Spazio Vettoriale Euclideo

maurizio candilera

October 1, 2019

Introduzione

Nello studio dello spazio affine non sono mai entrate in gioco considerazioni relative alla lunghezza dei segmenti e alle distanze tra i punti, ma solo considerazioni basate sulla "proporzionalità" tra figure (dedotta dal prodotto di scalari per vettori) o sulla decomposizione dei vettori tramite la "regola del parallelogramma" (ovvero la somma tra vettori).
In questa ultima parte del corso vogliamo costruire un modello di spazio dotato di una metrica che generalizza la metrica euclidea nel piano e descrivere le trasformazioni rigide di questo spazio, cioè le trasformazioni che conservano le distanze tra punti. Per prima cosa introdurremo la metrica su uno spazio vettoriale (reale) e poi generalizzeremo la costruzione a spazi affini. Per fare questo lavoreremo su spazi vettoriali reali (sarebbe sufficiente avere un campo ordinato in cui tutti gli elementi positivi hanno una radice quadrata) e vedremo come l'introduzione di un prodotto scalare (definito positivo) su tali spazi permetta di definire la misura di distanze e angoli e di caratterizzare le trasformazioni tipiche dello spazio metrico, ovvero quelle che rispettano le distanze.

Prodotto Scalare

Definizione [prodotto scalare su $\mathbb{R}^n$]

Il prodotto scalare su $\mathbb{R}^n$ è l'applicazione $\cdot : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, definita ponendo

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = x_1 y_1 + \cdots + x_n y_n .$$

Scriveremo $v \cdot w$ per indicare il prodotto scalare dei due vettori v e w.
Si chiama *Spazio Vettoriale Euclideo* lo spazio $\mathbb{R}^n$ dotato del prodotto scalare.

Prodotto Scalare

Definizione [prodotto scalare su $\mathbb{R}^n$]

Il prodotto scalare su $\mathbb{R}^n$ è l'applicazione $\cdot : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, definita ponendo

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = x_1 y_1 + \cdots + x_n y_n.$$

Scriveremo $v \cdot w$ per indicare il prodotto scalare dei due vettori v e w.
Si chiama *Spazio Vettoriale Euclideo* lo spazio $\mathbb{R}^n$ dotato del prodotto scalare.

Le seguenti proprietà del prodotto scalare sono di verifica immediata

Proposizione

Il prodotto scalare di $\mathbb{R}^n$ è un'applicazione *bilineare, simmetrica, definita positiva*. Ovvero, per ogni scelta di v, v_1, v_2, w, w_1, w_2 in $\mathbb{R}^n$ e a_1, a_2, b_1, b_2 in $\mathbb{R}$, si ha

(B1) $(a_1 v_1 + a_2 v_2) \cdot w = a_1(v_1 \cdot w) + a_2(v_2 \cdot w)$;

(B2) $v \cdot (b_1 w_1 + b_2 w_2) = b_1(v \cdot w_1) + b_2(v \cdot w_2)$;

(S) $v \cdot w = w \cdot v$;

(P) $v \neq 0 \Rightarrow v \cdot v > 0$.

Ricordiamo che dare un'applicazione bilineare non degenere (qual è il prodotto scalare) determina un isomorfismo tra lo spazio $\mathbb{R}^n$ e il suo duale e permette perciò di rappresentare le forme lineari con vettori di $\mathbb{R}^n$ e di introdurre la nozione di ortogonalità tra i vettori e i sottospazi di $\mathbb{R}^n$. Vedremo che, più in generale, potremo parlare di lunghezze di vettori e di angoli. Procediamo con ordine.

Definizione [norma di un vettore]

Nello spazio vettoriale euclideo si definisce la *norma* di un vettore ponendo

$$\|v\| = \sqrt{v \cdot v} \qquad \text{per ogni } v \in \mathbb{R}^n.$$

Ricordiamo che dare un'applicazione bilineare non degenere (qual è il prodotto scalare) determina un isomorfismo tra lo spazio $\mathbb{R}^n$ e il suo duale e permette perciò di rappresentare le forme lineari con vettori di $\mathbb{R}^n$ e di introdurre la nozione di ortogonalità tra i vettori e i sottospazi di $\mathbb{R}^n$. Vedremo che, più in generale, potremo parlare di lunghezze di vettori e di angoli. Procediamo con ordine.

