

Lezione 21

21.1 Prodotti scalari

Definizione 21.1 (Prodotto scalare). Sia V uno spazio vettoriale su \mathbb{R} .

Un *prodotto scalare* su V è un'applicazione

$$\begin{aligned} \langle \cdot, \cdot \rangle: V \times V &\longrightarrow \mathbb{R} \\ (v_1, v_2) &\mapsto \langle v_1, v_2 \rangle \end{aligned}$$

tale che valgono i seguenti assiomi:

- (PS1) per ogni $v_1, v_2 \in V$, si ha $\langle v_1, v_2 \rangle = \langle v_2, v_1 \rangle$ (*il prodotto scalare è commutativo*);
- (PS2) per ogni $v_1, v_2, v_3 \in V$, si ha $\langle v_1, v_2 + v_3 \rangle = \langle v_1, v_2 \rangle + \langle v_1, v_3 \rangle$ (*il prodotto scalare è distributivo rispetto alla somma*);
- (PS3) per ogni $\alpha \in \mathbb{R}$ e $v_1, v_2 \in V$, si ha $\alpha \langle v_1, v_2 \rangle = \langle \alpha v_1, v_2 \rangle$;
- (PS4) per ogni $v \in V \setminus \{0_V\}$, si ha $\langle v, v \rangle > 0$ (*il prodotto scalare è definito positivo*).

Osservazione 21.2. Elenchiamo alcune immediate conseguenze degli assiomi della definizione precedente.

1. Sia $v_0 \in V$ fissato. Allora l'applicazione

$$\begin{aligned} \langle \cdot, v_0 \rangle: V &\longrightarrow \mathbb{R} \\ v &\mapsto \langle v, v_0 \rangle \end{aligned}$$

è lineare. Per la commutatività del prodotto scalare segue anche la linearità dell'applicazione

$$\begin{aligned} \langle v_0, \cdot \rangle: V &\longrightarrow \mathbb{R} \\ v &\mapsto \langle v_0, v \rangle. \end{aligned}$$

Per questa doppia proprietà di linearità si dice spesso che il prodotto scalare è un'*applicazione bilineare*.

2. Per linearità, si ha che $\langle 0, v \rangle = \langle v, 0 \rangle = 0$ per ogni $v \in V$: in particolare $\langle v, v \rangle = 0$ se e solo se $v = 0$.
3. Se $W \subseteq V$ è un sottospazio, ha senso considerare la restrizione $\langle \cdot, \cdot \rangle|_{W \times W}$, che è un prodotto scalare su W .

Definizione 21.3 (Modulo di un vettore). Sia V uno spazio vettoriale su \mathbb{R} munito di prodotto scalare $\langle \cdot, \cdot \rangle$.

Per ogni $v \in V$, il numero $|v| = \sqrt{\langle v, v \rangle}$ si dice *modulo* di v . I vettori di modulo 1 si dicono *versori*.

Si noti che in uno spazio vettoriale V su \mathbb{R} munito di prodotto scalare l'unico vettore di modulo zero è il vettore nullo $\vec{0}_V$.

Esempio 21.4. Nello spazio $V_3(O)$ dei vettori applicati si può definire un prodotto scalare ponendo

$$\langle \vec{v}, \vec{w} \rangle = |\vec{v}| |\vec{w}| \cos(\widehat{\vec{v}\vec{w}}) \quad (21.1.1)$$

per ogni coppia di vettori $\vec{v}, \vec{w} \in V_3(O)$ non nulli, come avevamo visto nella Lezione 8.

Fissato un sistema di riferimento $O\vec{i}\vec{j}\vec{k}$, se $\vec{v} = v_x\vec{i} + v_y\vec{j} + v_z\vec{k}$ e $w = w_x\vec{i} + w_y\vec{j} + w_z\vec{k}$, è noto che

$$\langle \vec{v}, \vec{w} \rangle = v_x w_x + v_y w_y + v_z w_z. \quad \spadesuit$$

Esempio 21.5. Dati $x = (x_1, \dots, x_n)$ e $y = (y_1, \dots, y_n)$ in $V = \mathbb{R}^n$, si definisce *prodotto scalare standard* il prodotto scalare

$$\langle x, y \rangle = x_1 y_1 + \dots + x_n y_n. \quad (21.1.2)$$

Si noti che in tal caso risulta

$$\langle x, y \rangle = {}^t x y = {}^t x I_n y$$

come prodotto di matrici. Il fatto che tale applicazione soddisfi gli assiomi (PS1), (PS2) e (PS3) è evidente dalla definizione. Per quanto riguarda l'assioma (PS4), si noti che

$$\langle x, x \rangle = x_1^2 + \dots + x_n^2$$

è una somma di numeri reali non negativi, come lo sono i quadrati di numeri reali, quindi è non negativa ed è nulla se e solo se tutti gli addendi sono nulli.

