatrix} 2 \\ 0 \end{pmatrix} = 2\mathbf{e}_1 + 0\mathbf{e}_2,$$
$$f(\mathbf{e}_2) = f\left(\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}\right) = \begin{pmatrix} -1 \\ 1 \end{pmatrix} = -\mathbf{e}_1 + \mathbf{e}_2,$$
$$f(\mathbf{e}_3) = f\left(\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right) = \begin{pmatrix} 0 \\ 1 \end{pmatrix} = 0\mathbf{e}_1 + \mathbf{e}_2.$$

La matrice associata a f rispetto alle basi $\mathcal{U}_{\mathbb{R}^3}$ e $\mathcal{U}_{\mathbb{R}^2}$ è

$$A = \begin{bmatrix} 2 & -1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \in \mathrm{Mat}_{2,3}(\mathbb{R}).$$

Ritorniamo alla situazione generale. Se $v \in V$, essendo $\mathcal{B}_V = \{v_1, v_2, ..., v_n\}$ una base di V, possiamo scrivere

$$v = b_1 v_1 + b_2 v_2 + \cdots + b_n v_n$$

per un'unica scelta di $b_1, b_2, ..., b_n \in \mathbb{R}$.

D'altra parte, se $f : V \to W$ è un'applicazione lineare, allora $f(v_1) \in W$ si scrive in modo unico come combinazione lineare dei vettori di $\mathcal{B}_W = \{w_1, w_2, ..., w_m\}$:

$$f(v_1) = a_{11}w_1 + a_{21}w_2 + \cdots + a_{m1}w_m.$$

Essendo f un'applicazione lineare, abbiamo dunque

$$f(v) = f\left(\sum_{j=1}^{n} b_j v_j\right) = \sum_{j=1}^{n} f(b_j v_j) = \sum_{j=1}^{n} b_j f(v_j) = \sum_{j=1}^{n} b_j \sum_{i=1}^{m} a_{ij} w_i.$$

Questo risultato è formalizzato nella seguente proposizione.

Proposizione 3.14. *Siano V e W spazi vettoriali su $\mathbb{R}$, con $\dim(V) = n$ e $\dim(W) = m$. Fissiamo una base $\mathcal{B}_V = \{v_1, v_2, ..., v_n\}$ in V e una base $\mathcal{B}_W = \{w_1, w_2, ..., w_m\}$ in W. Sia $f : V \to W$ un'applicazione lineare.*
Se v ha coordinate $\begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$ rispetto alla base $\mathcal{B}_V$, allora $f(v)$ ha coordinate $A\begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$ rispetto alla base $\mathcal{B}_W$.

Esempio 3.15. *Sia $f : \mathbb{R}^3 \to \mathbb{R}^2$ l'applicazione lineare dell'Esempio 3.13. Il vettore $\mathbf{v} = 2\mathbf{e}_1 + \mathbf{e}_2 - \mathbf{e}_2$ ha coordinate*

$$\begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix}$$

rispetto alla base $\mathcal{U}_{\mathbb{R}^3} = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$. Allora $f(\mathbf{v})$ ha coordinate

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

rispetto alla base $\mathcal{U}_{\mathbb{R}^2} = \{\mathbf{e}_1, \mathbf{e}_2\}$, cioè $f(\mathbf{v}) = 3\mathbf{e}_1 + 0\mathbf{e}_2 = 3\mathbf{e}_1$.

Abbiamo visto che a $f : V \to W$ applicazione lineare (dove $\dim(V) = n$ e $\dim(W) = m$) possiamo associare una matrice $A \in \mathrm{Mat}_{m,n}(\mathbb{R})$. Viceversa, se $\bar{A}$ è una matrice in $\mathrm{Mat}_{m,n}(\mathbb{R})$, ad essa è associata un'applicazione lineare $\bar{f} : V \to W$ definita da

$$f\left(\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}\right) = \bar{A}\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix},$$

dove $\begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$ sono le coordinate di un generico vettore v rispetto alla base $\mathcal{B}_V$. Questo risultato viene formalizzato nella seguente proposizione (data nel caso $V = \mathbb{R}^n$ e $W = \mathbb{R}^m$).

Proposizione 3.16. *Vi è una corrispondenza biunivoca tra l'insieme delle applicazioni lineari $f : \mathbb{R}^n \to \mathbb{R}^m$ e l'insieme delle matrici $\mathrm{Mat}_{m,n}(\mathbb{R})$.*

Teorema 3.17 (Teorema della nullità + rango)**.** *Sia $f : V \to W$ un'applicazione lineare. Allora*

$$\dim(V) = \dim(\mathrm{Ker}(f)) + \dim(\mathrm{Im}(f)).$$

Osservazione 3.18. *Siano $\mathcal{B}_V = \{v_1, v_2, ..., v_m\}$ base di V, $\mathcal{B}_W$ base di W, e $f : V \to W$ applicazione lineare. Allora*

(i) Si ha

$$\mathrm{Im}(f) = \mathrm{span}\{f(v_1), f(v_2), ..., f(v_n)\}. \tag{3.1}$$

Infatti, se $y \in \mathrm{Im}(f)$, esiste $v \in V$ tale che $y = f(v)$. D'altra parte,

$$v = b_1 v_1 + \cdots + b_n v_n$$

per qualche $b_1, ..., b_n \in \mathbb{R}$. Pertanto,

$$y = b_1 f(v_1) + \cdots + b_n f(v_n),$$

ovvero ogni $y \in \mathrm{Im}(f)$ è combinazione lineare di $\{f(v_1), ..., f(v_n)\}$.

(ii) Se A è la matrice associata a f, $A = [f(v_1)|f(v_2)| \cdots |f(v_n)]$ e dunque da (3.1)

$$\dim(\mathrm{Im}(f)) = \mathrm{rg}(A). \tag{3.2}$$

(iii) Posto $n = \dim(V)$, il Teorema 3.17 assieme a (3.1)-(3.2) ci dice che

$$n = \dim(\mathrm{Ker}(f)) + \mathrm{rg}(A). \tag{3.3}$$

Esempio 3.19. *Sia $f : \mathbb{R}^3 \to \mathbb{R}^2$ definita da*

$$f\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = \begin{pmatrix} 2x \\ y+z \end{pmatrix}.$$

Consideriamo in $\mathbb{R}^3$ la base $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ e in $\mathbb{R}^2$ la base $\mathcal{U}_{\mathbb{R}^2} = \{\mathbf{e}_1, \mathbf{e}_2\}$, dove

$$\mathbf{v}_1 = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \mathbf{v}_3 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix},$$
$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Allora

$$f(\mathbf{v}_1) = \begin{pmatrix} 4 \\ 1 \end{pmatrix} = 4\mathbf{e}_1 + \mathbf{e}_2$$
$$f(\mathbf{v}_2) = \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \mathbf{e}_2$$

$$f(\mathbf{v}_3) = \begin{pmatrix} 2 \\ 1 \end{pmatrix} = 2\mathbf{e}_1 + \mathbf{e}_2,$$

e quindi $A = \begin{bmatrix} 4 & 0 & 2 \\ 1 & 1 & 1 \end{bmatrix}$. *Si ha*

$$\mathrm{Im}(f) = \mathrm{span}\left\{ \begin{pmatrix} 4 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix} \right\}.$$

Inoltre $\dim(\mathrm{Im}(f)) = \mathrm{rg}(A) = 2$, *e per il Teorema della nullità + rango nella versione* (3.3),

$$\dim(\mathrm{Ker}(f)) = n - \mathrm{rg}(A) = 3 - 2 = 1.$$

Una base per $\mathrm{Im}(f)$ *è data da*

$$\mathcal{B}_{\mathrm{Im}(f)} = \left\{ \begin{pmatrix} 4 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}.$$

Matrici associate ad applicazioni lineari notevoli.

1. *Applicazione identità* $\leftrightarrow$ *Matrice identità.* Fissata una base $\mathcal{B}$ di V, con $\dim(V) = n$, la matrice associata all'applicazione identità 1_V su V è la matrice identità
$$\mathbb{I}_n.$$

2. *Applicazione composta* $\leftrightarrow$ *Matrice prodotto.* Siano $f : V \to W$ e $g : W \to U$ due applicazioni lineari. Fissiamo $\mathcal{B}_V$ base di V, $\mathcal{B}_W$ base di W, B_U base di U. Siano A la matrice associata ad f (rispetto alle basi $\mathcal{B}_V$ e $\mathcal{B}_W$) e B la matrice associata a g (rispetto alle basi $\mathcal{B}_W$ e $\mathcal{B}_U$). Allora la matrice associata all'applicazione composta $g \circ f$, rispetto alle basi $\mathcal{B}_V$ e $\mathcal{B}_U$, è la matrice prodotto
$$BA.$$

3. *Applicazione inversa* $\leftrightarrow$ *Matrice inversa.* Sia $f : V \to V$ un'applicazione lineare invertibile. Fissiamo $\mathcal{B}$ base di V. Se A è la matrice associata a f rispetto alla base $\mathcal{B}$, allora la matrice associata all'applicazione inversa f^{-1} (sempre rispetto alla base $\mathcal{B}$) è la matrice inversa
$$A^{-1}.$$

3.2 Cambiamenti di base nello spazio $\mathbb{R}^n$

In $\mathbb{R}^n$, sia $\mathcal{U} = \{\mathbf{e}_1, ..., \mathbf{e}_n\}$ la base canonica e sia $\mathcal{B}' = \{\mathbf{v}_1, ..., \mathbf{v}_n\}$ un'altra base. Dato un vettore $\mathbf{x}$, avremo

$$\mathbf{x} = x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n,$$
$$\mathbf{x} = y_1\mathbf{v}_1 + \cdots + y_n\mathbf{v}_n.$$

Problema 1: se conosco la scrittura di $\mathbf{x}$ rispetto alla base canonica $\mathcal{U}$, come posso ottenere la scrittura di $\mathbf{x}$ rispetto alla base $\mathcal{B}'$?

Consideriamo la trasformazione lineare $f : \mathbb{R}^n \to \mathbb{R}^n$ tale che $f(\mathbf{e}_j) = \mathbf{v}_j$, $j = 1, ..., n$. Abbiamo

$$f\left(\sum_{j=1}^{n} y_j \mathbf{v}_j\right) = f\left(\sum_{j=1}^{n} x_j \mathbf{e}_j\right) = \sum_{j=1}^{n} x_j f(\mathbf{e}_j) = \sum_{j=1}^{n} x_j \mathbf{v}_j.$$

Se chiamiamo A la matrice di rappresentazione di f, le colonne di A sono i vettori della base $\mathcal{B}'$. La matrice A fornisce il cambio di base dalla base $\mathcal{B}'$ alla base fondamentale $\mathcal{U}$: data la rappresentazione $(y_1, ..., y_n)$ di un vettore $\mathbf{x}$ nella base $\mathcal{B}'$, $A\mathbf{x}$ fornisce la rappresentazione del vettore $(x_1, ..., x_n)$ nella base fondamentale $\mathcal{U}$.

Tutto questo giustifica la seguente definizione.

Definizione 3.20. *Siano $\mathcal{B} = \{v_1, ..., v_n\}$ e $\mathcal{B}' = \{v'_1, ..., v'_n\}$ due basi di $\mathbb{R}^n$. Chiamiamo matrice di cambio di base da $\mathcal{B}'$ a $\mathcal{B}$ la matrice $P_{\mathcal{B}\mathcal{B}'}$ tale che, se $(y_1, ..., y_n)$ è la rappresentazione di un qualunque vettore nella base $\mathcal{B}'$, allora $P_{\mathcal{B}\mathcal{B}'}(y_1, ..., y_n)^T$ è la rappresentazione del vettore nella base $\mathcal{B}$.*

Osservazione 3.21. *La matrice A ottenuta in precedenza disponendo in colonna i vettori della base $\mathcal{B}'$ non è altro che $P_{\mathcal{U}\mathcal{B}'}$, cioè la matrice di cambio di base dalla base $\mathcal{B}'$ alla base $\mathcal{U}$.*

Osservazione 3.22. *La matrice $P_{\mathcal{U}\mathcal{B}'}$ è invertibile dato che le sue colonne sono linearmente indipendenti. Senza ulteriori dettagli, diciamo che $P_{\mathcal{B}'\mathcal{U}} = P^{-1}_{\mathcal{U}\mathcal{B}'}$, e cioè che la matrice di cambio di base dalla base fondamentale $\mathcal{U}$ alla base $\mathcal{B}'$ è la matrice inversa dell'altra.*

Si ha dunque il seguente risultato.

Proposizione 3.23. *Sia $\mathcal{U}$ la base canonica su $\mathbb{R}^n$, e siano $\mathcal{B} = \{v_1, ..., v_n\}$ e $\mathcal{B}' = \{v'_1, ..., v'_n\}$ altre due basi di $\mathbb{R}^n$. Allora:*

- *la matrice di cambio di base $P_{\mathcal{U}\mathcal{B}}$ (rispettivamente $P_{\mathcal{U}\mathcal{B}'}$) dalla base $\mathcal{B}$ alla base $\mathcal{U}$ (rispettivamente dalla base $\mathcal{B}'$ alla base $\mathcal{U}$) è la matrice ottenuta disponendo in colonna i vettori della base $\mathcal{B}$ (rispettivamente della base $\mathcal{B}'$);*
- *la matrice di cambio di base $P_{\mathcal{B}\mathcal{B}'}$ da $\mathcal{B}'$ a $\mathcal{B}$ è data da*

$$P_{\mathcal{B}\mathcal{B}'} = P^{-1}_{\mathcal{U}\mathcal{B}} P_{\mathcal{U}\mathcal{B}'}. \tag{3.4}$$

Esempio 3.24. *Consideriamo in $\mathbb{R}^2$ la base canonica $\mathcal{U} = \{\mathbf{e}_1, \mathbf{e}_2\}$ e la base $\mathcal{B}' = \{\mathbf{v}_1, \mathbf{v}_2\}$, con $\mathbf{v}_1 = \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$, $\mathbf{v}_2 = \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix}$. Sia $\mathbf{x}$ il vettore di coordinate $\begin{pmatrix} 0 \\ -1 \end{pmatrix}$ rispetto alla base $\mathcal{B}'$. Vogliamo scrivere il vettore $\mathbf{x}$ rispetto alla base canonica $\mathcal{U}$.*

La matrice di cambio di base da $\mathcal{B}'$ *a* $\mathcal{U}$ *è*

$$P_{\mathcal{U}\mathcal{B}'} = \begin{bmatrix} -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}.$$

Il vettore $\mathbf{x}$ *rispetto alla base canonica* $\mathcal{U}$ *è indicato dalla coppia di coordinate*

$$P_{\mathcal{U}\mathcal{B}'} \begin{pmatrix} 0 \\ -1 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}.$$

Esempio 3.25. *Consideriamo in* $\mathbb{R}^2$ *le basi* $\mathcal{B} = \{\mathbf{w}_1, \mathbf{w}_2\}$ *e* $\mathcal{B}' = \{\mathbf{v}_1, \mathbf{v}_2\}$*, con*

$$\mathbf{w}_1 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \quad \mathbf{w}_2 = \begin{pmatrix} -1 \\ -1 \end{pmatrix},$$
$$\mathbf{v}_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Sia $\mathbf{x}$ *il vettore dato da*

$$\mathbf{x} = \mathbf{v}_1 + \mathbf{v}_2 = -\frac{1}{2}(\mathbf{w}_1 + \mathbf{w}_2) = -\frac{1}{2}\mathbf{w}_1 - \frac{1}{2}\mathbf{w}_2.$$

Le coordinate di $\mathbf{x}$ *sono* $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ *rispetto alla base* $\mathcal{B}'$ *e* $\begin{pmatrix} -\frac{1}{2} \\ -\frac{1}{2} \end{pmatrix}$ *rispetto alla base* $\mathcal{B}$.

Usiamo ora la formula (3.4): si ha

$$\begin{aligned} P_{\mathcal{B}\mathcal{B}'} &= P_{\mathcal{U}\mathcal{B}}^{-1} P_{\mathcal{U}\mathcal{B}'} \\ &= \begin{bmatrix} -1 & -1 \\ 1 & -1 \end{bmatrix}^{-1} \begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix} \\ &= \frac{1}{2} \begin{bmatrix} -1 & 1 \\ -1 & -1 \end{bmatrix}^{-1} \begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix} \\ &= \frac{1}{2} \begin{bmatrix} 1 & -2 \\ -1 & 0 \end{bmatrix}. \end{aligned}$$

Abbiamo dunque

$$P_{\mathcal{B}\mathcal{B}'} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{1}{2} \begin{bmatrix} 1 & -2 \\ -1 & 0 \end{bmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} -1 \\ -1 \end{pmatrix} = \begin{pmatrix} -\frac{1}{2} \\ -\frac{1}{2} \end{pmatrix}.$$

Infine

$$P_{\mathcal{B}'\mathcal{B}} = (P_{\mathcal{B}\mathcal{B}'})^{-1} = \begin{bmatrix} 0 & -2 \\ -1 & -1 \end{bmatrix}$$

realizza il cambio di base da $\mathcal{B}$ *a* $\mathcal{B}'$.