Definizione [norma di un vettore]

Nello spazio vettoriale euclideo si definisce la *norma* di un vettore ponendo

$$\|v\| = \sqrt{v \cdot v} \qquad \text{per ogni } v \in \mathbb{R}^n.$$

La norma gode delle seguenti proprietà fondamentali

- $\|v\| \geq 0$ per ogni $v \in \mathbb{R}^n$ e $\|v\| = 0 \;\Leftrightarrow\; v = 0$,
- $\|\lambda v\| = |\lambda| \, \|v\|$,
- $\|v + w\| \leq \|v\| + \|w\|$ (disuguaglianza triangolare),

qualunque siano $v, w \in \mathbb{R}^n$ e $\lambda \in \mathbb{R}$.

Le prime due proprietà sono una conseguenza immediata delle proprietà del prodotto scalare, mentre la terza si può dedurre da una più precisa disuguaglianza che andiamo ad enunciare e dimostrare.

Proposizione [disuguaglianza di Cauchy-Schwarz]

Per ogni coppia di vettori $v, w \in \mathbb{R}^n$, si ha

$$|v \cdot w| \leq \|v\| \, \|w\|,$$

e vale l'uguaglianza se, e solo se, v e w sono linearmente dipendenti.

dim. Siano fissati due vettori v e w e osserviamo che, se uno dei due vettori è nullo, la tesi è verificata. Supponiamo ora che v e w siano linearmente indipendenti. Per qualsiasi valore del parametro $t \in \mathbb{R}$, si ha $(v + tw) \cdot (v + tw) \geq 0$ e quindi,

$$v \cdot v + 2tv \cdot w + t^2 w \cdot w \geq 0 \qquad \text{per ogni } t \in \mathbb{R}.$$

Ciò significa che il discriminante di questo trinomio non può essere positivo e quindi deve aversi

$$0 \geq (v \cdot w)^2 - (v \cdot v)(w \cdot w), \qquad \text{ovvero} \qquad (v \cdot w)^2 \leq \|v\|^2 \|w\|^2,$$

da cui si deduce la tesi.

Proposizione [disuguaglianza di Cauchy-Schwarz]

Per ogni coppia di vettori $v, w \in \mathbb{R}^n$, si ha

$$|v \cdot w| \leq \|v\| \, \|w\|,$$

e vale l'uguaglianza se, e solo se, v e w sono linearmente dipendenti.

dim. Siano fissati due vettori v e w e osserviamo che, se uno dei due vettori è nullo, la tesi è verificata.
Supponiamo ora che v e w siano linearmente indipendenti. Per qualsiasi valore del parametro $t \in \mathbb{R}$, si ha $(v + tw) \cdot (v + tw) \geq 0$ e quindi,

$$v \cdot v + 2tv \cdot w + t^2 w \cdot w \geq 0 \qquad \text{per ogni } t \in \mathbb{R}.$$

Ciò significa che il discriminante di questo trinomio non può essere positivo e quindi deve aversi

$$0 \geq (v \cdot w)^2 - (v \cdot v)(w \cdot w), \qquad \text{ovvero} \qquad (v \cdot w)^2 \leq \|v\|^2 \|w\|^2,$$

da cui si deduce la tesi.
Resta quindi da dimostrare l'ultima affermazione sulla dipendenza lineare. Se $w = \lambda v$ per qualche $\lambda \in \mathbb{R}$, allora $\|w\| = |\lambda| \, \|v\|$ e si ha $|v \cdot w| = |\lambda(v \cdot v)| = |\lambda| \, \|v\|^2 = \|v\| \, \|w\|$.
Viceversa, se $|v \cdot w| = \|v\| \, \|w\|$, ovvero $(v \cdot w)^2 = (v \cdot v)(w \cdot w)$; allora, posto $\tau = -\frac{v \cdot w}{w \cdot w}$, si ha

$$(v + \tau w) \cdot (v + \tau w) = v \cdot v + 2\tau(v \cdot w) + \tau^2 (w \cdot w) = v \cdot v - 2\frac{(v \cdot w)^2}{w \cdot w} + \frac{(v \cdot w)^2}{w \cdot w} = 0$$

da cui si deduce che $v + \tau w = 0$ e quindi che v e w sono linearmente dipendenti. □