⚠ Quello standard non è l'unico possibile prodotto scalare che possiamo definire su \mathbb{R}^n !

Per esempio si verifichi che l'applicazione

$$((x_1, x_2), (y_1, y_2)) \mapsto 3x_1 y_1 + x_2 y_2 / 2$$

è un prodotto scalare in \mathbb{R}^2 diverso dal prodotto standard. Si noti che

$$3x_1 y_1 + x_2 y_2 / 2 = (x_1 \ x_2) \begin{pmatrix} 3 & 0 \\ 0 & 1/4 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}. \quad \spadesuit$$

Sia V uno spazio vettoriale su \mathbb{R} munito di prodotto scalare $\langle \cdot, \cdot \rangle$. Dati due vettori $v, w \in V \setminus \{0_V\}$, per ogni $t \in \mathbb{R}$ si ha

$$|v|^2 - 2t\langle v, w \rangle + t^2|w|^2 = \langle v - tw, v - tw \rangle \geq 0.$$

Il primo membro di tale trinomio non può avere radici distinte, dovendo altrimenti cambiare di segno, quindi $\langle v, w \rangle^2 - |v|^2|w|^2 \leq 0$: essendo $|v|, |w| > 0$, segue allora la cosiddetta *disuguaglianza di Cauchy-Schwartz* (si veda l'Osservazione 8.9).

Proposizione 21.6 (Disuguaglianza di Cauchy–Schwartz). Sia V uno spazio vettoriale su \mathbb{R} munito di prodotto scalare $\langle \cdot, \cdot \rangle$. Per ogni $v, w \in V$ si ha

$$|\langle v, w \rangle| \leq |v||w| \quad (21.1.3)$$

Inoltre vale l'uguaglianza in (21.1.3) se e solo se v e w sono proporzionali.

Dimostrazione. Rimane da dimostrare solo la seconda affermazione. Vale l'uguaglianza in (21.1.3) se e solo se l'equazione $|v|^2 - 2t\langle v, w \rangle + t^2|w|^2 = 0$ ha soluzione, ovvero se e solo se $\langle v - tw, v - tw \rangle = 0$ ha soluzione, cioè se e solo se $v = tw$. \square

Osservazione 21.7. Sia V uno spazio vettoriale su \mathbb{R} munito di prodotto scalare $\langle \cdot, \cdot \rangle$. Se $v, w \in V \setminus \{0_V\}$, allora

$$-1 \leq \frac{\langle v, w \rangle}{|v||w|} \leq 1 :$$

possiamo perciò definire l'angolo fra v e w come

$$\widehat{vw} = \arccos \left(\frac{\langle v, w \rangle}{|v||w|} \right).$$

Si ha quindi $\langle v, w \rangle = |v||w| \cos(\widehat{vw})$, che generalizza la formula (21.1.1) a prodotti scalari qualsiasi.

Osservazione 21.8. Sia V uno spazio vettoriale su \mathbb{R} munito di prodotto scalare $\langle \cdot, \cdot \rangle$. Dalla disuguaglianza (21.1.3) ricaviamo anche la *disuguaglianza triangolare* (si veda di nuovo l'Osservazione 8.9):

$$\begin{aligned} |v + w|^2 &= \langle v + w, v + w \rangle = |v|^2 + \langle v, w \rangle + \langle w, v \rangle + |w|^2 \\ &= |v|^2 + 2\langle v, w \rangle + |w|^2 \leq |v|^2 + 2|\langle v, w \rangle| + |w|^2 \\ &\leq |v|^2 + 2|v||w| + |w|^2 = (|v| + |w|)^2, \end{aligned}$$

quindi $|v + w| \leq |v| + |w|$. In modo analogo il lettore verifichi che $|v - w| \geq ||v| - |w||$.