Problema 2. Sia $f : \mathbb{R}^n \to \mathbb{R}^n$ un'applicazione lineare. Fissata una base $\mathcal{B} = \{\mathbf{v}_1, ..., \mathbf{v}_n\}$, abbiamo visto che possiamo associare ad f una matrice $A_{\mathcal{B}} \in \mathrm{Mat}_n(\mathbb{R})$ (relativamente a $\mathcal{B}$ come base sia del dominio che del codominio).
Consideriamo ora un'altra base $\mathcal{B}' = \{\mathbf{v}'_1, ..., \mathbf{v}'_n\}$ di $\mathbb{R}^n$. Come prima possiamo associare ad f una matrice $A_{\mathcal{B}'} \in \mathrm{Mat}_n(\mathbb{R})$ (relativamente a $\mathcal{B}'$ come base sia del dominio che del codominio). Che relazione c'è tra le matrici $A_{\mathcal{B}}$ e $A_{\mathcal{B}'}$?

Sia $A_{\mathcal{B}} = [f(\mathbf{v}_1)|f(\mathbf{v}_2)|\cdots|f(\mathbf{v}_n)]$ la matrice di rappresentazione di f rispetto a $\mathcal{B} = \{\mathbf{v}_1, ..., \mathbf{v}_n\}$ come base sia del dominio che del codominio.
Se $\mathbf{y} = (y_1, ..., y_n)^T$ è il vettore delle coordinate di $\mathbf{v} \in \mathbb{R}^n$ rispetto alla base $\mathcal{B}'$, allora

- $P_{\mathcal{B}\mathcal{B}'}\mathbf{y}$ è il vettore delle coordinate di $\mathbf{v}$ rispetto alla base $\mathcal{B}$,
- $A_{\mathcal{B}}P_{\mathcal{B}\mathcal{B}'}\mathbf{y}$ è l'immagine di $\mathbf{v}$ attraverso l'applicazione f, scritta rispetto alla base $\mathcal{B}$,
- $P_{\mathcal{B}'\mathcal{B}}A_{\mathcal{B}}P_{\mathcal{B}\mathcal{B}'}\mathbf{y} = P^{-1}_{\mathcal{B}\mathcal{B}'}A_{\mathcal{B}}P_{\mathcal{B}\mathcal{B}'}\mathbf{y}$ è l'immagine di $\mathbf{v}$ attraverso l'applicazione f, scritta rispetto alla base $\mathcal{B}'$.

Si ha dunque il seguente risultato.

Proposizione 3.26. *Siano $\mathcal{B} = \{\mathbf{v}_1, ..., \mathbf{v}_n\}$ e $\mathcal{B}' = \{\mathbf{v}'_1, ..., \mathbf{v}'_n\}$ due basi di $\mathbb{R}^n$, e $f : \mathbb{R}^n \to \mathbb{R}^n$ un'applicazione lineare. Sia $A_{\mathcal{B}} \in \mathrm{Mat}_n(\mathbb{R})$ la matrice associata ad f relativamente a $\mathcal{B}$ come base sia del dominio che del codominio.*
Allora la matrice $A_{\mathcal{B}'} \in \mathrm{Mat}_n(\mathbb{R})$ associata a $\mathcal{B}'$ come base sia del dominio che del codominio è data da

$$A_{\mathcal{B}'} = P^{-1}_{\mathcal{B}\mathcal{B}'}A_{\mathcal{B}}P_{\mathcal{B}\mathcal{B}'}, \tag{3.5}$$

dove $P_{\mathcal{B}\mathcal{B}'}$ è la matrice la cui j-esima colonna ha le coordinate di v'_j rispetto a $\mathcal{B}$.

Esempio 3.27. *Sia $f : \mathbb{R}^2 \to \mathbb{R}^2$ definita da*

$$f\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} x+2y \\ x-y \end{pmatrix}.$$

Siano $\mathcal{U} = \{\mathbf{e}_1, \mathbf{e}_2\}$ la base canonica e $\mathcal{B}' = \{\mathbf{v}'_1, \mathbf{v}'_2\}$ con $\mathbf{v}'_1 = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$, $\mathbf{v}'_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Si ha

$$f(\mathbf{e}_1) = f\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) = \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \mathbf{e}_1 + \mathbf{e}_2$$
$$f(\mathbf{e}_2) = f\left(\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right) = \begin{pmatrix} 2 \\ -1 \end{pmatrix} = 2\mathbf{e}_1 - \mathbf{e}_2,$$

e quindi $A_{\mathcal{B}} = \begin{bmatrix} 1 & 2 \\ 1 & -1 \end{bmatrix}$.
D'altra parte,

$$f(\mathbf{v}_1) = f\left(\begin{pmatrix} 2 \\ -1 \end{pmatrix}\right) = \begin{pmatrix} 0 \\ 3 \end{pmatrix} = -\mathbf{v}'_1 + 2\mathbf{v}'_2$$

$$f(\mathbf{v}_2) = f\left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right) = \begin{pmatrix} 3 \\ 0 \end{pmatrix} = \mathbf{v}_1' + \mathbf{v}_2',$$

e quindi $A_{\mathcal{B}'} = \begin{bmatrix} -1 & 1 \\ 2 & 1 \end{bmatrix}$.

Utilizzo ora la formula (3.5) di cambiamento di base. La matrice che ha come prima colonna le coordinate di $\mathbf{v}_1'$ *rispetto alla base* $\mathcal{U}$ *e come seconda colonna le coordinate di* $\mathbf{v}_2'$ *rispetto alla base* $\mathcal{U}$ *è data da*

$$P_{\mathcal{U}\mathcal{B}'} = \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix}.$$

Risulta

$$P_{\mathcal{U}\mathcal{B}'}^{-1} = \frac{1}{3}\begin{bmatrix} 1 & -1 \\ 1 & 2 \end{bmatrix},$$

e quindi

$$\begin{aligned} A_{\mathcal{B}'} = P_{\mathcal{U}\mathcal{B}'}^{-1} A_{\mathcal{U}} P_{\mathcal{U}\mathcal{B}'} &= \frac{1}{3}\begin{bmatrix} 1 & -1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix} \\ &= \frac{1}{3}\begin{bmatrix} 1 & -1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 0 & 3 \\ 3 & 0 \end{bmatrix} \\ &= \begin{bmatrix} -1 & 1 \\ 2 & 1 \end{bmatrix}. \end{aligned}$$

Definizione 3.28. *Due matrici* $A, B \in \mathrm{Mat}_n(\mathbb{R})$ *si dicono simili se esiste una matrice* $S \in \mathrm{Mat}_n(\mathbb{R})$ *invertibile con*

$$B = S^{-1}AS.$$

Abbiamo dunque visto che se due matrici rappresentano la stessa applicazione lineare sono simili. Vale anche il viceversa.

Teorema 3.29. *Due matrici sono simili e e solo se rappresentano la stessa applicazione lineare.*

Capitolo 4

Proiezioni

4.1 Prodotto interno

Definizione 4.1. *Sia V uno spazio vettoriale su $\mathbb{R}$. Si dice prodotto interno (o prodotto scalare) una funzione*

$$\begin{aligned}\langle\cdot,\cdot\rangle : V \times V &\to \mathbb{R}\\ (v,w) &\mapsto \langle v,w\rangle\end{aligned}$$

tale che

1. $\forall v \in V$, $\langle v,v\rangle \geq 0$, *e* $\langle v,v\rangle = 0$ *se e solo se* $v = 0_V$;
2. $\forall v,w \in V$, $\langle v,w\rangle = \langle w,v\rangle$;
3. $\forall \alpha,\beta \in \mathbb{R}$, $\forall v,u,w \in V$, $\langle \alpha v + \beta u, w\rangle = \alpha\langle v,w\rangle + \beta\langle u,w\rangle$.

Esempio 4.2. *In $\mathbb{R}^n$, dati $\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ e $\mathbf{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$, si definisce prodotto interno euclideo*

$$\langle \mathbf{x},\mathbf{y}\rangle := x_1y_1 + x_2y_2 + \cdots + x_ny_n = \mathbf{x}^T\mathbf{y} = \sum_{i=1}^{n} x_iy_i.$$

Si lascia per esercizio la verifica delle proprietà 1), 2) e 3) della Definizione 4.1.

Esempio 4.3. *In $\mathbb{P}_n$, dati i polinomi $p(t)$ e $q(t)$, definiamo un prodotto interno ponendo*

$$\langle p(t), q(t)\rangle := \int_{-1}^{1} p(t)q(t)dt. \tag{4.1}$$

Si lascia per esercizio la verifica delle proprietà 1), 2) e 3) della Definizione 4.1.

Definizione 4.4. *Sia V uno spazio vettoriale su $\mathbb{R}$. Una norma in V è una funzione*

$$\begin{aligned}||\cdot|| : V &\to \mathbb{R}\\ v &\mapsto ||v||\end{aligned}$$

tale che

1. $\forall v \in V$, $||v|| \geq 0$, *e* $||v|| = 0$ *se e solo se* $v = 0_V$*;*

2. $\forall v, w \in V$, $||v + w|| \leq ||v|| + ||w||$*;*

3. $\forall \alpha \in \mathbb{R}$, $\forall v \in V$, $||\alpha v|| = |\alpha|\, ||v||$*.*

Osservazione 4.5. *Se V è uno spazio vettoriale munito di prodotto interno $\langle \cdot, \cdot \rangle$, si ottiene una norma in V ponendo*

$$||v|| = \sqrt{\langle v, v \rangle}. \tag{4.2}$$

Esempio 4.6. *In $\mathbb{R}^n$ con il prodotto interno euclideo, dato $\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, otteniamo la norma euclidea*

$$||\mathbf{x}|| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle} = \sqrt{\sum_{i=1}^{n} x_i^2}.$$

Teorema 4.7 (Disuguaglianza di Cauchy-Schwarz)**.** *Sia V uno spazio vettoriale con prodotto interno $\langle \cdot, \cdot \rangle$, e sia $|| \cdot ||$ la norma associata tramite (4.2). Allora*

$$\langle v, w \rangle \leq ||v||\, ||w|| \quad \forall v, w \in V. \tag{4.3}$$

Dim. Siano $v, w \in V$, e consideriamo il vettore $v + tw$, con $t \in \mathbb{R}$. Si ha

$$\begin{aligned} \langle v + tw, v + tw \rangle &= \langle v, v \rangle + \langle v, tw \rangle + \langle tw, v \rangle + \langle tw, tw \rangle \\ &= t^2 ||w||^2 + 2t\langle v, w \rangle + ||v||^2. \end{aligned} \tag{4.4}$$

D'altra parte,

$$\langle v + tw, v + tw \rangle \geq 0, \ \forall t \in \mathbb{R},$$

dunque il polinomio di secondo grado in t in (4.4) soddisfa

$$at^2 + 2bt + c \geq 0, \ \forall t \in \mathbb{R}, \tag{4.5}$$

con $a := ||w||^2$, $b := \langle v, w \rangle$, $c := ||v||^2$. Ora, essendo $a \geq 0$, (4.5) implica $\Delta = b^2 - ac \leq 0$, ovvero

$$\langle v, w \rangle^2 \leq ||v||^2 ||w||^2$$

da cui segue (4.3). □

4.2 Vettori ortonormali

Definizione 4.8. *Sia V uno spazio vettoriale su $\mathbb{R}$ con prodotto interno $\langle \cdot, \cdot \rangle$. Due vettori $v, w \in V$ si dicono ortogonali se*

$$\langle v, w \rangle = 0.$$

www.unidocs.it - Appunti e dispense per superare i tuoi esami universitari

Esempio 4.9. *In* $\mathbb{R}^3$ *con prodotto interno euclideo i vettori* $\mathbf{x} = \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix}$ *e* $\mathbf{y} = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$ *sono ortogonali. Infatti* $\langle \mathbf{x}, \mathbf{y} \rangle = 2 - 1 - 1 = 0$.

Definizione 4.10. *Sia* V *uno spazio vettoriale su* $\mathbb{R}$ *con prodotto interno* $\langle \cdot, \cdot \rangle$. *I vettori di un insieme* $S \subseteq V$ *si dicono ortonormali se*

$$\forall\, v, w \in S \qquad \langle v, w \rangle = \begin{cases} 0 & \text{se } v \neq w \\ 1 & \text{se } v = w. \end{cases}$$

Esempio 4.11. *I vettori* $\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n$ *della base canonica di* $\mathbb{R}^n$ *sono ortonormali.*

Proposizione 4.12. *Sia* V *uno spazio vettoriale su* $\mathbb{R}$ *con prodotto interno* $\langle \cdot, \cdot \rangle$. *Siano* $v_1, ..., v_m$ *vettori ortonormali di* V. *Allora* $v_1, ..., v_m$ *sono linearmente indipendenti.*

Dim. Sia

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_m v_m = 0_V \quad \text{per certi } \alpha_1, \alpha_2, ..., \alpha_m \in \mathbb{R}.$$

Fissato $i \in I_m$, si ha

$$\begin{aligned} \langle \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_m v_m, v_i \rangle &= \alpha_1 \langle v_1, v_i \rangle + \cdots + \alpha_m \langle v_m, v_i \rangle \\ &= \alpha_i \langle v_i, v_i \rangle = \alpha_i. \end{aligned}$$

D'altra parte,

$$\langle \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_m v_m, v_i \rangle = \langle 0_V, v_i \rangle = 0.$$

Segue che $\alpha_i = 0$ per ogni $i \in I_m$, e dunque $v_1, ..., v_m$ sono linearmente indipendenti. □

Proposizione 4.13. *Sia* V *uno spazio vettoriale su* $\mathbb{R}$ *con prodotto interno* $\langle \cdot, \cdot \rangle$. *Sia* $\mathcal{B} = \{v_1, v_2, ..., v_n\}$ *una base di* V *costituita da vettori ortonormali. Allora, per* $w \in V$, *si ha*

1. $w = \sum_{i=1}^m \langle w, v_i \rangle v_i$;
2. $||w||^2 = \sum_{i=1}^m \langle w, v_i \rangle^2$.

Dim. (1) Poiché $w \in V$ e $\{v_1, v_2, ..., v_n\}$ è una base di V, ogni vettore di V si può scrivere, in modo unico, come

$$w = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n. \tag{4.6}$$

Allora, per ogni $i \in I_n$,

$$\langle w, v_i \rangle = \langle \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n, v_i \rangle = \alpha_i.$$

Segue da (4.6) che $w = \sum_{i=1}^m \langle w, v_i \rangle v_i$.

(2) Da (1) abbiamo

$$||w||^2 = \langle w, \rangle w = \left\langle \sum_{i=1}^{m} \langle w, v_i \rangle v_i, \sum_{j=1}^{m} \langle w, v_j \rangle v_j \right\rangle$$
$$= \sum_{i=1}^{m} \sum_{j=1}^{m} \langle w, v_i \rangle \langle w, v_j \rangle \langle v_i, v_j \rangle.$$

Essendo

$$\langle v_i, v_j \rangle = \begin{cases} 0 & \text{se } v_i \neq v_j \\ 1 & \text{se } v_i = v_j, \end{cases}$$

si ottiene $||w||^2 = \sum_{i=1}^{m} \langle w, v_i \rangle^2$. □

Procedimento di ortonormalizzazione di Gram-Schmidt.

<u>*Problema*</u>: Sia V uno spazio vettoriale su $\mathbb{R}$ con prodotto interno $\langle \cdot, \cdot \rangle$, e sia $\mathcal{B} = \{v_1, ..., v_n\}$ una base di V. Si vuole costruire un'altra base di V, $\mathcal{B}' = \{w_1, ..., w_n\}$, i cui vettori siano ortonormali.

Per fare ciò seguiamo il seguente procedimento, detto di ortogonalizzazione di Gram-Schmidt.

1. Poniamo
$$w_1 := \frac{v_1}{||v_1||},$$
così $||w_1|| = \frac{||v_1||}{||v_1||} = 1$.

2. Definiamo
$$u_2 := v_2 - \langle v_2, w_1 \rangle w_1.$$
I vettori u_2 e w_1 sono ortogonali:
$$\langle u_2, w_1 \rangle = \langle v_2, w_1 \rangle - \langle v_2, w_1 \rangle \langle w_1, w_1 \rangle = \langle v_2, w_1 \rangle - \langle v_2, w_1 \rangle = 0.$$
Inoltre, dal punto (1) si ha
$$u_2 = v_2 - \langle v_2, w_1 \rangle w_1 = v_2 - \langle v_2, \frac{v_1}{||v_1||} \rangle \frac{v_1}{||v_1||} = v_2 - \frac{\langle v_2, v_1 \rangle}{||v_1||^2} v_1.$$
Segue che $u_2 \neq 0_V$, perché altrimenti v_2 e v_1 sarebbero dipendenti. Possiamo dunque porre
$$w_2 := \frac{u_2}{||u_2||},$$
così $\langle w_2, w_1 \rangle = 0$ e $||w_2|| = 0$.