Dalla disuguaglianza di Schwarz si deduce facilmente la disuguaglianza triangolare. Infatti, $\|v + w\|$ e $\|v\| + \|w\|$ sono due numeri reali non-negativi e

$$\|v + w\|^2 = (v + w) \cdot (v + w) = v \cdot v + 2v \cdot w + w \cdot w \leq$$
$$\leq v \cdot v + 2|v \cdot w| + w \cdot w \leq \|v\|^2 + 2\|v\|\|w\| + \|w\|^2 = (\|v\| + \|w\|)^2$$

ove la disuguaglianza nella seconda riga è conseguenza della disuguaglianza di Schwarz. □

Dalla disuguaglianza di Schwarz si deduce facilmente la disuguaglianza triangolare. Infatti, $\|v+w\|$ e $\|v\|+\|w\|$ sono due numeri reali non-negativi e

$$\|v+w\|^2 = (v+w)\cdot(v+w) = v\cdot v + 2v\cdot w + w\cdot w \leq$$
$$\leq v\cdot v + 2|v\cdot w| + w\cdot w \leq \|v\|^2 + 2\|v\|\|w\| + \|w\|^2 = (\|v\|+\|w\|)^2$$

ove la disuguaglianza nella seconda riga è conseguenza della disuguaglianza di Schwarz. □

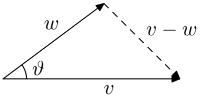
Definizione [angolo]

Si definisce l'*angolo non-orientato* tra due vettori non nulli v e w, come il numero reale $\vartheta \in [0,\pi]$, per cui si abbia $\frac{v\cdot w}{\|v\|\,\|w\|} = \cos\vartheta$.

Essendo i due vettori non-nulli, dalla disuguaglianza di Cauchy-Schwarz discende che $-1 \leq \frac{v\cdot w}{\|v\|\,\|w\|} \leq 1$ e quindi la definizione è ben posta (e l'angolo dipende dalle semirette generate dai due vettori, più che dai vettori stessi).

Inoltre, dalle proprietà del prodotto scalare discende

$$\|v-w\|^2 = \|v\|^2 + \|w\|^2 - 2v\cdot w = \|v\|^2 + \|w\|^2 - 2\|v\|\|w\|\cos\vartheta;$$



e quindi la definizione di angolo è coerente col cosiddetto "Teorema dei Coseni".

Osserviamo che, in base alla definizione di angolo, due vettori non nulli v e w in $\mathbb{R}^n$ formano un angolo di $\frac{\pi}{2}$ se, e solo se, $v \cdot w = 0$. Generalizziamo il fatto nella seguente

Definizione [ortogonale]

Due vettori v e w di $\mathbb{R}^n$ sono *ortogonali* se, e solo se, $v \cdot w = 0$. Dato un sottoinsieme, $S \neq \varnothing$, di $\mathbb{R}^n$, indicheremo con

$$S^\perp = \{\, v \in \mathbb{R}^n \mid v \cdot s = 0,\ \forall s \in S \,\}$$

il *sottospazio ortogonale* a S; ovvero il sottospazio dei vettori ortogonali ad ogni elementi di S.