Esempio 21.9. Sia $I = [a, b] \subseteq \mathbb{R}$ non vuoto e si consideri nello spazio $\mathcal{C}^0(I)$ l'applicazione

$$\langle f, g \rangle = \int_a^b f(x)g(x)dx.$$

È immediato verificare che gli assiomi di prodotto scalare (PS1), (PS2), (PS3) sono soddisfatti; inoltre il teorema della permanenza del segno per funzioni continue ci assicura che lo stesso vale per l'assioma (PS4).

In questo caso la disuguaglianza di Cauchy–Schwartz diviene

$$\left| \int_a^b f(x)g(x)dx \right| \leq \sqrt{\int_a^b f(x)^2 dx} \sqrt{\int_a^b g(x)^2 dx}.$$

Per ogni $f \in \mathcal{C}^0(I)$ la quantità

$$|f|_2 = \sqrt{\int_a^b f(x)^2 dx}$$

viene detta *norma L^2 di f* . \spadesuit

21.2 Basi ortonormali

Definizione 21.10. Sia V uno spazio vettoriale su \mathbb{R} munito di prodotto scalare $\langle \cdot, \cdot \rangle$.

- I vettori $v_1, v_2 \in V$ si dicono *ortogonali* (o *perpendicolari*) se $\langle v_1, v_2 \rangle = 0$ e in tal caso si scrive $v_1 \perp v_2$.
- L'insieme $\{v_1, \dots, v_n\} \subseteq V$ si dice *ortogonale* se $v_i \perp v_j$ per $i, j \in I$ con $i \neq j$.
- L'insieme $\{v_1, \dots, v_n\} \subseteq V$ si dice *ortonormale* se è ortogonale e i v_i sono versori.
- Se V è finitamente generato, una *base ortonormale* $\mathcal{B} = (v_1, \dots, v_n)$ è una base di V tale che l'insieme $\{v_1, \dots, v_n\}$ sia ortonormale.

Dalla definizione sopra possiamo dedurre che un insieme $\{v_1, \dots, v_n\} \subseteq V$ è ortonormale se per ogni $i, j = 1, \dots, n$

$$\langle v_i, v_j \rangle = \delta_{i,j} = \begin{cases} 1 & i = j, \\ 0 & i \neq j. \end{cases}$$

Esempio 21.11. Si fissi un sistema di riferimento $O\vec{i}\vec{j}\vec{k}$ nello spazio. Allora l'insieme $\{\vec{i}, \vec{j}, \vec{k}\}$ è ortonormale in $V_3(O)$ rispetto al prodotto scalare geometrico definito nell'Esempio 21.4. Quindi $\mathcal{B} = (\vec{i}, \vec{j}, \vec{k})$ è una base ortonormale di $V_3(O)$. ♠

Esempio 21.12. Si consideri lo spazio \mathbb{R}^n munito del prodotto scalare standard $\langle \cdot, \cdot \rangle$ definito nell'Esempio 21.5. I vettori $\{e_1, \dots, e_n\}$ formano un insieme ortonormale: perciò la base canonica $\mathcal{C} = (e_1, \dots, e_n)$ è una base ortonormale rispetto al prodotto scalare standard.

Invece $\{e_1, e_2\}$ non è ortonormale rispetto al prodotto scalare

$$((x_1, x_2), (y_1, y_2)) \mapsto 3x_1y_1 + x_2y_2/2$$

introdotto nello stesso Esempio. Infatti è vero che $e_1 \perp e_2$, ma $|e_1| = \sqrt{3}$ e $|e_2| = 1/\sqrt{2}$. Concludiamo che, rispetto a tale prodotto scalare, $\{e_1/\sqrt{3}, \sqrt{2}e_2\}$ è ortonormale. ♠

Esempio 21.13. Si consideri lo spazio V delle funzioni continue e periodiche di periodo 2π : per esempio $1, \cos px, \sin px \in V$ per ogni $p \in \mathbb{N}$ (si noti che $\sin px$ e $\cos px$ hanno periodo minimo $2\pi/p$). In V definiamo l'applicazione

$$\langle f, g \rangle = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x)g(x)dx;$$