3. Definiamo
$$u_3 := v_3 - \langle v_3, w_1 \rangle w_1 - \langle v_3, w_2 \rangle w_2.$$
Ragionando come al punto (2) si prova che u_3 è ortogonale a w_1 e w_2 e che $u_3 \neq 0_V$. Poniamo
$$w_2 := \frac{u_3}{||u_3||}.$$

$\vdots$

(n) Definiamo

$$u_n := v_n - \langle v_n, w_1\rangle w_1 - \langle v_n, w_2\rangle w_2 - \cdots - \langle v_n, w_{n-1}\rangle w_{n-1}.$$

u_n è ortogonale a $w_1, w_2, ..., w_{n-1}$ e $u_n \neq 0_V$. Poniamo

$$w_n := \frac{u_n}{||u_n||}.$$

Esempio 4.14. *In* $\mathbb{R}^3$ *sia* $B = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ *la base data da*

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \quad \mathbf{v}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Abbiamo $||\mathbf{v}_1|| = \sqrt{3}$, *e*

$$\mathbf{w}_1 = \frac{\mathbf{v}_1}{||\mathbf{v}_1||} = \frac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Inoltre $\langle \mathbf{v}_2, \mathbf{w}_1\rangle = \frac{2}{\sqrt{3}}$, *e*

$$\mathbf{u}_2 = \mathbf{v}_2 - \langle \mathbf{v}_2, \mathbf{w}_1\rangle \mathbf{w}_1 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} - \frac{2}{3}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \frac{1}{3}\begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix}.$$

Abbiamo $||\mathbf{u}_2|| = \frac{\sqrt{6}}{3}$, *e*

$$\mathbf{w}_2 = \frac{\mathbf{u}_2}{||\mathbf{u}_2||} = \frac{1}{\sqrt{6}}\begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix}.$$

Infine, $\langle \mathbf{v}_3, \mathbf{w}_1\rangle = \frac{1}{\sqrt{3}}$ *e* $\langle \mathbf{v}_3, \mathbf{w}_2\rangle = \frac{1}{\sqrt{6}}$, *e*

$$\mathbf{u}_3 = \mathbf{v}_3 - \langle \mathbf{v}_3, \mathbf{w}_1\rangle \mathbf{w}_1 - \langle \mathbf{v}_3, \mathbf{w}_2\rangle \mathbf{w}_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} - \frac{1}{3}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - \frac{1}{6}\begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}$$

Pertanto, essendo $||\mathbf{u}_3|| = \frac{\sqrt{2}}{2}$,

$$\mathbf{w}_3 = \frac{\mathbf{u}_3}{||\mathbf{u}_3||} = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}.$$

4.3 Problema della miglior approssimazione

Definizione 4.15. *Sia V uno spazio vettoriale con prodotto interno $\langle\cdot,\cdot\rangle$, e sia $||\cdot||$ la norma indotta dal prodotto interno data da (4.2). Per $v, w \in V$, la distanza tra v e w è definita da*

$$d(v,w) := ||v-w||.$$

Sia V uno spazio vettoriale con prodotto interno $\langle\cdot,\cdot\rangle$, e sia $||\cdot||$ la norma indotta dal prodotto interno data da (4.2). Assegnato un sottospazio S di V, ha senso porre il seguente problema, che sorge stimolato da questioni di interesse geometrico:

Problema. Dato il sottospazio $S \subseteq V$ e dato un vettore $u \in V$, determinare un elemento di S la cui distanza da u sia minima rispetto a quella degli altri elementi di S.

Formulato in altri termini, il problema consiste nella ricerca di un punto di minimo della funzione $d(u,\cdot) = ||u-\cdot||$ sull'insieme S, ovvero

$$\text{Trovare } s \in S: \ ||u-s|| < ||u-s'|| \quad \forall s' \in S. \tag{P}$$

Definizione 4.16. *Sia S un sottoinsieme di V.*

(i) *Un vettore $v \in V$ è ortogonale a S se*

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

(ii) *L'insieme*

$$S^\perp = \{v \in V: \ \langle v,s\rangle = 0 \quad \forall s \in S\}$$

si dice l'ortogonale di S. Se S è un sottospazio, $S^\perp$ si dice complemento ortogonale di S.

Esempio 4.17. *Nello spazio $\mathbb{R}^2$, munito del prodotto interno euclideo, il vettore*

$$\mathbf{v} = \begin{pmatrix} 3 \\ 0 \end{pmatrix}$$

è ortogonale all'insieme $S = \operatorname{span}\left\{\begin{pmatrix} 0 \\ -2 \end{pmatrix}\right\}$. Si ha $S^\perp = \operatorname{span}\left\{\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right\}$.

Esempio 4.18. *Nello spazio $\mathbb{P}_2$ munito del prodotto interno dato in (4.1), sia*

$$S = \{p(t) \in \mathbb{P}_2: \ p(t) = a_1t^2 + a_0, \ a_0, a_1 \in \mathbb{R}\}.$$

Si verifica facilmente che il vettore $q(t) = t$ è ortogonale a S. Si ha $S^\perp = \operatorname{span}\{q(t)\}$.

Proposizione 4.19. *Sia S un sottoinsieme di V. Allora $S^\perp$ è un sottospazio di V.*

Dim.

1. $0_V \in S^\perp$ perché $\langle 0_V, s\rangle = 0 \quad \forall s \in S$;

2. $\forall v, w \in S^\perp \Rightarrow v + w \in S^\perp$ perché

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

3. $\forall \alpha \in \mathbb{R}$, $\forall v \in S^\perp \Rightarrow \alpha v \in S^\perp$ perché

$$\langle \alpha v, s\rangle = \alpha\langle v, s\rangle = 0 \quad \forall s \in S.$$

$\square$

Osservazione 4.20. *Per ogni S sottospazio di V si ha*

$$S \cap S^\perp = \{0_V\}.$$

Infatti, se $v \in S \cap S^\perp$, risulta $||v||^2 = \langle v, v\rangle = 0$, e la conclusione segue ricordando che $||v|| = 0$ se e solo se $v = 0_V$.

Teorema 4.21 (Teorema della decomposizione ortogonale)**.** *Sia V uno spazio vettoriale con prodotto interno $\langle\cdot,\cdot\rangle$ e norma indotta $||\cdot||$, e sia S un sottospazio di V. Allora ogni vettore $v \in V$ si scrive in modo unico come*

$$v = \bar{s} + s^\perp \quad \text{con } \bar{s} \in S \text{ e } s^\perp \in S^\perp. \tag{4.7}$$

Inoltre

$$||v||^2 = ||\bar{s}||^2 + ||s^\perp||^2. \tag{4.8}$$

Osservazione 4.22. *Il Teorema 4.21 ci dice in particolare che, preso un sottospazio S di uno spazio vettoriale V con prodotto interno, ogni vettore $v \in V$ può essere rappresentato in maniera univoca come somma di un elemento $\bar{s} \in S$ e di un elemento $s^\perp \in S^\perp$, ovvero vale la seguente decomposizione in somma diretta ortogonale:*

$$V = S \oplus S^\perp. \tag{4.9}$$

Dim. Esistenza della decomposizione. Posto $n = \dim(S)$, sia $\mathcal{B}_S = \{s_1, ..., s_n\}$ una base di S. Possiamo supporre che $\mathcal{B}_S$ sia ortonormale (altrimenti applichiamo il procedimento di Gram-Schmidt). Sia $v \in V$. Definiamo

$$\bar{s} := \sum_{i=1}^{n} \langle v, s_i\rangle s_i, \quad s^\perp := v - \bar{s}.$$

È chiaro che $v = \bar{s} + s^\perp$, e che $\bar{s} \in S$. Resta da mostrare che $s^\perp \in S^\perp$.

Sia $s \in S$. Esistono $\alpha_1, \alpha_2, ..., \alpha_n \in \mathbb{R}$ tali che

$$s = \alpha_1 s_1 + \alpha_2 s_2 + \cdots + \alpha_n s_n = \sum_{j=1}^{n} \alpha_j s_j.$$

Si ha

$$\begin{aligned}\langle s^{\perp}, s\rangle &= \langle v-\bar{s}, \sum_{j=1}^{n}\alpha_j s_j\rangle = \langle v-\sum_{i=1}^{n}\langle v, s_i\rangle s_i, \sum_{j=1}^{n}\alpha_j s_j\rangle \\ &= \langle v, \sum_{j=1}^{n}\alpha_j s_j\rangle - \langle \sum_{i=1}^{n}\langle v, s_i\rangle s_i, \sum_{j=1}^{n}\alpha_j s_j\rangle \\ &= \sum_{j=1}^{n}\alpha_j\langle v, s_j\rangle - \sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_j\langle v, s_i\rangle\langle s_i, s_j\rangle.\end{aligned}$$

Essendo la base B_S ortonormale, $\langle s_i, s_j\rangle$ è 0 per $i \neq j$ ed è 1 per $i = j$. Si ottiene

$$\langle s^{\perp}, s\rangle = \sum_{j=1}^{n}\alpha_j\langle v, s_j\rangle - \sum_{j=1}^{n}\alpha_j\langle v, s_j\rangle = 0.$$

Unicità della decomposizione. Supponiamo sia

$$v = \bar{s} + s^{\perp}, \quad v = \bar{t} + t^{\perp}$$

con $\bar{s}, \bar{t} \in S$, $s^{\perp}, t^{\perp} \in S^{\perp}$. Allora

$$S \ni (\bar{s} - \bar{t}) = (t^{\perp} - s^{\perp}) \in S^{\perp}$$

perché per ipotesi S è un sottospazio di V, e $S^{\perp}$ è sempre un sottospazio di V, si veda la Proposizione 4.19. Quindi $(\bar{s} - \bar{t}) \in S \cap S^{\perp} = \{0_V\}$, si veda l'Osservazione 4.20. Segue che $\bar{s} = \bar{t}$ e $s^{\perp} = t^{\perp}$.
Formula (4.8). Si ha

$$||v||^2 = \langle v, v\rangle = \langle \bar{s} + s^{\perp}, \bar{s} + s^{\perp}\rangle = ||\bar{s}||^2 + 2\langle \bar{s}, s^{\perp}\rangle + ||s^{\perp}||^2 = ||\bar{s}||^2 + ||s^{\perp}||^2.$$

□

Definizione 4.23. *Siano V uno spazio vettoriale con prodotto interno $\langle\cdot,\cdot\rangle$, S un sottospazio di V di dimensione n, e $\mathcal{B}_S = \{s_1, ..., s_n\}$ una base ortonormale di S. Dato $v \in V$, si dice proiezione ortogonale di v su S il vettore*

$$\bar{s} = \sum_{i=1}^{n}\langle v, s_i\rangle s_i \in S.$$

Il seguente teorema rivela come il concetto di proiezione di un vettore su un sottospazio permette di risolvere il problema della migliore approssimazione (P). Più precisamente, esso stabilisce che tale problema ammette un'unica soluzione, e fornisce un procedimento per calcolarla.

Teorema 4.24 (Teorema della migliore approssimazione)**.** *Sia V uno spazio vettoriale con prodotto interno $\langle\cdot,\cdot\rangle$ e norma indotta $||\cdot||$. Sia S un sottospazio di V, e u un vettore in V. Detta $\bar{s}$ la proiezione ortogonale di u su S, risulta*

$$||u - \bar{s}|| < ||u - s'|| \quad \forall s' \in S.$$

Dim. Per il Teorema 4.21 possiamo scrivere

$$u = \bar{s} + s^{\perp}$$

dove $\bar{s}$ è proprio la proiezione ortogonale di u su S, e $s^{\perp} \in S^{\perp}$. Sia $s' \in S$. Si ha

$$u - s' = (u - \bar{s}) + (\bar{s} - s') \tag{4.10}$$

dove $(\bar{s} - s') \in S$, mentre $u - \bar{s} = s^{\perp} \in S^{\perp}$. Per il Teorema 4.21, (4.10) fornisce l'unica decomposizione di ortogonale $u - s'$. Inoltre, sempre per il Teorema 4.21,

$$||u - s'||^2 = ||u - \bar{s}||^2 + ||\bar{s} - s'||^2.$$

Ricordando che $||\bar{s} - s'||^2 \geq 0$, e $||\bar{s} - s'||^2 = 0$ se e solo se $\bar{s} = s'$, per $s' \neq \bar{s}$ si ha

$$||u - \bar{s}||^2 < ||u - s'||^2.$$

ovvero $||u - \bar{s}|| < ||u - s'||$. □

Definizione 4.25. *Sia V uno spazio vettoriale con prodotto interno $\langle \cdot, \cdot \rangle$, S un sottospazio di V, e u un vettore in V. Indicando con $\bar{s}$ la proiezione ortogonale di u su S nella Definizione 4.23, il numero (non negativo) $||u - \bar{s}||$ è detto distanza di u da S.*

Esercizio 4.26. *In $\mathbb{R}^3$ con il prodotto interno euclideo, si consideri il sottospazio*

$$S = \text{span}\left\{ \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 4 \\ 2 \end{pmatrix} \right\},$$

e il vettore $\mathbf{u} = \begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix}$. Determinare il vettore di S che meglio approssima $\mathbf{u}$.

Cominciamo con l'osservare che $\dim(S) = \text{rg}(A)$ con

$$A = \begin{bmatrix} 1 & 1 & 2 \\ 2 & 1 & 4 \\ 1 & 1 & 2 \end{bmatrix}.$$

Si ha $\dim(S) = 2$. Una base di S è $\mathcal{B}_S = \{\mathbf{t}_1, \mathbf{t}_2\}$, con

$$\mathbf{t}_1 = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \mathbf{t}_2 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Costruiamo una base ortonormale di S con il procedimento di Gram-Schmidt. Abbiamo $||\mathbf{t}_1|| = \sqrt{6}$, e

$$\mathbf{s}_1 = \frac{\mathbf{t}_1}{||\mathbf{t}_1||} = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}.$$

Inoltre $\langle \mathbf{t}_2, \mathbf{s}_1 \rangle = \frac{4}{\sqrt{6}}$, *e*

$$\mathbf{t}_2' = \mathbf{t}_2 - \langle \mathbf{t}_2, \mathbf{s}_1 \rangle \mathbf{s}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - \frac{4}{6} \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}.$$

Abbiamo $||\mathbf{t}_2'|| = \frac{1}{\sqrt{3}}$, *e*

$$\mathbf{s}_2 = \frac{\mathbf{t}_2'}{||\mathbf{t}_2'||} = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}.$$

Abbiamo pertanto la base ortonormale di S data da $\{\mathbf{s}_1, \mathbf{s}_2\}$. Per il Teorema 4.24, il vettore di S che meglio approssima $\mathbf{u}$ è la proiezione di $\mathbf{u}$ su S, cioè

$$\bar{\mathbf{s}} = \langle \mathbf{u}, \mathbf{s}_1 \rangle \mathbf{s}_1 + \langle \mathbf{u}, \mathbf{s}_2 \rangle \mathbf{s}_2.$$

Si ha

$$\langle \mathbf{u}, \mathbf{s}_1 \rangle = \langle \begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix}, \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} \rangle = \sqrt{6},$$

$$\langle \mathbf{u}, \mathbf{s}_2 \rangle = \langle \begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix}, \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} \rangle = 0,$$

quindi

$$\bar{\mathbf{s}} = \sqrt{6}\, \mathbf{s}_1 = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}.$$

4.3.1 Proiettori ortogonali: forma matriciale ed esempi

Definizione 4.27. *Dato V spazio vettoriale su $\mathbb{R}$ e $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\} \subseteq V$, chiamiamo matrice di Gram associata a $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ la matrice $G \in \mathrm{Mat}_k(\mathbb{R})$ i cui elementi sono definiti da*

$$g_{ij} = \langle \mathbf{v}_i, \mathbf{v}_j \rangle.$$

Osservazione 4.28. *Sia S un sottospazio k-dimensionale di $\mathbb{R}^n$, e sia $\mathcal{B}_S = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ una base di S con matrice associata di Gram G. Presi due vettori $\mathbf{x}^a, \mathbf{x}^b \in S$, il loro prodotto scalare è dato da*

$$\langle \mathbf{x}^a, \mathbf{x}^b \rangle = \sum_{i=1}^{k} \sum_{j=1}^{k} \lambda_i^a \lambda_j^b \langle \mathbf{v}_i, \mathbf{v}_j \rangle = \boldsymbol{\lambda}^{a^T} G \boldsymbol{\lambda}^b, \tag{4.11}$$

dove i vettori $\boldsymbol{\lambda}^a$ e $\boldsymbol{\lambda}^b$ hanno per componenti i coefficienti delle espansioni di $\mathbf{x}^a$ e $\mathbf{x}^b$ sulla base $\mathbf{v}_1, \ldots, \mathbf{v}_k$.

Definizione 4.29. *Data la decomposizione ortogonale $V = S \oplus S^\perp$, l'operatore lineare $p_S : V \to S$ che associa ad ogni vettore di V la propria componente $\bar{s}$ su S è detto proiettore ortogonale su S.*

Deriviamo ora la rappresentazione matriciale di un proiettore ortogonale su un dato sottospazio di $V = \mathbb{R}^n$. .

Proposizione 4.30 (Rappresentazione matriciale di un proiettore ortogonale)**.** *Sia $\mathcal{B} = \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ una base di $\mathbb{R}^n$ con matrice di Gram G. Sia $S = \operatorname{span}\{\mathbf{x}_1, \dots, \mathbf{x}_k\}$ e sia p_S il proiettore su S. Allora la matrice che rappresenta p_S sulla base $\mathcal{B}$ è data da:*

$$P_S = X(X^T G X)^{-1} X^T G, \tag{4.12}$$

dove $X = [\mathbf{x}_1 | \cdots | \mathbf{x}_k] \in \mathrm{Mat}_{n,k}(\mathbb{R})$.