Il sottospazio definito sopra è l'immagine del sottospazio ortogonale contenuto nello spazio duale tramite l'identificazione tra $\mathbb{R}^n$ e il suo duale indotta dal prodotto scalare. Valgono quindi le seguenti affermazioni

- Se S è un sottoinsieme di $\mathbb{R}^n$, allora $S^\perp$ è un sottospazio di $\mathbb{R}^n$ e $S_1 \subseteq S_2 \;\Rightarrow\; S_2^\perp \subseteq S_1^\perp$.
- Se S è un sottoinsieme di $\mathbb{R}^n$, allora $S^\perp = \langle S \rangle^\perp$.
- Se W è un sottospazio di $\mathbb{R}^n$ di dimensione k, allora $\dim W^\perp = n - k$.
- Se W è un sottospazio di $\mathbb{R}^n$, allora $(W^\perp)^\perp = W$.
- Dati due sottospazi, U e W di $\mathbb{R}^n$, si ha $(U+W)^\perp = U^\perp \cap W^\perp$ e $(U \cap W)^\perp = U^\perp + W^\perp$.

Basi ortogonali

Definizione [base ortogonale]

Una *base ortogonale* di $\mathbb{R}^n$ è una base $w_1, \dots, w_n$ di $\mathbb{R}^n$ costituita da vettori, a due a due ortogonali; ovvero tali che $w_i \cdot w_j = 0$ se $i \neq j$. Se, inoltre, tutti i vettori della base hanno norma 1, si dirà una *base ortonormale*.

A partire da una base qualunque di $\mathbb{R}^n$ si può ottenere una base ortonormale dello spazio; andiamo a descrivere un procedimento costruttivo per fare questo.

Procedimento di ortonormalizzazione di Gram-Schmidt : Vogliamo mostrare che, a partire da k vettori, $v_1, \dots, v_k$, linearmente indipendenti in $\mathbb{R}^n$, è possibile costruire k vettori ortonormali, $w_1, \dots, w_k$, tali che $\langle v_1, \dots, v_i \rangle = \langle w_1, \dots, w_i \rangle$, per ogni $i = 1, \dots, k$.

- Dovendo aversi $\langle v_1 \rangle = \langle w_1 \rangle$ e $\|w_1\| = 1$, basta porre $w_1 = \frac{v_1}{\|v_1\|} = \frac{1}{\|v_1\|} v_1$.
- Il vettore successivo, sarà un multiplo di $v_2 + a w_1$, determinato dalla condizione $(v_2 + a w_1) \cdot w_1 = 0$. Si porrà quindi $w_2' = v_2 - (v_2 \cdot w_1) w_1$ e $w_2 = \frac{w_2'}{\|w_2'\|}$. Per costruzione si ha $\langle v_1, v_2 \rangle = \langle w_1, w_2 \rangle$.
- Supponiamo ora di aver costruito i vettori ortonormali, $w_1, \dots, w_{i-1}$, tali che $\langle v_1, \dots, v_{i-1} \rangle = \langle w_1, \dots, w_{i-1} \rangle$ e cerchiamo un vettore $w_i' = v_i + a_1 w_1 + \cdots + a_{i-1} w_{i-1}$, tale che $w_i' \cdot w_1 = \cdots = w_i' \cdot w_{i-1} = 0$. Deve essere $a_1 = -v_i \cdot w_1, \dots, a_{i-1} = -v_i \cdot w_{i-1}$ e, si pone $w_i = \frac{w_i'}{\|w_i'\|}$. I vettori $w_1, \dots, w_i$ sono un sistema ortonormale e si ha $\langle v_1, \dots, v_i \rangle = \langle w_1, \dots, w_i \rangle$.
- Si procede in questo modo fino all'esaurimento dei vettori $v_1, \dots, v_k$.

Esercizio: Nello spazio $\mathbb{R}^4$ con l'usuale prodotto scalare siano dati i vettori $v_1 = \begin{pmatrix} 1 \\ 0 \\ 2 \\ -2 \end{pmatrix}$, $v_2 = \begin{pmatrix} 1 \\ 0 \\ 4 \\ 0 \end{pmatrix}$, $v_3 = \begin{pmatrix} 1 \\ -1 \\ 6 \\ 2 \end{pmatrix}$. Tramite il procedimento di Gram-Schmidt si determini una base ortogonale, $\mathcal{W} = \{w_1, w_2, w_3\}$, del sottospazio $V = \langle v_1, v_2, v_3 \rangle$ e si scriva la matrice di cambiamento di base $\alpha_{\mathcal{W},\mathcal{V}}(\mathrm{id}_V)$.