è facile verificare che $\langle \cdot, \cdot \rangle$ è un prodotto scalare. Inoltre dall'analisi è noto che

$$\begin{aligned} \left\langle \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right\rangle &= 1 \\ \langle \cos px, \cos qx \rangle &= \begin{cases} 1 & \text{se } p = q \neq 0, \\ 0 & \text{se } p \neq q, \end{cases} \\ \langle \sin px, \cos qx \rangle &= 0 \\ \langle \sin px, \sin qx \rangle &= \begin{cases} 1 & \text{se } p = q \neq 0, \\ 0 & \text{se } p \neq q, \end{cases} \end{aligned}$$

quindi, per ogni $N \in \mathbb{N}$, l'insieme $\left\{ \frac{1}{\sqrt{2}}, \cos px, \sin qx \right\}_{p,q=1,\dots,N}$ è ortonormale. ♠

Osserviamo che in tutti gli esempi trattati è stato possibile determinare una base ortonormale. Questo è un risultato generale che si può dimostrare in modo algoritmico, come spiegato nel seguente risultato.

Proposizione 21.14 (Metodo di ortonormalizzazione di Gram–Schmidt).

Sia $V \neq \{0_V\}$ uno spazio vettoriale finitamente generato su \mathbb{R} munito di prodotto scalare $\langle \cdot, \cdot \rangle$. Allora esistono in V basi ortonormali.

Dimostrazione. Sia $\mathcal{B} = (v_1, \dots, v_n)$ una base di V ; vogliamo costruire una base ortonormale a partire da \mathcal{B} . Procediamo nel modo seguente.

Passo (1). Il vettore u_1 è semplicemente uguale a v_1 :

$$u_1 = v_1 .$$

Passo (2). Il vettore u_2 è v_2 meno la proiezione ortogonale di v_2 nella direzione di u_1 , si veda l'Osservazione 8.10:

$$u_2 = v_2 - \frac{\langle v_2, u_1 \rangle}{|u_1|^2} u_1 .$$

Si noti che, per costruzione, $\langle u_2, u_1 \rangle = 0$ e $\mathcal{L}(u_1, u_2) = \mathcal{L}(v_1, v_2)$, quindi u_1 e u_2 sono linearmente indipendenti.

Passo (3). Il vettore u_3 è v_3 cui vengono sottratte le proiezioni ortogonali di v_3 nelle direzioni di u_1 e u_2 :

$$u_3 = v_3 - \frac{\langle v_3, u_1 \rangle}{|u_1|^2} u_1 - \frac{\langle v_3, u_2 \rangle}{|u_2|^2} u_2 .$$

Si noti che, per costruzione, $\langle u_3, u_2 \rangle = \langle u_3, u_1 \rangle = 0$ e $\mathcal{L}(u_1, u_2, u_3) = \mathcal{L}(v_1, v_2, v_3)$, quindi u_1, u_2, u_3 sono linearmente indipendenti.

⋮

Passo (i). Il vettore u_i è v_i cui vengono sottratte le proiezioni ortogonali di v_i nelle direzioni di tutti i vettori precedenti u_1, \dots, u_{i-1} :

$$u_i = v_i - \frac{\langle v_i, u_1 \rangle}{|u_1|^2} u_1 - \frac{\langle v_i, u_2 \rangle}{|u_2|^2} u_2 - \dots - \frac{\langle v_i, u_{i-1} \rangle}{|u_{i-1}|^2} u_{i-1} .$$

Si noti che, per costruzione, $\langle u_i, u_{i-1} \rangle = \dots = \langle u_i, u_1 \rangle = 0$ e $\mathcal{L}(u_1, \dots, u_i) = \mathcal{L}(v_1, \dots, v_i)$, quindi u_1, \dots, u_i sono linearmente indipendenti.

Alla fine otteniamo una base ortogonale (u_1, \dots, u_n) . Normalizzando i vettori ricaviamo la base ortonormale cercata. \square

Esempio 21.15. Data la matrice

$$A = \begin{pmatrix} 1 & -1 & 4 \\ 1 & 4 & -2 \\ 1 & 4 & -2 \\ 1 & -1 & 0 \end{pmatrix} \in \mathbb{R}^{4,3},$$

vogliamo trovare una base ortonormale per lo spazio W generato dalle sue colonne. È immediato verificare che le tre colonne sono linearmente indipendenti, quindi come base \mathcal{B} di partenza possiamo prendere $\mathcal{B} = (v_1, v_2, v_3)$ con

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \quad v_2 = \begin{pmatrix} -1 \\ 4 \\ 4 \\ -1 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 4 \\ -2 \\ -2 \\ 0 \end{pmatrix}.$$

Seguendo l'algoritmo di Gram-Schmidt costruiamo la nuova base ortonormale $\mathcal{B}' = (w_1, w_2, w_3)$ nel modo seguente.