Dim. Sia $\mathbf{y} \in \mathbb{R}^n$ un generico vettore. Poiché $p_S(\mathbf{y}) \in S$, si ha

$$p_S(\mathbf{y}) = X\boldsymbol{\lambda} \tag{4.13}$$

per un opportuno vettore $\boldsymbol{\lambda} \in \mathbb{R}^k$. Per definizione, il vettore $\mathbf{y} - p_S(\mathbf{y})$ deve essere ortogonale a S e quindi a tutte le colonne di X, che di S formano una base, pertanto (ricordando che $X^T G X$ è invertibile):

$$\mathbf{0} = X^T G(\mathbf{y} - p_S(\mathbf{y}))$$

da cui

$$X^T G\mathbf{y} = X^T G X \boldsymbol{\lambda}$$

e quindi

$$\boldsymbol{\lambda} = (X^T G X)^{-1} X^T G \mathbf{y}.$$

L'identità (4.13) diventa

$$P_S \mathbf{y} := p_S(\mathbf{y}) = X(X^T G X)^{-1} X^T G \mathbf{y}.$$

Poiché questa espressione vale per qualunque vettore $\mathbf{y} \in \mathbb{R}^n$, si ottiene la formula (4.12).
□

Osservazione 4.31. *Tipicamente si sceglie $G = \mathbb{I}_k$ e quindi la matrice di rappresentazione del proiettore assume la forma semplificata*

$$P_S = X(X^T X)^{-1} X^T. \tag{4.14}$$

Nell'analisi dei dati, tuttavia, è tipico anche scegliere come matrice di Gram la matrice $G = diag(\frac{1}{n}, \dots, \frac{1}{n})$, che però non muta la forma del proiettore, dato che i fattori $\frac{1}{n}$ presenti nella sua espressione originaria si semplificano.

Vediamo ora alcuni esempi di rappresentazioni matriciali per proiettori ortogonali di uso comune nell'analisi dei dati.

- **Proiettore su un sottospazio unidimensionale**. Se il sottospazio di proiezione S è unidimensionale, la matrice di rappresentazione del proiettore assume una forma molto semplice.
  Sia infatti S tale che $\dim(S) = 1$. Preso un vettore $\mathbf{x} \in S$ (che quindi è una base per S), e applicando la formula (4.14), si ottiene

$$P_{\mathbf{x}} = \mathbf{x}(\mathbf{x}^T\mathbf{x})^{-1}\mathbf{x}^T = \frac{1}{\|\mathbf{x}\|^2}\mathbf{x}\mathbf{x}^T. \tag{4.15}$$

Nel caso in cui $||\mathbf{x}|| = 1$, l'espressione precedente si riduce a

$$P_{\mathbf{x}} = \mathbf{x}\mathbf{x}^T.$$

Proiezione e centratura delle variabili. Sia $\mathbf{1}_n \in \mathbb{R}^n$ il vettore $(1, 1, \ldots, 1)^T$ e sia $S = \operatorname{span}\{\mathbf{1}_n\}$ il sottospazio unidimensionale generato da $\mathbf{1}_n$. Il proiettore su S ha la seguente matrice di rappresentazione data da (4.15):

$$P_{\mathbf{1}_n} = \mathbf{1}_n(\mathbf{1}_n^T\mathbf{1}_n)^{-1}\mathbf{1}_n^T = \frac{1}{n}\mathbf{1}_n\mathbf{1}_n^T.$$

L'operatore $p_{\mathbf{1}_n} : \mathbb{R}^n \to S$ associa ad ogni vettore $\mathbf{y} \in \mathbb{R}^n$ il vettore $\boldsymbol{\mu}_y \in S$ avente come componenti la media μ_y di $\mathbf{y}$. Infatti:

$$p_{\mathbf{1}_n}(\mathbf{y}) = P_{\mathbf{1}_n}\mathbf{y} = \frac{1}{n}\mathbf{1}_n\mathbf{1}_n^T\mathbf{y} = \mathbf{1}_n\left(\frac{1}{n}\mathbf{1}_n^T\mathbf{y}\right) = \mathbf{1}_n\left(\frac{1}{n}\sum_{i=1}^n y_i\right) = \mu_y\mathbf{1}_n = \boldsymbol{\mu}_y.$$

Di conseguenza, l'operatore che proietta sull'ortogonale di S,

$$p_{\mathbf{1}_n^\perp} := 1_n - p_{\mathbf{1}_n} : \mathbb{R}^n \to S^\perp,$$

è l'operatore di centratura. Infatti:

$$p_{\mathbf{1}_n^\perp}(\mathbf{y}) = (\mathbb{I}_n - P_{\mathbf{1}_n})\mathbf{y} = \mathbf{y} - \boldsymbol{\mu}_y.$$

- **Proiettore e basi ortogonali**. Sia $\mathcal{B} = \{\mathbf{x}_1, \ldots, \mathbf{x}_k\}$ una base ortogonale per il sottospazio S, e sia p_S il proiettore su S con matrice di rappresentazione rispetto a $\mathcal{B}$ data da (4.14). Per l'ortogonalità della base, la matrice X^TX è diagonale ed è pari a

$$X^TX = diag(\|\mathbf{x}_1\|^2, \ldots, \|\mathbf{x}_k\|^2).$$

Di conseguenza, il proiettore su S è esprimibile come somma dei proiettori sui singoli sottospazi generati dagli elementi della base. Infatti:

$$P_S = X\, diag\left(\frac{1}{\|\mathbf{x}_1\|^2}, \ldots, \frac{1}{\|\mathbf{x}_k\|^2}\right)X^T = \sum_{i=1}^k \frac{1}{\|\mathbf{x}_i\|^2}\mathbf{x}_i\mathbf{x}_i^T = \sum_{i=1}^k P_{\mathbf{x}_i}.$$

- **Proiezioni, covarianza, varianza e correlazione**. La covarianza tra due vettori $\mathbf{z}, \mathbf{y} \in \mathbb{R}^n$ è data dal prodotto scalare (con matrice di Gram $G = diag(\frac{1}{n}, \ldots, \frac{1}{n})$) tra i rispettivi vettori centrati. Infatti

$$\begin{aligned} Cov(\mathbf{z}, \mathbf{y}) &:= \frac{1}{n}\sum_{i=1}^{n} z_i y_i - \boldsymbol{\mu_z^T}\boldsymbol{\mu_y} \\ &= \frac{1}{n}\langle \mathbf{z}, \mathbf{y}\rangle - \frac{1}{n}\langle P_{\mathbf{1}_n}\mathbf{z}, P_{\mathbf{1}_n}\mathbf{y}\rangle \\ &= \frac{1}{n}\langle (\mathbb{I}_n - P_{\mathbf{1}_n})\mathbf{z}, (\mathbb{I}_n - P_{\mathbf{1}_n})\mathbf{y}\rangle \\ &= \frac{1}{n}\langle P_{\mathbf{1}_n^\perp}\mathbf{z}, P_{\mathbf{1}_n^\perp}\mathbf{y}\rangle. \end{aligned}$$

Ponendo $\mathbf{y} = \mathbf{z}$ nella formula precedente, si ottiene l'espressione della varianza:

$$Var(\mathbf{y}) \;:=\; \frac{1}{n}\sum_{i=1}^{n} y_i^2 - ||\boldsymbol{\mu_y}||^2 = \frac{1}{n}||P_{\mathbf{1}_n^\perp}\mathbf{y}||^2.$$

Infine, la correlazione fra i medesimi vettori si esprime come

$$corr(\mathbf{z}, \mathbf{y}) := \frac{Cov(\mathbf{z}, \mathbf{y})}{\sqrt{Var(\mathbf{z})}\sqrt{Var(\mathbf{z})}} = \frac{\langle P_{\mathbf{1}_n^\perp}\mathbf{z}, P_{\mathbf{1}_n^\perp}\mathbf{y}\rangle}{\|P_{\mathbf{1}_n^\perp}\mathbf{z}\|\|P_{\mathbf{1}_n^\perp}\mathbf{y}\|}.$$

- **Proiettori ortogonali e cambiamenti di base**. Consideriamo un sottospazio $S \subset \mathbb{R}^n$ con $\dim(S) = k$, e fissiamone una base $\mathcal{B} = \{\mathbf{x}_1, \ldots, \mathbf{x}_k\}$. Un qualunque vettore $\mathbf{y} \in S$ può essere rappresentato in modo unico, sulla base scelta, mediante un opportuno insieme di coefficienti $\lambda_1, \ldots, \lambda_k$. In termini matriciali:

$$\mathbf{y} = X\boldsymbol{\lambda}.$$

Supponiamo, adesso, di scegliere una diversa base $\mathcal{B}' = \{\mathbf{x}'_1, \ldots, \mathbf{x}'_k\}$, per cui si abbia

$$\mathbf{y} = X'\boldsymbol{\lambda}'.$$

<u>*Problema 1*</u>. Qual è la relazione tra i coefficienti (o coordinate) $\lambda_1, \ldots, \lambda_k$ e i coefficienti $\lambda'_1, \ldots, \lambda'_k$?

Sia p_S il proiettore ortogonale su S. Indichiamo con P_S la matrice di rappresentazione di p_S rispetto alla base $\mathcal{B}$ e con P'_S la matrice di rappresentazione di p_S rispetto alla base $\mathcal{B}'$:

$$P_S = X(X^TX)^{-1}X^T, \quad P'_S = X'(X'^TX')^{-1}X'^T.$$

P'_S agisce come la matrice identità su ogni $\mathbf{y} \in S$:

$$P'_S\mathbf{y} = P'_SX'\boldsymbol{\lambda}' = X'(X'^TX')^{-1}X'^TX'\boldsymbol{\lambda}' = X'\boldsymbol{\lambda}' = \mathbf{y}.$$

Di conseguenza

$$X'\boldsymbol{\lambda}' = \mathbf{y} = P'_S\mathbf{y} = P'_SX\boldsymbol{\lambda} = X'(X'^TX')^{-1}X'^TX\boldsymbol{\lambda}$$

da cui, per l'unicità della rappresentazione del vettore $\mathbf{y}$ sulla base $\mathcal{B}'$:

$$\boldsymbol{\lambda}' = (X'^T X')^{-1} X'^T X \boldsymbol{\lambda}. \tag{4.16}$$

In sostanza, per cambiare base è sufficiente scrivere la matrice di rappresentazione del proiettore rispetto alla base desiderata, applicarlo al vettore rappresentato sulla base originaria, e ricavare i nuovi coefficienti.

Osservazione 4.32. *Si noti che*

$$X = P_{\mathcal{U}\mathcal{B}}, \quad X' = P_{\mathcal{U}\mathcal{B}'}.$$

Poiché $(X'^T X')^{-1} = (X')^{-1}(X'^T)^{-1}$*, la formula* (4.16) *fornisce*

$$\boldsymbol{\lambda}' = (X')^{-1}(X'^T)^{-1} X'^T X \boldsymbol{\lambda} = (X')^{-1} X \boldsymbol{\lambda} = P_{\mathcal{U}\mathcal{B}'}^{-1} P_{\mathcal{U}\mathcal{B}} \boldsymbol{\lambda},$$

ovvero una riscrittura della classica formula di cambiamento di base (3.4)*.*

Capitolo 5

Sistemi di equazioni lineari

Definizione 5.1. *Un sistema di m equazioni lineari in n incognite $x_1, x_2..., x_n \in \mathbb{R}$ è un sistema della forma*

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases} \tag{5.1}$$

dove $a_{ij} \in \mathbb{R}$ e $b_l \in \mathbb{R}$ sono fissati, per ogni $(i,j) \in I_m \times I_n$ e $l \in I_m$.

Osservazione 5.2. *È immediato osservare che il sistema (5.1) si può scrivere in forma compatta come*

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

con

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

$A \in \mathrm{Mat}_{m,n}(\mathbb{R})$ si dice matrice dei coefficienti (o matrice incompleta), $\mathbf{x} \in \mathbb{R}^n$ si dice vettore delle incognite, $\mathbf{b} \in \mathbb{R}^m$ si dice vettore dei termini noti.

Definizione 5.3. *Una soluzione del sistema (5.1) è un vettore $\mathbf{y} \in \mathbb{R}^n$ tale che $A\mathbf{y} = \mathbf{b}$.*

Risolvere il sistema (5.1) significa trovare *tutte* le sue soluzioni. Un sistema si dice

- *possibile* se ha almeno una soluzione
  - (i) *determinato* se ha un'unica soluzione
  - (ii) *indeterminato* se ha infinite soluzioni
- *impossibile* se non ha soluzioni.

Osservazione 5.4. *Siano $A \in \mathrm{Mat}_{m,n}(\mathbb{R})$ e $\mathbf{b} \in \mathbb{R}^m$. Risolvere il sistema $A\mathbf{x} = \mathbf{b}$ significa trovare la controimmagine di $\mathbf{b}$ attraverso l'applicazione lineare $f : \mathbb{R}^n \to \mathbb{R}^m$ rappresentata dalla matrice A (si ha $f(\mathbf{x}) = A\mathbf{x}$, si veda la Proposizione 3.16). Dunque il sistema è possibile se e solo se $\mathbf{b} \in \mathrm{Im}(f)$, ovvero se e solo se $\mathbf{b}$ si può scrivere come combinazione lineare delle colonne di A.*

Notazione. La matrice

$$A|\mathbf{b} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{bmatrix} \in \mathrm{Mat}_{m,n+1}(\mathbb{R})$$

si dice matrice completa del sistema.

Teorema 5.5 (Rouché-Capelli)**.** *Siano $A \in \mathrm{Mat}_{m,n}(\mathbb{R})$ e $\mathbf{b} \in \mathbb{R}^m$.*
Il sistema lineare $A\mathbf{x} = \mathbf{b}$ ammette soluzioni se e solo se

$$\mathrm{rg}(A) = \mathrm{rg}(A|\mathbf{b}),$$

ovvero il rango della matrice dei coefficienti è uguale al rango della matrice completa del sistema.

Dim. Consideriamo i vettori colonna della matrice A

$$\mathbf{a}_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}, \quad \mathbf{a}_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix}, \quad \cdots \mathbf{a}_n = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix}$$

che sono n vettori di $\mathbb{R}^m$. Dire che $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$ è soluzione del sistema $A\mathbf{x} = \mathbf{b}$ equivale a dire che

$$y_1\mathbf{a}_1 + y_2\mathbf{a}_2 + \cdots + y_n\mathbf{a}_n = \mathbf{b},$$

ovvero che $\mathbf{b}$ è composizione lineare di $\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_n$. Questo accade se e solo se $\mathbf{b} \in \mathrm{span}\{\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_n\}$, ovvero se e solo se

$$\mathrm{span}\{\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_n\} = \mathrm{span}\{\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_n, \mathbf{b}\}. \tag{5.2}$$

D'altra parte $\mathrm{span}\{\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_n\} \subseteq \mathrm{span}\{\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_n, \mathbf{b}\}$, quindi (5.2) vale se e solo se

$$\mathrm{rg}(A) = \dim(\mathrm{span}\{\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_n\}) = \dim(\mathrm{span}\{\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_n, \mathbf{b}\}) = \mathrm{rg}(A|\mathbf{b}). \ \square$$

Definizione 5.6. *Un sistema di m equazioni lineari in n incognite $x_1, x_2, ..., x_n \in \mathbb{R}$ si dice omogeneo se è della forma (5.1) con $\mathbf{b} = \mathbf{0}_{\mathbb{R}^m}$, ovvero $A\mathbf{x} = \mathbf{0}_{\mathbb{R}^m}$, con $A \in \mathrm{Mat}_{m,n}(\mathbb{R})$.*

Osservazione 5.7. *Sia $A \in \mathrm{Mat}_{m,n}(\mathbb{R})$. Risolvere il sistema omogeneo $A\mathbf{x} = \mathbf{0}_{\mathbb{R}^m}$ equivale a determinare il nucleo della trasformazione lineare $f : \mathbb{R}^n \to \mathbb{R}^m$ rappresentata da A:*

$$\mathrm{Ker}(f) = \{\mathbf{x} \in \mathbb{R}^n : \ f(\mathbf{x}) = \mathbf{0}_{\mathbb{R}^m}\} = \{\mathbf{x} \in \mathbb{R}^n : \ A\mathbf{x} = \mathbf{0}_{\mathbb{R}^m}\}. \tag{5.3}$$

Osservazione 5.8. *Un sistema omogeneo è sempre possibile, avendo sicuramente almeno la soluzione banale $\mathbf{0}_{\mathbb{R}^n} \in \mathbb{R}^n$. La domanda interessante nel caso del sistema omogeneo è quindi se esso ha altre soluzioni oltre a quella nulla.*

Proposizione 5.9. *Sia $A \in \mathrm{Mat}_{m,n}(\mathbb{R})$. Le soluzioni del sistema $A\mathbf{x} = \mathbf{0}_{\mathbb{R}^m}$ formano un sottospazio di $\mathbb{R}^n$, di dimensione $n - \mathrm{rg}(A)$.*

Dim. Mostriamo dapprima che le soluzioni del sistema $A\mathbf{x} = \mathbf{0}_{\mathbb{R}^m}$ formano un sottospazio di $\mathbb{R}^n$.