Svolg.: Ci limitiamo a determinare una base ortogonale senza normalizzare i vettori, per cui $w_1 = v_1$ e

$$w_2 = v_2 - \frac{v_2 \cdot w_1}{w_1 \cdot w_1} w_1 = \begin{pmatrix} 1 \\ 0 \\ 4 \\ 0 \end{pmatrix} - \frac{9}{9} \begin{pmatrix} 1 \\ 0 \\ 2 \\ -2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 2 \\ 2 \end{pmatrix},$$

$$w_3 = v_3 - \frac{v_3 \cdot w_1}{w_1 \cdot w_1} w_1 - \frac{v_3 \cdot w_2}{w_2 \cdot w_2} w_2 = \begin{pmatrix} 1 \\ -1 \\ 6 \\ 2 \end{pmatrix} - \frac{9}{9} \begin{pmatrix} 1 \\ 0 \\ 2 \\ -2 \end{pmatrix} - \frac{16}{8} \begin{pmatrix} 0 \\ 0 \\ 2 \\ 2 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \\ 0 \\ 0 \end{pmatrix}.$$

La matrice di cambiamento di base è $\alpha_{\mathcal{W},\mathcal{V}}(\mathrm{id}_V) = \begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix}$, non a caso, triangolare superiore. □

Esercizio : Nello spazio $\mathbb{R}^4$ con l'usuale prodotto scalare siano dati i vettori $v_1 = \begin{pmatrix} 1 \\ 0 \\ 2 \\ -2 \end{pmatrix}$, $v_2 = \begin{pmatrix} 1 \\ 0 \\ 4 \\ 0 \end{pmatrix}$, $v_3 = \begin{pmatrix} 1 \\ -1 \\ 6 \\ 2 \end{pmatrix}$. Tramite il procedimento di Gram-Schmidt si determini una base ortogonale, $\mathcal{W} = \{w_1, w_2, w_3\}$, del sottospazio $V = \langle v_1, v_2, v_3 \rangle$ e si scriva la matrice di cambiamento di base $\alpha_{\mathcal{W},\mathcal{V}}(\mathrm{id}_V)$.

Svolg.: Ci limitiamo a determinare una base ortogonale senza normalizzare i vettori, per cui $w_1 = v_1$ e

$$w_2 = v_2 - \frac{v_2 \cdot w_1}{w_1 \cdot w_1} w_1 = \begin{pmatrix} 1 \\ 0 \\ 4 \\ 0 \end{pmatrix} - \frac{9}{9} \begin{pmatrix} 1 \\ 0 \\ 2 \\ -2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 2 \\ 2 \end{pmatrix},$$

$$w_3 = v_3 - \frac{v_3 \cdot w_1}{w_1 \cdot w_1} w_1 - \frac{v_3 \cdot w_2}{w_2 \cdot w_2} w_2 = \begin{pmatrix} 1 \\ -1 \\ 6 \\ 2 \end{pmatrix} - \frac{9}{9} \begin{pmatrix} 1 \\ 0 \\ 2 \\ -2 \end{pmatrix} - \frac{16}{8} \begin{pmatrix} 0 \\ 0 \\ 2 \\ 2 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \\ 0 \\ 0 \end{pmatrix}.$$

La matrice di cambiamento di base è $\alpha_{\mathcal{W},\mathcal{V}}(\mathrm{id}_V) = \begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix}$, non a caso, triangolare superiore. $\square$

Esercizio : Sia $\phi : \mathbb{R}^n \to \mathbb{R}^n$ un endomorfismo triangolarizzabile. Si mostri che esiste una base ortonormale $\mathcal{W} = \{w_1, \dots, w_n\}$ di $\mathbb{R}^n$ tale che la matrice $\alpha_{\mathcal{W},\mathcal{W}}(\phi)$ sia triangolare superiore (Teorema di Schur reale).