Passo (1). Si ha che

$$u_1 = v_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

Passo (2). Calcoliamo il prodotto scalare $\langle v_2, u_1 \rangle = -1 + 4 + 4 - 1 = 6$. Definiamo poi il vettore

$$u_2 = v_2 - \frac{\langle v_2, u_1 \rangle}{|u_1|^2} u_1 = \begin{pmatrix} -1 \\ 4 \\ 4 \\ -1 \end{pmatrix} - \frac{3}{2} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -5/2 \\ 5/2 \\ 5/2 \\ -5/2 \end{pmatrix}.$$

Passo (3). Calcoliamo i due prodotti scalari $\langle v_3, u_1 \rangle = 0$ e $\langle v_3, u_2 \rangle = -20$ e definiamo il vettore

$$u_3 = v_3 - \frac{\langle v_3, u_1 \rangle}{|u_1|^2} u_1 - \frac{\langle v_3, u_2 \rangle}{|u_2|^2} u_2 = \begin{pmatrix} 4 \\ -2 \\ -2 \\ 0 \end{pmatrix} - 0 \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} + \frac{4}{5} \begin{pmatrix} -5/2 \\ 5/2 \\ 5/2 \\ -5/2 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \\ 0 \\ -2 \end{pmatrix}.$$

Concludiamo normalizzando i 3 vettori, quindi la base ortonormale di W cercata è

$$w_1 = \frac{u_1}{|u_1|} = \begin{pmatrix} 1/2 \\ 1/2 \\ 1/2 \\ 1/2 \end{pmatrix}, \quad w_2 = \frac{u_2}{|u_2|} = \begin{pmatrix} -1/2 \\ 1/2 \\ 1/2 \\ -1/2 \end{pmatrix}, \quad w_3 = \frac{u_3}{|u_3|} = \begin{pmatrix} 1/\sqrt{2} \\ 0 \\ 0 \\ -1/\sqrt{2} \end{pmatrix}$$

Il lettore verifichi che la base \mathcal{B}' così ottenuta è effettivamente ortonormale. \spadesuit

L'importanza e utilità delle basi ortonormali è data dalla seguente proposizione.

Proposizione 21.16. *Sia V uno spazio vettoriale su \mathbb{R} munito di prodotto scalare $\langle \cdot, \cdot \rangle$. Se $\{v_1, \dots, v_n\} \subseteq V$ è un insieme ortonormale, allora:*

- (i) v_1, \dots, v_n sono linearmente indipendenti;
- (ii) se V è finitamente generato e $\dim_{\mathbb{R}}(V) = n$, allora $\mathcal{B} = (v_1, \dots, v_n)$ è una base ortonormale di V e si ha $v = \langle v, v_1 \rangle v_1 + \dots + \langle v, v_n \rangle v_n$ per ogni $v \in V$.

Dimostrazione. Per definizione $\{v_1, \dots, v_n\} \subseteq V$ è un insieme di vettori ortonormali se e solo se

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

quindi, se $\alpha_1 v_1 + \dots + \alpha_n v_n = 0$ è una relazione di dipendenza lineare, si ha

$$0 = \langle 0, v_j \rangle = \langle \alpha_1 v_1 + \dots + \alpha_n v_n, v_j \rangle = \alpha_1 \langle v_1, v_j \rangle + \dots + \alpha_n \langle v_n, v_j \rangle = \alpha_j.$$

In particolare v_1, \dots, v_n sono linearmente indipendenti.

Se V è finitamente generato e $\dim_{\mathbb{R}}(V) = n$, per la Proposizione 16.7, segue che \mathcal{B} è una base di V . In particolare per ogni $v \in V$ esistono $\alpha_1, \dots, \alpha_n \in \mathbb{R}$ tali che $v = \alpha_1 v_1 + \dots + \alpha_n v_n$. Quindi

$$\langle v, v_j \rangle = \langle \alpha_1 v_1 + \dots + \alpha_n v_n, v_j \rangle = \alpha_1 \langle v_1, v_j \rangle + \dots + \alpha_n \langle v_n, v_j \rangle = \alpha_j$$

per ogni $j = 1, \dots, n$. □

Il coefficiente $\langle v, v_j \rangle$ viene detto *coefficiente di Fourier* (di v rispetto a v_j).