1. $\mathbf{0}_{\mathbb{R}^n}$ è soluzione di $A\mathbf{x} = \mathbf{0}_{\mathbb{R}^m}$.
2. se $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{R}^n$ sono due soluzioni di $A\mathbf{x} = \mathbf{0}_{\mathbb{R}^m}$, anche $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ è soluzione di $A\mathbf{x} = \mathbf{0}_{\mathbb{R}^m}$. Infatti
$$A\left(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}\right) = A\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + A\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$$
3. se $\alpha \in \mathbb{R}$ e se $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$ è soluzione di $A\mathbf{x} = \mathbf{0}_{\mathbb{R}^m}$, anche $\alpha \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ è soluzione di $A\mathbf{x} = \mathbf{0}_{\mathbb{R}^m}$. Infatti
$$A\left(\alpha\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}\right) = \alpha A\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \alpha\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Infine, sia $f : \mathbb{R}^n \to \mathbb{R}^m$ l'applicazione lineare associata alla matrice A. Dal Teorema 3.17 (della nullità + rango) si ha

$$\dim(\mathrm{Ker}(f)) = n - \mathrm{rg}(A).$$

Da (5.3) segue che lo spazio delle soluzioni del sistema $A\mathbf{x} = \mathbf{0}_{\mathbb{R}^m}$ ha dimensione $n - \mathrm{rg}(A)$.
□

Osservazione 5.10. *Sia* $A \in \mathrm{Mat}_{m,n}(\mathbb{R})$. *Se* $\mathrm{rg}(A) = n$*, segue direttamente dalla Proposizione 5.9 che il sistema* $A\mathbf{x} = \mathbf{0}_{\mathbb{R}^m}$ *ammette solo la soluzione banale.*

Torniamo a un sistema della forma $A\mathbf{x} = \mathbf{b}$. Il teorema che segue illustra in generale qual è la struttura dell'insieme delle soluzioni di un sistema lineare.

Teorema 5.11. *Siano* $A \in \mathrm{Mat}_{m,n}(\mathbb{R})$ *e* $\mathbf{b} \in \mathbb{R}^m$.
Se il sistema $A\mathbf{x} = \mathbf{b}$ *ha la soluzione* $\mathbf{x}_0 \in \mathbb{R}^n$*, allora le soluzioni di* $A\mathbf{x} = \mathbf{b}$ *sono tutti e soli i vettori dell'insieme*

$$S = \{\mathbf{x}_0 + \mathbf{w} : \ A\mathbf{w} = \mathbf{0}_{\mathbb{R}^m}\} \subseteq \mathbb{R}^n. \tag{5.4}$$

Dim. Proviamo dapprima che se $\mathbf{w}$ è soluzione del sistema omogeneo associato $A\mathbf{x} = \mathbf{0}_{\mathbb{R}^m}$, allora $\mathbf{x}_0 + \mathbf{w}$ è soluzione di $A\mathbf{x} = \mathbf{b}$. Si ha

$$A(\mathbf{x}_0 + \mathbf{w}) = A\mathbf{v}_0 + A\mathbf{w} = \mathbf{b} + \mathbf{0}_{\mathbb{R}^m} = \mathbf{b}.$$

Mostriamo ora che, per qualunque soluzione $\mathbf{y}$ di $A\mathbf{x} = \mathbf{b}$, allora $\mathbf{y} = \mathbf{x}_0 + \mathbf{w}$, per una qualche $\mathbf{w}$ soluzione del sistema omogeneo associato $A\mathbf{x} = \mathbf{0}$. Sia dunque $\mathbf{y}$ soluzione di $A\mathbf{x} = \mathbf{b}$. Sappiamo che $A\mathbf{y} = \mathbf{b}$ e $A\mathbf{x}_0 = \mathbf{b}$. Pertanto

$$A(\mathbf{y} - \mathbf{x}_0) = \mathbf{0}$$

ovvero $\mathbf{w} := \mathbf{y} - \mathbf{x}_0$ è soluzione del sistema omogeneo associato $A\mathbf{x} = \mathbf{0}_{\mathbb{R}^m}$. In altre parole, $\mathbf{y} = \mathbf{x}_0 + \mathbf{w}$ con $\mathbf{w}$ soluzione del sistema omogeneo associato $A\mathbf{x} = \mathbf{0}_{\mathbb{R}^m}$. □

Osservazione 5.12. *Segue dalle Proposizioni 5.9 e 5.11 che le soluzioni di* $A\mathbf{x} = \mathbf{b}$*, quando esistono, dipendono da* $n - \mathrm{rg}(A)$ *parametri.*

Osservazione 5.13. *L'insieme delle soluzioni di un sistema non omogeneo non è un sottospazio di* $\mathbb{R}^n$ *(non contiene il vettore nullo).*

Teorema 5.14 (Teorema di Cramer)**.** *Siano* $A \in \mathrm{Mat}_n(\mathbb{R})$ *e* $\mathbf{b} \in \mathbb{R}^n$.
Il sistema quadrato $A\mathbf{x} = \mathbf{b}$ *ha una ed una sola soluzione se e solo se* $\det A \neq 0$.

Dim. ($\Leftarrow$) Se $\det A \neq 0$ allora A è invertibile ed esiste A^{-1}, si vedano la Definizione 2.22 e la Proposizione 2.37. Quindi, se $A\mathbf{x} = \mathbf{b}$, allora $\mathbf{x} = A^{-1}\mathbf{b}$ e quindi vi è una e una sola soluzione.
($\Rightarrow$) Se $A\mathbf{x} = \mathbf{b}$ ha una ed una sola soluzione, da (5.4) segue che $\mathrm{Ker}(f) = \{\mathbf{0}_{\mathbb{R}^n}\}$, dove $f : \mathbb{R}^n \to \mathbb{R}^n$ è la trasformazione lineare rappresentata da A. Per la Proposizione 3.10 f è invertibile dunque esiste f^{-1} (e di conseguenza A^{-1}), pertanto A è non singolare ancora per la Proposizione 2.37. □

Osservazione 5.15. *Come già osservato in precedenza, i sistemi omogenei hanno sempre almeno la soluzione nulla. Sull'esistenza di altre soluzioni, distinguiamo due casi.*

- $A \in \text{Mat}_n(\mathbb{R})$. *Per il Teorema di Cramer 5.14*, $A\mathbf{x} = \mathbf{0}_{\mathbb{R}^m}$ *ha soluzioni non nulle se e solo se* $\det A = 0$, *ovvero se e solo se* $\text{rg}(A) < n$, *ovvero se e solo se* $\text{Ker}(f) \neq \{\mathbf{0}_{\mathbb{R}^n}\}$, *dove* $f : \mathbb{R}^n \to \mathbb{R}^n$ *è la trasformazione lineare rappresentata da* A.

- $A \in \text{Mat}_{m,n}(\mathbb{R})$, $n \neq m$. $A\mathbf{x} = \mathbf{0}_{\mathbb{R}^m}$ *ha soluzioni non nulle se e solo se* $\text{rg}(A) < n$, *ovvero se e solo se* $\text{Ker}(f) \neq \{\mathbf{0}_{\mathbb{R}^n}\}$, *dove* $f : \mathbb{R}^n \to \mathbb{R}^m$ *è la trasformazione lineare rappresentata da* A. *Questo in particolare è verificato quando* $m < n$.

Relativamente al calcolo delle soluzioni di un sistema quadrato $A\mathbf{x} = \mathbf{b}$ con $\det A \neq 0$, il modo più naturale per trovare la soluzione è sicuramente $\mathbf{x} = A^{-1}\mathbf{b}$. Esiste un modo equivalente, che consente di trovare la soluzione componente per componente, evitando così il calcolo della matrice inversa. Si tratta della cosiddetta

Regola di Cramer: dati $A \in \text{Mat}_n(\mathbb{R})$ non singolare, $\mathbf{b} \in \mathbb{R}^n$, l'unica soluzione del sistema $A\mathbf{x} = \mathbf{b}$ è il vettore $\mathbf{x} \in \mathbb{R}^n$ la cui i-esima componente è

$$x_i = \frac{\det A_i}{\det A}$$

dove A_i è la matrice ottenuta da A sostituendo all'i-esima colonna il vettore $\mathbf{b}$.

Osservazione 5.16. *La Regola di Cramer si giustifica osservando che*

$$\mathbf{x} = A^{-1}\mathbf{b} = \frac{1}{\det A} C^T \mathbf{b},$$

dove C *è la matrice dei complementi algebrici di* A. *Pertanto*

$$x_i = \frac{1}{\det A}(b_1 c_{1i} + b_2 c_{2i} + \cdots + b_n c_{ni}) = \frac{1}{\det A} \det A_i$$

dove A_i *è la matrice ottenuta da* A *sostituendo all'*i*-esima colonna il vettore* $\mathbf{b}$.

Esempio 5.17. *Risolvere il sistema*

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

Abbiamo

$$A = \begin{bmatrix} 1 & -2 & 0 \\ 1 & 1 & 1 \\ 0 & 3 & -1 \end{bmatrix}, \quad \mathbf{b} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad A|\mathbf{b} = \begin{bmatrix} 1 & -2 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 3 & -1 & 0 \end{bmatrix}.$$

La matrice dei coefficienti è quadrata. Si ha $\det A = -6 \neq 0$. *Dal Teorema di Cramer 5.14 sappiamo che esiste un'unica soluzione* $\mathbf{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$. *Calcoliamo dunque la soluzione* *con la regola di Cramer. Si ha*

$$x = \frac{\det A_1}{\det A} = \frac{\begin{bmatrix} 0 & -2 & 0 \\ 1 & 1 & 1 \\ 0 & 3 & -1 \end{bmatrix}}{-6} = \frac{-2}{-6} = \frac{1}{3}$$

$$y = \frac{\det A_2}{\det A} = \frac{\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 0 & -1 \end{bmatrix}}{-6} = \frac{-1}{-6} = \frac{1}{6}$$

$$z = \frac{\det A_3}{\det A} = \frac{\begin{bmatrix} 1 & -2 & 0 \\ 1 & 1 & 1 \\ 0 & 3 & 0 \end{bmatrix}}{-6} = \frac{-3}{-6} = \frac{1}{2}.$$

Vediamo ora come possiamo calcolare le soluzioni di un sistema in tutti i casi diversi da quello quadrato con $\det A \neq 0$.

Dato il sistema $A\mathbf{x} = \mathbf{b}$, supponiamo di aver trovato che $\mathrm{rg}(A) = \mathrm{rg}(A|\mathbf{b}) = r$, dunque per il Teorema di Rouché-Capelli 5.5 il sistema ammette soluzioni. Indichiamo con $\bar{A}$ una sottomatrice di A di ordine r. Riscriviamo il sistema eliminando le eventuali equazioni corrispondenti a righe di A che non figurano in $\bar{A}$, e "portando a secondo membro" le eventuali incognite relative a colonne di A che non figurano in $\bar{A}$. Abbiamo così ottenuto un sistema di r equazioni, in cui le r incognite relative alle colonne di A che figurano in $\bar{A}$ vengono espresse in funzione delle altre $n - r$, che a questo punto diventano parametri arbitrari. È chiaro che, per ogni scelta dei valori di questi $n - r$ parametri, il sistema ha una e una soluzione perché è un sistema quadrato del tipo $\bar{A}\bar{\mathbf{x}} = \bar{\mathbf{b}}$ con $\det \bar{A} \neq 0$. Per il calcolo esplicito delle soluzioni basta dunque risolvere il nuovo sistema con la regola di Cramer, considerando parametri le incognite che appaiono a secondo membro.

Esempio 5.18. *Risolvere il sistema*

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

Abbiamo

$$A = \begin{bmatrix} 1 & -1 & 1 \\ -1 & 1 & 1 \end{bmatrix}, \quad \mathbf{b} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \quad A|\mathbf{b} = \begin{bmatrix} 1 & -1 & 1 & 2 \\ 1 & 1 & 1 & 1 \end{bmatrix}.$$

Risulta $\mathrm{rg}(A) = \mathrm{rg}(A|\mathbf{b}) = 2$. *Dal Teorema di Rouché-Capelli 5.5 il sistema è possibile. Inoltre per il Teorema 5.11 lo spazio delle sue soluzioni ha dimensione* $n - \mathrm{rg}(A) = 3 - 2 = 1$. *Scegliamo come $\bar{A}$ la matrice formata dalla prima a dalla terza colonna di A:*

$$\bar{A} = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}, \quad \det \bar{A} = 2 \neq 0.$$

Riscriviamo il sistema dato nel modo equivalente

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

ovvero $\bar{A}\bar{\mathbf{x}} = \bar{\mathbf{b}}$, *con* $\bar{\mathbf{b}} = \begin{pmatrix} 2+y \\ 1-y \end{pmatrix}$, $\bar{\mathbf{x}} = \begin{pmatrix} x \\ z \end{pmatrix}$ *Applichiamo la regola di Cramer a* $\bar{A}\bar{\mathbf{x}} = \bar{\mathbf{b}}$. *Si ottiene*

$$x = \frac{\det \bar{A}_1}{\det \bar{A}} = \frac{\begin{bmatrix} 2+y & 1 \\ 1-y & 1 \end{bmatrix}}{2} = \frac{2+y-1+y}{2} = y + \frac{1}{2}$$

$$z = \frac{\det \bar{A}_2}{\det \bar{A}} = \frac{\begin{bmatrix} 1 & 2+y \\ -1 & 1-y \end{bmatrix}}{2} = \frac{1-y+2+y}{2} = \frac{3}{2}.$$

Concludiamo che le soluzioni del sistema $A\mathbf{x} = \mathbf{b}$ *sono date dall'insieme*

$$S = \left\{ \begin{pmatrix} y+\frac{1}{2} \\ y \\ \frac{3}{3} \end{pmatrix}, \ y \in \mathbb{R} \right\} = \left\{ y \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \begin{pmatrix} \frac{1}{2} \\ 0 \\ \frac{3}{2} \end{pmatrix}, \ y \in \mathbb{R} \right\}$$
$$= \left\{ \operatorname{span} \left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \right\} + \begin{pmatrix} \frac{1}{2} \\ 0 \\ \frac{3}{2} \end{pmatrix} \right\}.$$

Osservazione 5.19. *L'esempio 5.18 poteva anche essere risolto scegliendo come* $\bar{A}$ *la matrice formata dalla seconda a dalla terza colonna di* A*:*

$$\bar{A} = \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix}, \qquad \det \bar{A} = -2 \neq 0.$$

In tal caso il sistema viene riscritto come

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

e la regola di Cramer fornisce

$$y = \frac{\begin{bmatrix} 2-x & 1 \\ 1+x & 1 \end{bmatrix}}{-2} = \frac{2-x-1-x}{-2} = x - \frac{1}{2}$$

$$z = \frac{\begin{bmatrix} -1 & 2-x \\ 1 & 1+x \end{bmatrix}}{2} = \frac{-1-x-2+x}{-2} = \frac{3}{2}.$$

Concludiamo che le soluzioni

$$S = \left\{ \begin{pmatrix} x \\ x-\frac{1}{2} \\ \frac{3}{2} \end{pmatrix}, \ x \in \mathbb{R} \right\} = \left\{ x \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ -\frac{1}{2} \\ \frac{3}{2} \end{pmatrix}, \ x \in \mathbb{R} \right\}$$
$$= \left\{ \operatorname{span} \left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \right\} + \begin{pmatrix} 0 \\ -\frac{1}{2} \\ \frac{3}{2} \end{pmatrix} \right\}.$$

È chiaro che questo è un modo solo formalmente diverso di scrivere le soluzioni trovate nell'Esempio 5.18: al variare dei parametri in $\mathbb{R}$ *si ottiene lo stesso insieme di* $\mathbb{R}^3$. *Si osservi che non era invece possibile esprimere le soluzioni in funzione di* z *in quanto la sottomatrice formata dalla prima e seconda colonna è singolare.*

Capitolo 6

Diagonalizzazione

6.1 Autovalori ed autovettori

Sia V uno spazio vettoriale su $\mathbb{R}$ di dimensione n, e $f : V \to V$ un'applicazione lineare. Fissata una base $\mathcal{B}_V$ in V, sia $A \in \mathrm{Mat}_n(\mathbb{R})$ la matrice associata ad f rispetto a $\mathcal{B}_V$.

Definizione 6.1. *Uno scalare $\lambda \in \mathbb{R}$ si dice autovalore di f (o di A) se esiste un vettore $v \in V$, con $v \neq 0_V$, tale che*

$$f(v) = \lambda v \quad (Av = \lambda v).$$

In tal caso, v si dice autovettore di f (o di A) relativo all'autovalore λ.

Osservazione 6.2.

1. *Nella Definizione 6.1 si chiede che sia $v \neq 0_V$. Infatti, se fosse $v = 0_V$, allora ogni scalare λ sarebbe un autovalore perché*

$$f(0_V) = 0_V = \lambda\, 0_V \quad \forall \lambda \in \mathbb{R}.$$

2. *Può accadere che $\lambda = 0$, cioè $\lambda = 0$ può essere un autovalore. Più precisamente, questo accade se e solo se esiste $v \in V$, $v \neq 0_V$, tale che*

$$f(v) = 0\, v = 0_V$$

cioè se e solo se $\mathrm{Ker}(f) \neq \{0_V\}$.

Definizione 6.3. *Sia λ un autovalore per f. Definiamo*

$$V_\lambda = \{v \in V : \; v \neq 0_V \text{ e } f(v) = \lambda v\} \cup \{0_V\}, \tag{6.1}$$

cioè V_λ è l'insieme di tutti gli autovettori di f relativi all'autovalore λ, a cui aggiungiamo il vettore nullo.