Svolg.: Sappiamo che ϕ è triangolarizzabile se, e solo se, ha tutti gli autovalori reali. In tal caso, esiste una base (non necessariamente ortonormale) $\mathcal{V} = \{v_1, \dots, v_n\}$, tale che la matrice $\alpha_{\mathcal{V},\mathcal{V}}(\phi)$ sia triangolare superiore; ovvero si abbia $\phi(v_j) \in \langle v_1, \dots, v_j \rangle$, per $j = 1, \dots, n$. Applicando a questa base il procedimento di Gram-Schmidt, si ottiene una base ortonormale $\mathcal{W} = \{w_1, \dots, w_n\}$ tale che $\langle w_1, \dots, w_j \rangle = \langle v_1, \dots, v_j \rangle$, per $j = 1, \dots, n$. Dunque, si ha anche

$$\phi(w_j) \in \langle \phi(v_1), \dots, \phi(v_j) \rangle \subseteq \langle v_1, \dots, v_j \rangle = \langle w_1, \dots, w_j \rangle, \quad \text{per } j = 1, \dots, n;$$

ovvero $\alpha_{\mathcal{W},\mathcal{W}}(\phi)$ è triangolare superiore. $\square$

Proposizione

Si consideri lo spazio $\mathbb{R}^n$ dotato del prodotto scalare.

(a) Dati i vettori, $w_1, \dots, w_k$, non nulli e a due a due ortogonali, allora essi sono linearmente indipendenti.

(b) Per ogni sottospazio vettoriale W, si ha $\mathbb{R}^n = W \oplus W^\perp$.

(c) Se $\{w_1, \dots, w_k\}$ è una base ortogonale di W allora, fissato comunque un vettore v in $\mathbb{R}^n$, si ha $p_W(v) = \frac{w_1 \cdot v}{w_1 \cdot w_1} w_1 + \cdots + \frac{w_k \cdot v}{w_k \cdot w_k} w_k$, ove $p_W : \mathbb{R}^n \to \mathbb{R}^n$ è la *proiezione ortogonale* su W, ovvero la proiezione su W parallelamente a $W^\perp$.

(d) Per ogni vettore $w \in W$, si ha $\|v - p_W(v)\| \leq \|v - w\|$.

Proposizione

Si consideri lo spazio $\mathbb{R}^n$ dotato del prodotto scalare.

(a) Dati i vettori, $w_1, \dots, w_k$, non nulli e a due a due ortogonali, allora essi sono linearmente indipendenti.

(b) Per ogni sottospazio vettoriale W, si ha $\mathbb{R}^n = W \oplus W^\perp$.

(c) Se $\{w_1, \dots, w_k\}$ è una base ortogonale di W allora, fissato comunque un vettore v in $\mathbb{R}^n$, si ha $p_W(v) = \frac{w_1 \cdot v}{w_1 \cdot w_1} w_1 + \cdots + \frac{w_k \cdot v}{w_k \cdot w_k} w_k$, ove $p_W : \mathbb{R}^n \to \mathbb{R}^n$ è la *proiezione ortogonale* su W, ovvero la proiezione su W parallelamente a $W^\perp$.

(d) Per ogni vettore $w \in W$, si ha $\|v - p_W(v)\| \leq \|v - w\|$.