Esempio 21.17. Si consideri lo spazio \mathbb{R}^3 munito del prodotto scalare $\langle \cdot, \cdot \rangle$ definito nell'Esempio 21.5. I tre vettori

$$v_1 = \frac{1}{3}(2, 2, 1), \quad v_2 = \frac{1}{3}(1, -2, 2), \quad v_3 = \frac{1}{3}(-2, 1, 2)$$

formano un insieme $\{v_1, v_2, v_3\}$ ortonormale, quindi $\mathcal{B} = (v_1, v_2, v_3)$ è una base di \mathbb{R}^3 .

Sia $v = (1, 1, 1) \in \mathbb{R}^3$. Allora $\langle v, v_1 \rangle = 5/3$, $\langle v, v_2 \rangle = \langle v, v_3 \rangle = 1/3$: quindi, come è anche facile verificare direttamente, risulta

$$v = \frac{5}{3}v_1 + \frac{1}{3}v_2 + \frac{1}{3}v_3. \quad \spadesuit$$

21.3 Matrici ortogonali

In questo paragrafo introduciamo un'importante famiglia di matrici quadrate invertibili.

Definizione 21.18 (Matrici ortogonali). Una matrice $P \in \mathbb{R}^{n,n}$ si dice *ortogonale* se vale l'uguaglianza ${}^t P P = I_n$.

Osservazione 21.19. Sia $P \in \mathbb{R}^{n,n}$ ortogonale.

1. La matrice identità I_n è ortogonale; anche ogni matrice ottenuta da I_n cambiando segno a una o più delle sue entrate è ortogonale.

- Poiché ${}^tPP = I_n$, segue che P è invertibile e $P^{-1} = {}^tP$: in particolare si ha anche $P{}^tP = I_n$. In maniera analoga si dimostra che se $P{}^tP = I_n$ allora anche ${}^tPP = I_n$, cioè P è ortogonale se e solo se $P{}^tP = I_n$.
- Si ha $1 = \det(I_n) = \det({}^tPP) = \det({}^tP) \det(P) = \det(P)^2$, dunque

$$\det(P) = \pm 1 :$$

quanto osservato sulla matrice identità ci permette di affermare che esistono matrici di entrambi i tipi.

- Poiché la riga i -esima di tP è la colonna i -esima P_i di P , la condizione ${}^tPP = I_n$ si può leggere dicendo che il prodotto scalare standard dell'Esempio 21.5 delle colonne P_i e P_j di P è $\delta_{i,j}$: in altre parole una matrice è ortogonale se e solo se le sue colonne sono un insieme ortonormale, rispetto al prodotto scalare standard, di n vettori di \mathbb{R}^n .
- Identificando $\mathbb{R}^{1,n}$ con \mathbb{R}^n tramite la trasposta, si verifica in maniera simile che la condizione $P{}^tP = I_n$ equivale a dire che le righe di P formano un insieme ortonormale di n vettori di \mathbb{R}^n rispetto al prodotto scalare standard.

Le matrici ortogonali si dividono, quindi, in due classi non vuote, quelle con determinante 1 e quelle con determinante -1 .

Definizione 21.20 (Matrici ortogonali speciali e non speciali). Una matrice ortogonale $P \in \mathbb{R}^{n,n}$ si dice *speciale* se $\det(P) = 1$, *non speciale* se $\det(P) = -1$.

Esempio 21.21. Le matrici di $\mathbb{R}^{3,3}$

$$P_1 = \frac{1}{3} \begin{pmatrix} 1 & 2 & 2 \\ 2 & 1 & -2 \\ 2 & -2 & 1 \end{pmatrix}, \quad P_2 = \frac{1}{3} \begin{pmatrix} 1 & 2 & 2 \\ 2 & 1 & -2 \\ -2 & 2 & -1 \end{pmatrix},$$

sono ortogonali. La prima è non speciale, la seconda speciale. ♠

Esempio 21.22. Determiniamo tutte le matrici ortogonali di ordine due; sia

$$P = \begin{pmatrix} p_{1,1} & p_{1,2} \\ p_{2,1} & p_{2,2} \end{pmatrix} \in \mathbb{R}^{2,2}$$

ortogonale. La condizione $P{}^tP = I_2$ si traduce nel sistema

$$\begin{cases} p_{1,1}^2 + p_{1,2}^2 = 1 \\ p_{1,1}p_{2,1} + p_{1,2}p_{2,2} = 0 \\ p_{2,1}^2 + p_{2,2}^2 = 1. \end{cases}$$