Proposizione 6.4. *V_λ è un sottospazio di V.*

Dim. Si ha

(i) $0_V \in V_\lambda$;

(ii) $\forall v, w \in V_\lambda \Rightarrow v + w \in V_\lambda$. Infatti

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

(iii) $\forall v \in V_\lambda$, $\forall \alpha \in \mathbb{R} \Rightarrow \alpha v \in V_\lambda$. Infatti

$$f(\alpha v) = \alpha f(v) = \alpha \lambda v = \lambda(\alpha v).$$

□

Definizione 6.5. *Il sottospazio V_λ in (6.1) si dice autospazio relativo all'autovalore λ. La dimensione di V_λ si dice molteplicità geometrica dell'autovalore λ e si indica con $m_g(\lambda)$.*

Proposizione 6.6. *Lo scalare $\lambda \in \mathbb{R}$ è un autovalore per f se e solo se l'applicazione lineare*

$$\begin{aligned} f - \lambda 1_V : & V \to V \\ & v \mapsto f(v) - \lambda v \end{aligned}$$

non è invertibile.

Dim. Per la Definizione 6.1, λ è un autovalore di f se e solo se esiste $v \in V$, $v \neq 0_V$, tale che

$$f(v) = \lambda v,$$

ovvero se e solo se

$$0_V = f(v) - \lambda v = (f - \lambda 1_V)(v),$$

ovvero se e solo se

$$v \in \operatorname{Ker}(f - \lambda 1_V).$$

Ricordando che un'applicazione lineare $g : V \to V$ è invertibile se e solo se $\operatorname{Ker}(g) = \{0_V\}$, concludiamo che λ è un autovalore per f se e solo se l'applicazione $f - \lambda 1_v$ non è invertibile.
□

Osservazione 6.7. *In termini di matrici, la Proposizione 6.6 ci dice che lo scalare $\lambda \in R$ è un autovalore per la matrice $A \in \operatorname{Mat}_n(\mathbb{R})$ se e solo se la matrice*

$$A - \lambda \mathbb{I}_n$$

non è invertibile, ovvero se e solo $\det(A - \lambda \mathbb{I}_n) = 0$.

Definizione 6.8. *Data una matrice $A \in \operatorname{Mat}_n(\mathbb{R})$, si dice polinomio caratteristico di A il polinomio nell'indeterminata x:*

$$p_A(x) := \det(A - x\mathbb{I}_n).$$

Si dice inoltre equazione caratteristica l'equazione

$$p_A(x) = 0.$$

Osservazione 6.9. *Il polinomio caratteristico $p_A(x)$ di $A \in \mathrm{Mat}_n(\mathbb{R})$ ha grado n.*

Esempio 6.10. *Sia*

$$A = \begin{bmatrix} 1 & 3 \\ 1 & -1 \end{bmatrix}.$$

Il polinomio caratteristico di A è

$$p_A(x) = \det(A - x\mathbb{I}_2) = \det\left(\begin{bmatrix} 1 & 3 \\ 1 & -1 \end{bmatrix} - x\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right) = \det\begin{bmatrix} 1-x & 3 \\ 1 & -(1+x) \end{bmatrix}$$
$$= -(1-x)(1+x) - 3 = -(1-x^2) - 3 = x^2 - 4.$$

L'equazione caratteristica è dunque $x^2 - 4 = 0$.

Osservazione 6.11. *Per quanto visto nella Proposizione 6.4 e nell'Osservazione 6.7, si ha che λ è un autovalore di A se e solo se λ è soluzione di $p_A(x) = 0$ ovvero se e solo se $p_A(\lambda) = 0$.*

Definizione 6.12. *Se λ è un autovalore di A, si dice molteplicità algebrica di λ, e si indica con $m_a(\lambda)$, la molteplicità di λ come radice di $p_A(x)$.*

Esempio 6.13. *Torniamo all'Esempio 6.10. Abbiamo*

$$p_A(x) = x^2 - 4 = (x-2)(x+2).$$

Gli autovalori di A sono

$$\lambda_1 = 2, \quad m_a(2) = 1$$
$$\lambda_2 = -2, \quad m_a(-2) = 1$$

Calcoliamo gli autospazi associati ai due autovalori.

$$V_2 = \left\{\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 : \ A\begin{pmatrix} x \\ y \end{pmatrix} = 2\begin{pmatrix} x \\ y \end{pmatrix}\right\}.$$

Si ha

$$A\begin{pmatrix} x \\ y \end{pmatrix} = 2\begin{pmatrix} x \\ y \end{pmatrix} \quad \Leftrightarrow \quad \begin{bmatrix} 1 & 3 \\ 1 & -1 \end{bmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2x \\ 2y \end{pmatrix}$$
$$\Leftrightarrow \quad \begin{cases} x + 3y = 2x \\ x - y = 2y \end{cases}$$
$$\Leftrightarrow \quad \begin{cases} x = 3y \\ y \end{cases}$$

quindi

$$V_2 = \left\{\begin{pmatrix} 3y \\ y \end{pmatrix} : \ y \in \mathbb{R}\right\} = \operatorname{span}\left\{\begin{pmatrix} 3 \\ 1 \end{pmatrix}\right\}$$

e $m_g(2) = \dim(V_2) = 1$. Analogamente,

$$V_{-2} = \left\{\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 : \ A\begin{pmatrix} x \\ y \end{pmatrix} = -2\begin{pmatrix} x \\ y \end{pmatrix}\right\}.$$

Si ha

$$A\begin{pmatrix} x \\ y \end{pmatrix} = -2\begin{pmatrix} x \\ y \end{pmatrix} \quad \Leftrightarrow \quad \begin{bmatrix} 1 & 3 \\ 1 & -1 \end{bmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -2x \\ -2y \end{pmatrix}$$
$$\Leftrightarrow \quad \begin{cases} x + 3y = -2x \\ x - y = -2y \end{cases}$$
$$\Leftrightarrow \quad \begin{cases} x = -y \\ y \end{cases}$$

quindi

$$V_{-2} = \left\{ \begin{pmatrix} -y \\ y \end{pmatrix} : y \in \mathbb{R} \right\} = \operatorname{span}\left\{ \begin{pmatrix} -1 \\ 1 \end{pmatrix} \right\}$$

e $m_g(-2) = \dim(V_{-2}) = 1$.

Esempio 6.14. *Sia*

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

Si ha

$$p_A(x) = \det \begin{bmatrix} 1-x & 0 & 0 \\ 0 & -x & 0 \\ 1 & 0 & 1-x \end{bmatrix} = -x\,(1-x)^2,$$

pertanto risolvendo

$$-x\,(1-x)^2 = 0$$

si trovano gli autovalori

$$\lambda_1 = 0, \quad m_a(1) = 1,$$
$$\lambda_2 = 1, \quad m_a(1) = 2.$$

Calcoliamo gli autospazi associati. Si ha

$$V_0 = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 : A\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\}.$$

$$A\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad \Leftrightarrow \quad \begin{cases} x = 0 \\ x + z = 0 \end{cases}$$
$$\Leftrightarrow \quad \begin{cases} x = 0 \\ y \\ z = 0 \end{cases}$$

quindi

$$V_0 = \left\{ \begin{pmatrix} 0 \\ y \\ 0 \end{pmatrix} : y \in \mathbb{R} \right\} = \operatorname{span}\left\{ \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\}$$

e $m_g(0) = \dim(V_0) = 1$. *Analogamente,*

$$V_1 = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 : \; A \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \right\}$$
$$= \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 : \; (A - \mathbb{I}_3) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\}.$$

Si ha

$$A - \mathbb{I}_3 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

e dunque

$$(A - \mathbb{I}_3) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad \Leftrightarrow \quad \begin{cases} x = 0 \\ y = 0 \\ z \end{cases}$$

Si ottiene

$$V_1 = \left\{ \begin{pmatrix} 0 \\ 0 \\ z \end{pmatrix} : \; z \in \mathbb{R} \right\} = \operatorname{span} \left\{ \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$$

e $m_g(1) = \dim(V_1) = 1$.

Osservazione 6.15. *Se λ è un autovalore di A, allora*

$$1 \leq m_g(\lambda) \leq m_a(\lambda).$$

Teorema 6.16. *Se $A, B \in \mathrm{Mat}_n(\mathbb{R})$ sono due matrici simili, allora*

1. *A e B hanno lo stesso determinante;*
2. *A e B hanno lo stesso polinomio caratteristico;*
3. *A e B hanno gli stessi autovalori.*

Osservazione 6.17. *Si dice che la traccia, il determinante, il polinomio caratteristico e gli autovalori sono* invarianti *della trasformazione. Infatti, mentre le matrici che la rappresentano possono variare, le entità indicate invece sono proprie della trasformazione e non della sua rappresentazione.*

Dim. Poiché A e B sono simili, esiste una matrice invertibile S tale che $B = S^{-1}AS$, si veda la Definizione 3.28.
(1) Usando il Teorema di Binet si ha

$$\det B = \det(S^{-1}AS) = \det S^{-1} \det A \det S = \det S^{-1} \det S \det A = \det(S^{-1}S) \det A = \det A.$$

(2) Si ha

$$\det(B-\lambda\mathbb{I}_n) = \det(S^{-1}AS-\lambda\mathbb{I}_n) = \det(S^{-1}AS-S^{-1}\lambda\mathbb{I}_nS)$$
$$= \det(S^{-1}(A-\lambda\mathbb{I}_n)S) = \det S^{-1}\det(A-\lambda\mathbb{I}_n)\det S = \det(A-\lambda\mathbb{I}_n).$$

(3) Segue direttamente da (2) essendo gli autovalori le radici del polinomio caratteristico.
□

Definizione 6.18. *Se $f: V \to V$ è una applicazione lineare e A è la matrice associata ad f rispetto ad una qualunque base $\mathcal{B}_V$ di V, chiameremo polinomio caratteristico di f il polinomio caratteristico di A.*

6.2 Diagonalizzabilità

6.2.1 Diagonalizzabilità per una applicazione lineare

Definizione 6.19. *Una applicazione lineare $f: V \to V$ è diagonalizzabile se esiste una base di V composta da autovettori di f.*

Osservazione 6.20. *Sia f una applicazione lineare diagonalizzabile, e sia $\mathcal{B}_V = \{v_1, ..., v_n\}$ una base di V composta da autovettori per f. Sarà quindi*

$$f(v_i) = \lambda_i v_i \quad \text{per } \lambda_i \in \mathbb{R}.$$

La matrice A associata a f rispetto a $\mathcal{B}_V$ è dunque la matrice diagonale

$$A = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_n \end{bmatrix}.$$

Vogliamo trovare un criterio di diagonalizzabilità. Cominciamo con il provare il seguente risultato.

Proposizione 6.21. *Sia $f: V \to V$ una applicazione lineare e siano $v_1, ..., v_k$ autovettori di f relativi agli autovalori $\lambda_1, ..., \lambda_k$ rispettivamente. Se $\lambda_s \neq \lambda_t$ per $s \neq t$ allora i vettori $v_1, ..., v_k$ sono linearmente indipendenti.*

Dim. Dimostriamo il risultato per induzione su k.

- *Caso* $k = 1$. Sia v_1 autovettore di f relativo all'autovalore λ_1. Abbiamo $v_1 \neq 0_V$ (v_1 è autovettore), quindi v_1 è linearmente indipendente.
- *Caso* $k-1$. Supponiamo valga il risultato (ipotesi induttiva).
- *Caso* k. Siano $v_1, ..., v_k$ autovettori di f relativi agli autovalori $\lambda_1, ..., \lambda_k$ rispettivamente, con $\lambda_s \neq \lambda_t$ per $s \neq t$. Si ha

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k = 0_V \tag{6.2}$$

per certi $\alpha_1, ..., \alpha_k \in \mathbb{R}$. Applicando f troviamo

$$\begin{aligned} 0_V &= f(0_V) \\ &= f(\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k) \\ &= \alpha_1 f(v_1) + \alpha_2 f(v_2) + \cdots + \alpha_k f(v_k) \\ &= \lambda_1 \alpha_1 f(v_1) + \lambda_2 \alpha_2 f(v_2) + \cdots + \lambda_k \alpha_k f(v_k). \end{aligned} \tag{6.3}$$

Moltiplichiamo (6.2) per λ_k e otteniamo

$$0_V = \lambda_k \alpha_1 v_1 + \lambda_k \alpha_2 v_2 + \cdots + \lambda_k \alpha_k v_k. \tag{6.4}$$

Sottraendo la (6.4) dalla (6.3) si ottiene

$$0_V = (\lambda_1 - \lambda_k)\alpha_1 v_1 + (\lambda_2 - \lambda_k)\alpha_2 v_2 + \cdots + (\lambda_{k-1} - \lambda_k)\alpha_{k-1} v_{k-1}. \tag{6.5}$$

Ma per ipotesi $v_1, ..., v_{k-1}$ sono indipendenti, quindi

$$\begin{aligned} \alpha_1(\lambda_1 - \lambda_k) &= 0 \\ \alpha_2(\lambda_2 - \lambda_k) &= 0 \\ &\vdots \\ \alpha_{k-1}(\lambda_{k-1} - \lambda_k) &= 0 \end{aligned}$$

da cui abbiamo, essendo per ipotesi $\lambda_s \neq \lambda_t$, $t \neq s$,

$$\alpha_1 = \alpha_2 = ... = \alpha_{k-1} = 0.$$

Ritornando alla (6.2), otteniamo dunque

$$\alpha_k v_k = 0_V$$

da cui, essendo $v_k \neq 0_V$ (v_k è un autovettore) si conclude che $\alpha_k = 0$.

□

Come conseguenza, abbiamo il seguente risultato.

Corollario 6.22. *Sia $f : V \to V$ una applicazione lineare, e sia $n = \dim(V)$. Se f ha n autovettori distinti, allora f è diagonalizzabile.*

Dim. Siano $\lambda_1, ..., \lambda_n$ gli autovalori distinti di f, e $v_1, ..., v_n$ i rispettivi autovettori. Per la Proposizione 6.21, i vettori $v_1, ..., v_n$ sono linearmente indipendenti e pertanto formano una base di V. □

Il Corollario 6.26 fornisce una condizione *sufficiente* per la diagonalizzabilità di una applicazione lineare. Vediamo ora, senza dimostrarlo, il caso generale.

Teorema 6.23. *Una applicazione lineare $f : V \to V$ è diagonalizzabile se e solo se il polinomio caratteristico di f ha tutte le sue radici in $\mathbb{R}$, e la molteplicità geometrica di ciascun autovalore coincide con la sua molteplicità algebrica.*

Osservazione 6.24. *Dire che il polinomio caratteristico $p(x)$ di f ha tutte le radici reali significa che, una volta fattorizzato, non ha fattori irriducibili di grado maggiore o uguale a 2.*

6.2.2 Diagonalizzabilità per una matrice

Ripetiamo quanto visto nella Sezione 6.2.1 per le matrici.

Definizione 6.25. *Una matrice $A \in \mathrm{Mat}_n(\mathbb{R})$ è diagonalizzabile se è simile ad una matrice diagonale, cioè se esiste una matrice S invertibile tale che*

$$S^{-1}AS$$

è una matrice diagonale.

Corollario 6.26. *Se $A \in \mathrm{Mat}_n(\mathbb{R})$ ha n autovettori distinti, allora A è diagonalizzabile.*

Teorema 6.27. *Una matrice $A \in \mathrm{Mat}_n(\mathbb{R})$ è diagonalizzabile se e solo se il polinomio caratteristico $p_A(x)$ di A ha tutte le sue radici in $\mathbb{R}$, e la molteplicità geometrica di ciascun autovalore coincide con la sua molteplicità algebrica.*

Osservazione 6.28. *Se $A \in \mathrm{Mat}_n(\mathbb{R})$ diagonalizzabile, allora la matrice*

$$S := [v_1 \,|\, v_2 \,|\, \cdots \,|\, v_k]$$

le cui colonne $\{v_1, v_2, ..., v_k\}$ sono le componenti di una base di autovettori per A è tale che

$$S^{-1}AS = \begin{bmatrix} \lambda_1 & 0 \cdots & 0 \\ 0 & \ddots & 0 \\ 0 & 0 \cdots & \lambda_n \end{bmatrix}.$$

Teorema 6.29 (Teorema Spettrale)**.** *Sia $A \in \mathrm{Mat}_n(\mathbb{R})$ matrice simmetrica. Allora A è diagonalizzabile. Inoltre, esiste una matrice ortogonale U tale che*

$$U^T AU$$

è diagonale.

6.3 Procedimento di diagonalizzazione

Andiamo ora a considerare il procedimento di diagonalizzazione per una matrice $A \in \mathrm{Mat}_n(\mathbb{R})$, ovvero la costruzione (quando possibile) di una matrice invertibile S tale che $S^{-1}AS$ sia diagonale.

1. Calcolare il polinomio caratteristico di A

$$p_A(x) = \det(A - x\mathbb{I}_n).$$

2. Fattorizzare $p_A(x)$.

    (2.a) Nella fattorizzazione di A compare un polinomio irriducibile di ordine maggiore o uguale a 2
    $\Rightarrow A$ non è diagonalizzabile. STOP

    (2.b) Scriviamo
    $$p_A(x) = C(x-\lambda_1)^{a_1}(x-\lambda_2)^{a_2}\cdots(x-\lambda_s)^{a_s}$$
    con $C \in \mathbb{R}$, $\lambda_1, \lambda_2, ..., \lambda_s \in \mathbb{R}$, $a_1 + a_2 + \cdots + a_s = n$.