dim. (*a*) Da $a_1 w_1 + \cdots + a_k w_k = 0$ si deduce che, per ogni $i = 1, \dots, k$, si ha $0 = w_i \cdot 0 = w_i \cdot (a_1 w_1 + \cdots + a_k w_k) = a_i(w_i \cdot w_i)$. Essendo $w_i \cdot w_i \geq 0$, si conclude che tutti i coefficienti a_i sono nulli.
(*b*) Se $w \in W \cap W^\perp$, allora $w \cdot w = 0$ e quindi $w = 0$ perché il prodotto scalare è definito positivo. Quindi $W \cap W^\perp = \langle 0 \rangle$. Poiché $\dim W^\perp = n - k$, deve aversi $\mathbb{R}^n = W \oplus W^\perp$.
(*c*) Il vettore $p_W(v)$ dell'enunciato è ben definito e appartiene a W. Inoltre, per ogni vettore w_i della base $\mathcal{W} = \{w_1, \dots, w_k\}$ di W, si ha $w_i \cdot (v - p_W(v)) = w_i \cdot v - w_i \cdot w_i \frac{w_i \cdot v}{w_i \cdot w_i} = 0$. Dunque $v - p_W(v) \in W^\perp$ come si voleva.
(*d*) Per ogni vettore $w \in W$, si ha $w = p_W(v) + w'$, con $w' \in W$. Allora $\|v - w\|^2 = \|v - p_W(v) - w'\|^2 = \|v - p_W(v)\|^2 + \|w'\|^2$, perché i due addendi sono ortogonali. □

Esercizio : Si consideri il sottospazio $W = \langle 2e_1 - 3e_2 + e_4 \rangle \subset \mathbb{R}^4$. Si determinino le matrici in base canonica della proiezione ortogonale su W e della simmetria rispetto all'iperpiano $W^\perp$.

Svolg.: Posto $w = 2e_1 - 3e_2 + e_4$, si ha $p_W(v) = \frac{w \cdot v}{w \cdot w} w$, per ogni $v \in \mathbb{R}^n$. Dunque le matrici sono

$$A = \alpha_{\mathcal{E},\mathcal{E}}(p_W) = \frac{1}{14}\begin{pmatrix} 2 \\ -3 \\ 0 \\ 1 \end{pmatrix}(2, -3, 0, 1) = \begin{pmatrix} 2/7 & -3/7 & 0 & 1/7 \\ -3/7 & 9/14 & 0 & -3/14 \\ 0 & 0 & 0 & 0 \\ 1/7 & -3/14 & 0 & 1/14 \end{pmatrix},$$

$$B = \alpha_{\mathcal{E},\mathcal{E}}(s_{W^\perp}) = \mathbf{1}_4 - 2A = \begin{pmatrix} 3/7 & 6/7 & 0 & -2/7 \\ 6/7 & -2/7 & 0 & 3/7 \\ 0 & 0 & 1 & 0 \\ -2/7 & 3/7 & 0 & 6/7 \end{pmatrix}.$$

Ciò è quanto richiesto. □

Esercizio : Si consideri il sottospazio $W = \langle 2e_1 - 3e_2 + e_4 \rangle \subset \mathbb{R}^4$. Si determinino le matrici in base canonica della proiezione ortogonale su W e della simmetria rispetto all'iperpiano $W^\perp$.

Svolg.: Posto $w = 2e_1 - 3e_2 + e_4$, si ha $p_W(v) = \frac{w \cdot v}{w \cdot w} w$, per ogni $v \in \mathbb{R}^n$. Dunque le matrici sono

$$A = \alpha_{\mathcal{E},\mathcal{E}}(p_W) = \frac{1}{14} \begin{pmatrix} 2 \\ -3 \\ 0 \\ 1 \end{pmatrix} (2, -3, 0, 1) = \begin{pmatrix} 2/7 & -3/7 & 0 & 1/7 \\ -3/7 & 9/14 & 0 & -3/14 \\ 0 & 0 & 0 & 0 \\ 1/7 & -3/14 & 0 & 1/14 \end{pmatrix},$$

$$B = \alpha_{\mathcal{E},\mathcal{E}}(s_{W^\perp}) = \mathbf{1}_4 - 2A = \begin{pmatrix} 3/7 & 6/7 & 0 & -2/7 \\ 6/7 & -2/7 & 0 & 3/7 \\ 0 & 0 & 1 & 0 \\ -2/7 & 3/7 & 0 & 6/7 \end{pmatrix}.$$

Ciò è quanto richiesto. □

Esercizio : Si consideri il sottospazio $W = \langle e_1 - e_3, 4e_2 - 3e_4 \rangle \subset \mathbb{R}^4$. Si determinino le matrici in base canonica della proiezione e della simmetria ortogonale rispetto a W.