La prima e la terza equazione implicano l'esistenza di $\vartheta, \psi \in [0, 2\pi]$ tali che

$$p_{1,1} = \cos \vartheta, \quad p_{1,2} = -\sin \vartheta, \quad p_{2,1} = \sin \psi, \quad p_{2,2} = \cos \psi.$$

La seconda equazione è allora equivalente a

$$0 = \cos \vartheta \sin \psi - \sin \vartheta \cos \psi = \sin(\psi - \vartheta).$$

In particolare, a meno di multipli di 2π , si deve avere o $\psi = \vartheta$ ovvero $\psi = \vartheta + \pi$. Nel primo caso

$$P = \begin{pmatrix} \cos \vartheta & -\sin \vartheta \\ \sin \vartheta & \cos \vartheta \end{pmatrix},$$

(e P è ortogonale speciale); nel secondo

$$P = \begin{pmatrix} \cos \vartheta & -\sin \vartheta \\ -\sin \vartheta & -\cos \vartheta \end{pmatrix}$$

(e P è ortogonale non speciale).

Concludiamo l'esempio con un'interpretazione geometrica delle matrici ortogonali speciali: consideriamo nel piano due sistemi di riferimento $O\vec{i}\vec{j}$ e $O\vec{i}'\vec{j}'$ e sia ψ l'angolo misurato in senso antiorario fra i versori \vec{i} e \vec{i}' . Allora si deve avere $\vec{i}' = a\vec{i} + b\vec{j}$, $\vec{j}' = c\vec{i} + d\vec{j}$ e si ha, per la Proposizione 21.16,

$$a = \langle \vec{i}', \vec{i} \rangle = \cos \varphi, \quad b = \langle \vec{i}', \vec{j} \rangle = \sin \varphi,$$

$$c = \langle \vec{j}', \vec{i} \rangle = -\sin \varphi, \quad d = \langle \vec{j}', \vec{j} \rangle = \cos \varphi.$$

Se ora consideriamo $\vec{v} = x'\vec{i}' + y'\vec{j}' = x\vec{i} + y\vec{j}$, sostituendo le espressioni ottenute sopra di \vec{i}' e \vec{j}' in funzione di \vec{i} e \vec{j} , tenendo conto che (\vec{i}, \vec{j}) è una base di $V_2(O)$, si ottiene

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos \varphi & \sin \varphi \\ -\sin \varphi & \cos \varphi \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Concludiamo che le matrici ortogonali speciali di ordine due corrispondono alle rotazioni nel piano. Per questo spesso indichiamo con R_φ la matrice

$$\begin{pmatrix} \cos \varphi & \sin \varphi \\ -\sin \varphi & \cos \varphi \end{pmatrix}.$$

Nella figura 21.1 qui sotto vediamo un sistema $Ox'y'$ ottenuto da Oxy con una rotazione.

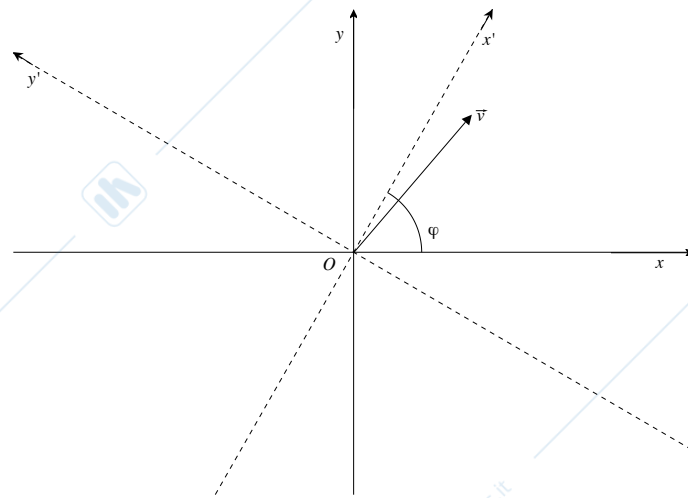


Figura 21.1

Un'analoga interpretazione può essere data per matrici ortogonali in $\mathbb{R}^{n,n}$ con $n \geq 3$. ♠