3. Gli autovalori di A sono le radici di $p_A(x)$ cioè sono $\lambda_1, \lambda_2, ..., \lambda_s$. Inoltre, l'autovalore λ_i ha molteplicità algebrica a_i, $i = 1, ..., s$, ovvero
$$\begin{aligned} m_a(\lambda_1) &= a_1 \\ m_a(\lambda_2) &= a_2 \\ &\vdots \\ m_a(\lambda_s) &= a_s \end{aligned}$$

4. Determinare la molteplicità geometrica di λ_1 data da
$$m_g(\lambda_1) = \dim(V_{\lambda_1})$$
dove V_{λ_1} è l'autospazio relativo all'autovalore λ_1:
$$V_{\lambda_1} = \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n : \ (A - \lambda_1 \mathbb{I}_n) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \right\}.$$
Lo spazio delle soluzioni del sistema
$$(A - \lambda_1 \mathbb{I}_n) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \tag{6.6}$$
ha dimensione $n - \text{rg}(A - \lambda_1 \mathbb{I}_n)$ (si veda la Proposizione 5.9), perciò
$$\dim(V_{\lambda_1}) = n - \text{rg}(A - \lambda_1 \mathbb{I}_n).$$

    (4.a) $m_g(\lambda_1) \neq m_a(\lambda_1) \Rightarrow A$ non è diagonalizzabile. STOP

    (4.b) $m_g(\lambda_1) = m_a(\lambda_1)$. Determinare una base $\{\mathbf{v}_1, \mathbf{v}_2, ...\mathbf{v}_{a_1}\}$ per V_{λ_1}.

5. Tornare al punto (4) con λ_2 al posto di λ_1, e così via fino all'autovalore λ_s.

6. Se A risulta diagonalizzabile, la matrice S che ha come colonne i vettori che formano una base per V_{λ_1}, V_{λ_2}, ..., V_{λ_s} è tale che
$$S^{-1}AS = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_n \end{bmatrix}.$$
dove λ_i è ripetuto a_i volte, $i = 1, ..., s$.

Esempio 6.30. *Si consideri la matrice*

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

1. *Si dica se la matrice A è diagonalizzabile.*
2. *In caso affermativo, si determinino una matrice diagonale D e una matrice di passaggio S tali che $D = S^{-1}AS$.*

1. Si ha

$$\begin{aligned} p_A(x) = \det(A - x\mathbb{I}_3) &= \det \begin{bmatrix} -1-x & 2 & -3 \\ 2 & 2-x & -6 \\ -1 & -2 & 1-x \end{bmatrix} \\ &= (-1-x)[(2-x)(1-x)-12] - 2[2(1-x)-6] - 3[-4+2-x] \\ &= (-1-x)(-3x+x^2-10) + 4(x+2) + 3(x+2) \\ &= (-x^3+2x^2+13x+10) + 7(x+2) \\ &= -x^3+2x^2+20x+24. \end{aligned}$$

Applicando la regola di Ruffini con $x_0 = -2$ a $-x^3+2x^2+20x+24$ ottengo

$$-x^3+2x^2+20x+24 = (-x^2+4x+12)(x+2)$$

perciò

$$p_A(x) = (-x^2+4x+12)(x+2).$$

Applicando la regola di Ruffini con $x_0 = 6$ a $-x^2+4x+12$ ottengo

$$-x^2+4x+12 = -(x+2)(x-6)$$

e concludiamo che

$$p_A(x) = -(x-6)(x+2)^2.$$

Abbiamo dunque gli autovalori

$$\begin{aligned} \lambda_1 = 6 \quad & m_a(6) = 1 \\ \lambda_2 = -2 \quad & m_a(-2) = 2. \end{aligned}$$

Consideriamo gli autospazi associati. Si ha

$$V_{\lambda_1} = V_6 = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 : \ (A - 6\mathbb{I}_3) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\}$$

dove

$$A - 6\mathbb{I}_3 = \begin{bmatrix} -7 & 2 & -3 \\ 2 & -4 & -6 \\ -1 & -2 & -5 \end{bmatrix}.$$

Poiché per esempio

$$\det \bar{A} = \det \begin{bmatrix} -7 & 2 \\ 2 & -4 \end{bmatrix} = 28 - 4 = 24 \neq 0,$$

si ha $\mathrm{rg}(A - 6\mathbb{I}_3) = 2$. *Segue che*

$$m_g(6) = \dim(V_6) = n - \mathrm{rg}(A - 6\mathbb{I}_3) = 3 - 2 = 1 = m_a(6).$$

Passiamo dunque all'autospazio associato al secondo autovalore. Si ha

$$V_{\lambda_2} = V_{-2} = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 : \ (A + 2\mathbb{I}_3) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\}$$

dove

$$A + 2\mathbb{I}_3 = \begin{bmatrix} 1 & 2 & -3 \\ 2 & 4 & -6 \\ -1 & -2 & 3 \end{bmatrix}.$$

Poiché tutti i minori di ordine 2 *sono nulli, concludiamo che* $\mathrm{rg}(A + 2\mathbb{I}_3) = 1$. *Segue che*

$$m_g(-2) = \dim(V_{-2}) = n - \mathrm{rg}(A + 2\mathbb{I}_3) = 3 - 1 = 2 = m_a(-2).$$

Dunque la matrice A è diagonalizzabile.

2. Resta da calcolare una base rispettivamente per V_6 e V_{-2}. Si ha

$$(A - 6\mathbb{I}_3) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad \Leftrightarrow \quad \begin{cases} -7x - 2y - 3z & = 0 \\ 2x - 4y - 6z & = 0 \\ -x - 2y - 5z & = 0 \end{cases}$$

$$\Leftrightarrow \quad \begin{cases} -7x + 2y = 3z \\ 2x - 4y = 6z \end{cases}$$

$$\Leftrightarrow \quad \bar{A} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 3z \\ 6z \end{pmatrix}$$

con $\bar{A} = \begin{bmatrix} -7 & 2 \\ 2 & -4 \end{bmatrix}$, $\det \bar{A} = 24$. *Applicando la regola di Cramer, otteniamo*

$$x = \frac{\begin{bmatrix} 3z & 2 \\ 6z & -4 \end{bmatrix}}{24} = \frac{-12z - 12z}{-24} = -z$$

$$y = \frac{\begin{bmatrix} -7 & 3z \\ 2 & 6z \end{bmatrix}}{24} = \frac{-42z - 6z}{-24} = -2z.$$

Concludiamo che

$$V_6 = \left\{ \begin{pmatrix} -z \\ -2z \\ z \end{pmatrix} : z \in \mathbb{R} \right\} = \mathrm{span} \left\{ \begin{pmatrix} -1 \\ -2 \\ 1 \end{pmatrix} \right\}.$$

Passiamo a V_{-2}. Si ha

$$A + 2\mathbb{I}_3 \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad \Leftrightarrow \quad \begin{cases} x + 2y - 3z & = 0 \\ 2x + 4y - 6z & = 0 \\ -x - 2y + 3z & = 0 \end{cases}$$

$$\Leftrightarrow \quad \begin{cases} x = -2y + 3z \\ y \\ z \end{cases}$$

e concludiamo che

$$V_{-2} = \left\{ \begin{pmatrix} -2y + 3z \\ y \\ z \end{pmatrix} : y,\, z \in \mathbb{R} \right\} = \operatorname{span} \left\{ \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} \begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix} \right\}.$$

Abbiamo dunque

$$S = \begin{bmatrix} -1 & -2 & 3 \\ -2 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}, \quad D = S^{-1}AS = \begin{bmatrix} 6 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{bmatrix}.$$

Capitolo 7

Forme quadratiche

7.1 Definizione e segno

Definizione 7.1. *Si dice forma quadratica (in breve f.q.) nelle indeterminate* $x_1, x_2, ..., x_n$ *ogni polinomio omogeneo di grado 2 in* $x_1, x_2, ..., x_n$ *(cioè tutti i monomi hanno grado 2).*
Posto $\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$, *scriviamo una f.q. come*

$$Q(\mathbf{x}) = \sum_{i,j=1}^{n} a_{ij} x_i x_j. \tag{7.1}$$

Esempio 7.2. 1. $Q(\mathbf{x}) = x_1^2 + 3x_1x_2 - 6x_2^2$ *è una f.q. in* x_1, x_2.

2. $Q(\mathbf{x}) = x_1x_2 + x_3^2$ *è una f.q. in* x_1, x_2, x_3.

3. *Il polinomio* $x_1^2 + x_2^2 - 2x_1$ *non definisce una f.q. in* x_1, x_2, *in quanto non è omogeneo di grado 2.*

Data la f.q. $Q(\mathbf{x})$ della forma (7.1), definiamo la matrice $\tilde{A} = (a_{ij}) \in \mathrm{Mat}_n(\mathbb{R})$ in modo che

$$Q(\mathbf{x}) = \mathbf{x}^T \tilde{A} \mathbf{x}.$$

Esempio 7.3. 1. *La matrice della f.q. dell'Esempio 7.2.1 è*

$$\tilde{A} = \begin{bmatrix} 1 & 3 \\ 0 & -6 \end{bmatrix}.$$

2. *La matrice della f.q. dell'Esempio 7.2.2 è*

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

Data la f.q. $Q(\mathbf{x})$ della forma (7.1) possiamo sempre scriverla nella forma

$$Q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}, \quad A \in \mathrm{Mat}_n(\mathbb{R}) \text{ simmetrica.}$$

Infatti, per ogni $i, j \in \{1, ..., n\}$,

$$a_{ij}x_ix_j + a_{ji}x_jx_i = \frac{a_{ij}+a_{ji}}{2}x_ix_j + \frac{a_{ij}+a_{ji}}{2}x_jx_i.$$

Esempio 7.4. *1. La f.q. dell'Esempio 7.2.1 può essere riscritta come*

$$\begin{aligned} Q(\mathbf{x}) &= x_1^2 + 3x_1x_2 - 6x_2^2 \\ &= x_1 + \frac{3}{2}x_1x_2 + \frac{3}{2}x_2x_1 - 6x_2^2 \\ &= \mathbf{x}^T A\mathbf{x}, \quad \mathbf{x} = (x_1, x_2)^T \in \mathbb{R}^2, \end{aligned}$$

con

$$A = \begin{bmatrix} 1 & \frac{3}{2} \\ \frac{3}{2} & -6 \end{bmatrix}.$$

2. La f.q. dell'Esempio 7.2.2 può essere riscritta come

$$\begin{aligned} Q(\mathbf{x}) &= x_1x_2 + x_3^2 \\ &= \frac{1}{2}x_1x_2 + \frac{1}{2}x_2x_1 + x_3^2 \\ &= \mathbf{x}^T A\mathbf{x}, \quad \mathbf{x} = (x_1, x_2, x_3)^T \in \mathbb{R}^3, \end{aligned}$$

con

$$A = \begin{bmatrix} 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Definizione 7.5. *Una f.q. (o la corrispondente $A \in \mathrm{Mat}_n(\mathbb{R})$ simmetrica) si dice*

1. *definita positiva se $Q(\mathbf{x}) > 0$ per ogni $\mathbf{x} \neq \mathbf{0}_{\mathbb{R}^n}$;*
2. *definita negativa se $Q(\mathbf{x}) < 0$ per ogni $\mathbf{x} \neq \mathbf{0}_{\mathbb{R}^n}$;*
3. *semidefinita positiva se $Q(\mathbf{x}) \geq 0$ per ogni $\mathbf{x} \in \mathbb{R}^n$, e esiste $\mathbf{x}_0 \neq \mathbf{0}_{\mathbb{R}^n}$ tale che $Q(\mathbf{x}_0) = 0$;*
4. *semidefinita positiva se $Q(\mathbf{x}) \leq 0$ per ogni $\mathbf{x} \in \mathbb{R}^n$, e esiste $\mathbf{x}_0 \neq \mathbf{0}_{\mathbb{R}^n}$ tale che $Q(\mathbf{x}_0) = 0$;*
5. *indefinita se esistono $\mathbf{x}_0, \mathbf{y}_0 \neq \mathbf{0}_{\mathbb{R}^n}$ tali che $Q(\mathbf{x}_0) > 0$, $Q(\mathbf{y}_0) < 0$.*

Diamo ora dei metodi operativi per riconoscere una forma quadratica.

Teorema 7.6. *Sia* $Q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ *una f.q. con* $A \in \mathrm{Mat}_n(\mathbb{R})$ *simmetrica. Dal Teorema Spettrale 6.29 sappiamo che* A *è diagonalizzabile. Allora la f.q. è*

1. *definita positiva se e solo se tutti gli autovalori di* A *sono strettamente positivi;*
2. *definita negativa se e solo se tutti gli autovalori di* A *sono strettamente negativi;*
3. *semidefinita positiva se e solo se tutti gli autovalori di* A *sono positivi o nulli;*
4. *semidefinita negativa se e solo se tutti gli autovalori di* A *sono negativi o nulli;*
5. *indefinita se esistono sia autovalori di* A *positivi che autovalori di* A *negativi.*

Esempio 7.7. *Si consideri la f.q.*

$$\begin{aligned} Q(\mathbf{x}) &= 9x_1^2 + 12x_1x_2 + 4x_2^2 \\ &= 9x_1^2 + 6x_1x_2 + 6x_2x_1 + 4x_2^2 \\ &= \mathbf{x}^T A\mathbf{x}, \quad \mathbf{x} = (x_1, x_2)^T \in \mathbb{R}^2, \end{aligned}$$

con

$$A = \begin{bmatrix} 9 & 6 \\ 6 & 4 \end{bmatrix}.$$

Gli autovalori di A *sono le soluzioni di*

$$0 = p_A(\lambda) = \det(A - \lambda\mathbb{I}_2) = (9-\lambda)(4-\lambda) - 36 = -13\lambda - \lambda^2 = \lambda(\lambda - 13),$$

ovvero, $\lambda_1 = 0$, $\lambda_2 = 13$. *Pertanto, per il Teorema 7.6-(3),* $Q(\mathbf{x})$ *è semidefinita positiva.*

Esempio 7.8. *Si consideri la f.q.*

$$\begin{aligned} Q(\mathbf{x}) &= 2x_1^2 - 2x_1x_2 + 2x_2^2 + 2x_2x_3 + 2x_3^2 \\ &= 2x_1^2 - x_1x_2 - x_2x_1 + 2x_2^2 + x_2x_3 + x_3x_2 + 2x_3^2 \\ &= \mathbf{x}^T A\mathbf{x}, \quad \mathbf{x} = (x_1, x_2, x_3)^T \in \mathbb{R}^3, \end{aligned}$$

con

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

Gli autovalori di A *sono le soluzioni di*

$$\begin{aligned} 0 = p_A(\lambda) = \det(A - \lambda\mathbb{I}_2) &= (2-\lambda)[(2-\lambda)^2 - 1] - (2-\lambda) \\ &= (2-\lambda)[(2-\lambda)^2 - (\sqrt{2})^2] \\ &= (2-\lambda)(2-\lambda-\sqrt{2})(2-\lambda+\sqrt{2}), \end{aligned}$$

ovvero, $\lambda_1 = 2$, $\lambda_2 = 2+\sqrt{2}$, $\lambda_3 = 2-\sqrt{2}$. *Pertanto, per il Teorema 7.6-(1),* $Q(\mathbf{x})$ *è definita positiva.*

Definizione 7.9. *Sia* $A \in \mathrm{Mat}_n(\mathbb{R})$ *simmetrica. Chiamiamo*

- *minori principali di ordine* k *di* A *i minori che si ottengono considerando* k *righe di* A *e le corrispondenti colonne;*
- *minori principali di Nord Ovest (NO in breve) di ordine* k *di* A *i minori principali che si ottengono considerando le prime* k *righe di* A *e le corrispondenti colonne.*

Esempio 7.10. *Consideriamo la generica* $A \in \mathrm{Mat}_3(\mathbb{R})$*:*

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

I minori principali di A *sono*

$$\begin{aligned} k = 1 \quad & a_{11} \quad a_{22}, \quad a_{33}; \\ k = 2 \quad & \det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \quad \det \begin{bmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{bmatrix}, \quad \det \begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix}; \\ k = 3 \quad & \det A. \end{aligned}$$

I minori principali di NO di A *sono*

$$\begin{aligned} k = 1 \quad & a_{11}; \\ k = 2 \quad & \det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}; \\ k = 3 \quad & \det A. \end{aligned}$$

Teorema 7.11. *Sia* $Q(\mathbf{x}) = \mathbf{x}^T A \mathbf{x}$ *una f.q. con* $A \in \mathrm{Mat}_n(\mathbb{R})$ *simmetrica. La f.q. è*

1. *definita positiva se e solo se tutti i minori principali di NO di* A *sono positivi;*
2. *definita negativa se e solo se tutti i minori principali di NO di* A *di ordine pari sono positivi e tutti i minori principali di NO di* A *di ordine dispari sono negativi;*
3. *semidefinita positiva se e solo tutti i minori principali di* A *sono positivi o nulli, e* $\det A = 0$*;*
4. *semidefinita negativa se e solo se tutti i minori principali di* A *di ordine pari sono positivi o nulli, tutti i minori principali di* A *di ordine dispari sono negativi o nulli, e* $\det A = 0$*;*
5. *indefinita se non rientra nei casi precedenti.*

Esempio 7.12. *Torniamo all'Esempio 7.8. I tre minori principali di NO di* A *sono*

$$k = 1 \quad a_{11} = 2;$$

$$
\begin{aligned}
&k = 2 \quad \det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \det \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} = 3; \\
&k = 3 \quad \det A = 2(4-1) + 1(-2) = 4.
\end{aligned}
$$

Dunque per il Teorema 7.11-(1) $Q(\mathbf{x})$ è definita positiva.

Esempio 7.13. *Si consideri la f.q.*

$$
\begin{aligned}
Q(\mathbf{x}) &= -2x_1^2 + 4x_1x_2 - 4x_2^2 + 5x_2x_3 - 2x_3^2 \\
&= -2x_1^2 + 2x_1x_2 + 2x_2x_1 - 4x_2^2 + \frac{5}{2}x_2x_3 + \frac{5}{2}x_3x_2 - 2x_3^2 \\
&= \mathbf{x}^T A\mathbf{x}, \quad \mathbf{x} = (x_1, x_2, x_3)^T \in \mathbb{R}^3,
\end{aligned}
$$

con

$$
A = \begin{bmatrix} -2 & 2 & 0 \\ 2 & -4 & \frac{5}{2} \\ 0 & \frac{5}{2} & -2 \end{bmatrix}.
$$

I tre minori principali di NO di A sono

$$
\begin{aligned}
&k = 1 \quad a_{11} = -2; \\
&k = 2 \quad \det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \det \begin{bmatrix} -2 & 2 \\ 2 & -4 \end{bmatrix} = 4; \\
&k = 3 \quad \det A = -2\left(8 - \frac{25}{4}\right) - 2(-4) = \frac{9}{2}.
\end{aligned}
$$

Dunque, per il Teorema 7.11-(5), $Q(\mathbf{x})$ è indefinita.

7.2 Generalizzazioni del Teorema Spettrale e applicazioni alla statistica

7.2.1 Decomposizione ai Valori Singolari (SVD)

La decomposizione a valori singolari è una generalizzazione della decomposizione spettrale e, grazie al Teorema di Eckart-Young 7.23, è probabilmente la più importante decomposizione matriciale per le applicazioni all'analisi dei dati. Essa, infatti, fornisce la soluzione ai problemi di approssimazione ottima di matrici di dati nella norma di Frobenius (si veda l'Osservazione 7.21) e permette di unificare numerose procedure di sintesi di dati multidimensionali, largamente utilizzate nella pratica statistica.

Osservazione 7.14. *Sia $X \in \mathrm{Mat}_{n,k}(\mathbb{R})$, $k \le n$, con $\mathrm{rg}(X) = k$. La matrice*

$$
X^T X \in \mathrm{Mat}_k(\mathbb{R})
$$

è simmetrica definita positiva e quindi per il Teorema 7.6-(1) ed il Teorema Spettrale 6.29, ammette k autovettori ortonormali $\mathbf{v}_1, \ldots, \mathbf{v}_k$, corrispondenti a k autovalori positivi $\sigma_1^2 \ge \sigma_2^2 \ge \ldots \ge \sigma_k^2 > 0$.

Definizione 7.15. *Sia $X \in \mathrm{Mat}_{n,k}(\mathbb{R})$, $k \leq n$, con $\mathrm{rg}(X) = k$. Le radici quadrate $\sigma_1, \sigma_2, \ldots, \sigma_k$ degli autovalori di X^TX (prese con il segno positivo) sono dette valori singolari di X.*

Teorema 7.16 (Decomposizione a valori singolari (SVD))**.** *Sia $X \in \mathrm{Mat}_{n,k}(\mathbb{R})$ con $\mathrm{rg}(X) = k$. Siano $\mathbf{v}_1, \ldots, \mathbf{v}_k$ gli autovettori ortonormali di X^TX, e $\sigma_1, \sigma_2 \ldots, \sigma_k$ i valori singolari di X, ordinati in senso decrescente. Vale la decomposizione a valori singolari:*

$$X = UDV^T \tag{7.2}$$

dove

- $U \in \mathrm{Mat}_{n,k}(\mathbb{R})$ *è tale che* $U^TU = \mathbb{I}_k$;
- $D = diag(\sigma_1, \sigma_2 \ldots, \sigma_k)$;
- $V \in \mathrm{Mat}_k(\mathbb{R})$ *è la matrice ortogonale avente come colonne* $\mathbf{v}_1, \ldots, \mathbf{v}_k$.

Dim. Poniamo $U := XVD^{-1}$. Moltiplicando a destra entrambi i membri per DV^T, otteniamo:

$$UDV^T = XVD^{-1}DV^T = X$$

dove abbiamo usato che $VV^T = \mathbb{I}_k$ per l'ortogonalità di V. La matrice U soddisfa le richieste del teorema:

$$U^TU = D^{-1}V^TX^TXVD^{-1} = D^{-1}V^TVD^2D^{-1} = D^{-1}D^2D^{-1} = \mathbb{I}_k,$$

dove abbiamo usato che $X^TXV = VD^2$. □

Osservazione 7.17. *Le colonne della matrice U sono autovettori normalizzati della matrice XX^T, relativi agli autovalori $\sigma_1^2, \sigma_2^2 \ldots, \sigma_k^2$. Infatti:*

$$\begin{aligned} XX^TU &= XX^TXVD^{-1} = X(X^TXV)D^{-1} = XVD^2D^{-1} \\ &= (XVD^{-1})D^2 = UD^2. \end{aligned}$$

Osservazione 7.18. *Sia $f : \mathbb{R}^k \to \mathbb{R}^n$ un'applicazione lineare associata a X. I vettori $\mathbf{v}_1, \ldots, \mathbf{v}_k$ costituiscono una base di $\mathbb{R}^k$ e i vettori $\mathbf{u}_1, \ldots, \mathbf{u}_k$ costituiscono una base di $Im(f)$ (che è un sottospazio k-dimensionale di $\mathbb{R}^n$). In particolare, da (7.2) si ha*

$$X\mathbf{v}_i = UDV^T\mathbf{v}_i = \sigma_i\mathbf{u}_i.$$

La matrice X associa pertanto ad ogni elemento della base ortonormale dello spazio di input, un elemento della base ortonormale del sottospazio di output, dilatato di un fattore σ_i. La specificità della decomposizione SVD è proprio quella di individuare due basi ortonormali collegate dalla relazione $\mathbf{v}_i \mapsto \sigma_i\mathbf{u}_i$.

Definizione 7.19. *Sia $X \in \mathrm{Mat}_{n,k}(\mathbb{R})$ con decomposizione a valori singolari data da (7.2). Gli autovettori di X^TX (colonne di V) sono detti vettori singolari destri di X, mentre gli autovettori di XX^T (colonne di U) sono detti vettori singolari sinistri di X.*

Osservazione 7.20. *Abbiamo formulato la decomposizione SVD nel caso di matrici rettangolari $X \in \mathrm{Mat}_{n,k}(\mathbb{R})$, con $\mathrm{rg}(X) = k \leq n$, perché questo è il caso più frequente nella pratica dell'analisi dei dati. Tuttavia, è possibile estendere il risultato anche agli altri casi:*

- $\mathrm{rg}(X) = n < k$. *È sufficiente trasporre la matrice di partenza e procedere come in precedenza.*
- $\mathrm{rg}(X) = p < k \leq n$. *È sufficiente notare che anche le matrici $X^T X$ e XX^T hanno rango p e che quindi la decomposizione SVD sarà costruita a partire dalla matrice $\tilde{V} \in \mathrm{Mat}_{k,p}(\mathbb{R})$, avente come colonne i p autovettori di $X^T X$ relativi ai p autovalori non nulli, dalla matrice diagonale $\tilde{D} \in \mathrm{Mat}_p(\mathbb{R})$, avente sulla diagonale i p valori singolari non nulli di X, e dalla matrice $\tilde{U} \in \mathrm{Mat}_{n,p}(\mathbb{R})$ tale che $\tilde{U} = X\tilde{V}\tilde{D}^{-1}$.*

Forma alternativa della SVD. La decomposizione (7.2) può anche essere posta nella seguente forma:

$$X = UDV^T = \sum_{i=1}^{k} \sigma_i \mathbf{u}_i \mathbf{v}_i^T = \sum_{i=1}^{k} \sigma_i Z_i, \tag{7.3}$$

dove $\mathbf{u}_i$ e $\mathbf{v}_i$ sono le colonne i-esime di U e di V, rispettivamente e $Z_i := \mathbf{u}_i \mathbf{v}_i^T$.

Osservazione 7.21. *Le matrici Z_i in (7.3) hanno, per costruzione, rango pari a 1 e formano un insieme di "vettori" ortonormali rispetto al prodotto scalare di Frobenius, definito come*

$$\langle A, B \rangle_F := \sum_{i,j=1}^{n} a_{ij} b_{ij} = Tr(A^T B).$$

Infatti:

$$\langle Z_i, Z_j \rangle_F = Tr(Z_i^T Z_j) = Tr(\mathbf{v}_i \mathbf{u}_i^T \mathbf{u}_j \mathbf{v}_j^T) = Tr(\mathbf{v}_i \delta_{ij} \mathbf{v}_j^T) = \delta_{ij} Tr(\mathbf{v}_j^T \mathbf{v}_i) = \delta_{ij}^2 = \delta_{ij},$$

dove δ_{ij}, detta delta di Kronecker, vale 1 se $i = j$ e 0 altrimenti.

Osservazione 7.22. *Scritta nella forma (7.3), la SVD ricostruisce la matrice di input X come sovrapposizione di "layers" (le matrici Z_i), e non per singole entrate.*

Data una matrice X, il seguente risultato stabilisce come costruire la matrice $\hat{X}$ di rango fissato, che meglio approssima X, nella norma di Frobenius.

Teorema 7.23 (Teorema di Eckart-Young). *Sia $X \in \mathrm{Mat}_{n,k}(\mathbb{R})$ con $\mathrm{rg}(X) = k$, e sia $X = UDV^T$ la decomposizione a valori singolari di X. Sia $0 < p \leq k$ un intero fissato, e*

- *$U_{[p]} \in \mathrm{Mat}_{n,p}(\mathbb{R})$ la matrice composta dalle prime p colonne di U;*
- *$V_{[p]} \in \mathrm{Mat}_{k,p}(\mathbb{R})$ la matrice composta dalle prime p colonne di V;*
- *$D_{[p]} \in \mathrm{Mat}_{p}(\mathbb{R})$ la matrice composta dalle prime p righe e prime p colonne di D.*

La matrice

$$\hat{X} := U_{[p]} D_{[p]} V_{[p]}^T = \sum_{i=1}^{p} \sigma_i Z_i$$

minimizza la distanza $\|X - \cdot\|_F$ nell'insieme $\mathrm{Mat}_{n,k}(\mathbb{R})$ di rango p, ovvero

$$\hat{X} = \min_{\substack{Y \in \mathrm{Mat}_{n,k}(\mathbb{R}): \\ \mathrm{rg}(Y)=p}} \|X - Y\|_F$$

Dim. Per le proprietà della norma di Frobenius, possiamo scrivere

$$\|X - \hat{X}\|_F^2 = \|UDV^T - \hat{X}\|_F^2 = \|D - U^T \hat{X} V\|_F^2 = \|D - M\|_F^2$$

dove $M := U^T \hat{X} V$ è una matrice di rango p. Pertanto:

$$\|X - \hat{X}\|_F^2 = \sum_{i,j=1}^{k} (D_{ij} - M_{ij})^2 = \sum_{i=1}^{k} (\sigma_i - M_{ii})^2 + \sum_{i \neq j=1}^{k} M_{ij}^2 \geq \sum_{i=1}^{k} (\sigma_j - M_{ii})^2.$$

Da ciò segue che la matrice minimizzante M deve essere diagonale, cosicché

$$\|X - \hat{X}\|_F^2 = \sum_{i=1}^{k} (\sigma_j - M_{ii})^2.$$

Avendo M rango p, la sua diagonale contiene solo p valori non nulli; tra tutte le matrici diagonali $n \times k$ di rango p, quella che minimizza $\|X - \hat{X}\|_F^2$ è quindi data da

$$M = diag(\sigma_1, \ldots, \sigma_p, 0, \ldots, 0)$$

e la matrice $\hat{X}$ è quindi

$$\hat{X} = UMV^T = U_{[p]} D_{[p]} V_{[p]}^T.$$

$\square$

Osservazione 7.24. *In pratica, per ottenere la migliore approssimazione di rango p della matrice X di input, è sufficiente costruire la SVD di X e porre a zero tutti gli elementi della diagonale di D, dal $k+1$-esimo in poi o, equivalentemente, espandere X in termini delle matrici Z_i e considerare solo i primi p termini. L'errore che si compie in questo modo è:*

$$e^2 := \|X - \hat{X}\|_F^2 = \left\| \sum_{i=1}^{k} \sigma_i Z_i - \sum_{i=1}^{p} \sigma_i Z_i \right\|_F^2 = \left\| \sum_{i=p+1}^{k} \sigma_i Z_i \right\|_F^2 = \sum_{i=p+1}^{k} \sigma_i^2.$$

7.2.2 Decomposizione QR

La decomposizione QR permette di rappresentare una qualunque matrice quadrata come prodotto di una matrice ortogonale Q e di una matrice triangolare superiore R.

Teorema 7.25. *Sia $X \in \mathrm{Mat}_n(\mathbb{R})$. Vale la decomposizione*

$$X = QR$$

dove

- *$Q \in \mathrm{Mat}_n(\mathbb{R})$ è ortogonale;*
- *$R \in \mathrm{Mat}_n(\mathbb{R})$ è triangolare superiore.*

Dim. Distinguiamo due casi.

- $rg(X) = n$. Le colonne di X sono linearmente indipendenti. Definita Q come la matrice di vettori ortonormali ottenuta attraverso la procedura di Gram-Schmidt applicata alle colonne di X stessa, la decomposizione non esprime altro che la ricostruzione delle colonne di X sulle colonne di Q. Infatti $R := Q^T X$ è triangolare superiore perché, per costruzione, la h-esima ($h > 1$) colonna di Q è ortogonale alle colonne $1, \ldots, h-1$ di X e quindi il generico elemento di matrice $R_{ij} = \mathbf{q}_i^T \mathbf{x}_j$ (dove $\mathbf{q}_i$ e $\mathbf{x}_j$ sono l'i-esima e la j-esima colonna di Q e di X, rispettivamente) è nullo per $i > j$.
- $rg(X) = k < n$. Si indichino con $i_1, \ldots, i_k$ gli indici delle prime k colonne di X linearmente indipendenti e con S il sottospazio lineare k-dimensionale di $\mathbb{R}^n$ da esse generato. Si costruisca Q inserendo

  (i) nelle colonne $i_1, \ldots, i_k$, i vettori ortonormali ottenuti applicando la procedura di Gram-Schmidt alle colonne $i_1, \ldots, i_k$ di X;

  (ii) nelle rimanenti colonne, i vettori di una qualunque base ortonormale del sottospazio $S^\perp$ ortogonale a S.

  La matrice $R := Q^T X$ è, come nel caso precedente, triangolare superiore.

□

Osservazione 7.26. *Si noti che la decomposizione QR non è unica. Infatti, se Q ed R realizzano la decomposizione, allora anche $-Q$ e $-R$ la realizzano. Inoltre, nel caso in cui X non sia a rango pieno, la base ortonormale di $S^\perp$ può essere scelta in modo arbitrario, e inserita nelle opportune colonne di Q in modi diversi, generando decomposizioni differenti.*

7.2.3 Decomposizione di Cholesky

La decomposizone di Cholesky garantisce che una qualunque matrice quadrata simmetrica definita positiva si possa scrivere come prodotto di due matrici, una triangolare inferiore e una triangolare superiore, l'una la trasposta dell'altra.

Teorema 7.27. *Sia $X \in \mathrm{Mat}_n(\mathbb{R})$ una matrice simmetrica definita positiva. Vale la decomposizione*

$$X = R^T R$$

con $R \in \mathrm{Mat}_n(\mathbb{R})$ triangolare superiore.

Dim. Essendo X simmetrica, dal Teorema Spettrale 6.29 esiste una matrice ortogonale U tale che

$$X = UDU^T.$$

Inoltre, poiché X è definita positiva, per il Teorema 7.6-(1) gli elementi di D sono tutti positivi e quindi X può essere scritta come

$$X = UD^{1/2}D^{1/2}U^T = B^T B$$

con $B := D^{1/2}U^T$. Applicando la decomposizione QR a B data nel Teorema 7.25, otteniamo (ricordando l'ortogonalità di Q)

$$X = R^T Q^T QR = R^T R,$$

con R matrice triangolare superiore. □

La decomposizione di Cholesky è molto utilizzata nelle simulazioni statistiche, anche per il fatto che essa può essere numericamente implementata in modo efficiente. In particolare, essa è spesso usata per generare osservazioni da variabili casuali correlate: è sufficiente generare variabili casuali centrate, normalizzate e indipendenti (cosa, in generale, facile) e poi trasformare i campioni così ottenuti.

Esempio 7.28. *Sia $\Sigma \in \mathrm{Mat}_k(\mathbb{R})$ una matrice simmetrica definita positiva con decomposizione di Cholesky data dal Teorema 7.27*

$$\Sigma = R^T R.$$

Allora una matrice $Y \in \mathrm{Mat}_{n,k}(\mathbb{R})$ avente matrice di varianze-covarianze Σ_y pari a Σ è data da

$$Y = AR,$$

con $A \in \mathrm{Mat}_{n,k}(\mathbb{R})$ con colonne centrate, di varianza unitaria e incorrelate:

$$\Sigma_a := \frac{A^T A}{n} = \mathbb{I}_n.$$

Infatti:

$$\Sigma_y := \frac{Y^T Y}{n} = R^T \frac{A^T A}{n} R = R^T \mathbb{I}_n R = R^T R = \Sigma.$$