Svolg.: I due generatori sono una base ortogonale di W. Per il punto (c) della proposizione precedente, si ha

$$P = \alpha_{\mathcal{E},\mathcal{E}}(p_W) = \begin{pmatrix} 1 & 0 \\ 0 & 4 \\ -1 & 0 \\ 0 & -3 \end{pmatrix} \begin{pmatrix} 1/2 & 0 \\ 0 & 1/25 \end{pmatrix} \begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 4 & 0 & -3 \end{pmatrix} = \begin{pmatrix} 1/2 & 0 & -1/2 & 0 \\ 0 & 16/25 & 0 & -12/25 \\ -1/2 & 0 & 1/2 & 0 \\ 0 & -12/25 & 0 & 9/25 \end{pmatrix}.$$

Come per ogni proiezione e simmetria, si ha $S = \alpha_{\mathcal{E},\mathcal{E}}(s_W) = 2P - \mathbf{1}_4 = \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 7/25 & 0 & -24/25 \\ -1 & 0 & 0 & 0 \\ 0 & -24/25 & 0 & 7/25 \end{pmatrix}$. □

In base alle definizioni date, la base canonica, $\mathcal{E} = \{e_1, \dots, e_n\}$, di $\mathbb{R}^n$ è una base ortonormale.
Una base ortonormale, $\mathcal{V} = \{v_1, \dots, v_n\}$, coincide con l'immagine della base duale nell'identificazione indotta dal prodotto scalare; dunque si possono calcolare le coordinate dei vettori tramite il prodotto scalare. Si ha

$$\boxed{v = \sum_{i=1}^{n} (v_i \cdot v) v_i, \quad \text{per ogni } v \in \mathbb{R}^n \qquad \text{[identità di Parseval]}.}$$

Se non si vuole usare l'identificazione col duale, si può lo stesso verificare l'uguaglianza dei due membri osservando che la differenza tra v e la somma è ortogonale a tutti i vettori della base $\mathcal{V}$ e perciò è nulla.

In base alle definizioni date, la base canonica, $\mathcal{E} = \{e_1, \dots, e_n\}$, di $\mathbb{R}^n$ è una base ortonormale.
Una base ortonormale, $\mathcal{V} = \{v_1, \dots, v_n\}$, coincide con l'immagine della base duale nell'identificazione indotta dal prodotto scalare; dunque si possono calcolare le coordinate dei vettori tramite il prodotto scalare. Si ha

$$\boxed{v = \sum_{i=1}^{n} (v_i \cdot v) v_i, \quad \text{per ogni } v \in \mathbb{R}^n \qquad \text{[identità di Parseval]}.}$$

Se non si vuole usare l'identificazione col duale, si può lo stesso verificare l'uguaglianza dei due membri osservando che la differenza tra v e la somma è ortogonale a tutti i vettori della base $\mathcal{V}$ e perciò è nulla.

Inoltre, dati due vettori $v = a_1 v_1 + \cdots + a_n v_n$ e $w = b_1 v_1 + \cdots + b_n v_n$, si ha

$$v \cdot w = \sum_{1 \le i,j \le n} a_i b_j (v_i \cdot v_j) = \sum_{i=1}^{n} a_i b_i,$$

sempre perché la base $\mathcal{V} = \{v_1, \dots, v_n\}$ è ortonormale e $v_i \cdot v_j = \delta_{ij}$ (simbolo di Kronecker).

In base alle definizioni date, la base canonica, $\mathcal{E} = \{e_1, \ldots, e_n\}$, di $\mathbb{R}^n$ è una base ortonormale.
Una base ortonormale, $\mathcal{V} = \{v_1, \ldots, v_n\}$, coincide con l'immagine della base duale nell'identificazione indotta dal prodotto scalare; dunque si possono calcolare le coordinate dei vettori tramite il prodotto scalare. Si ha

$$v = \sum_{i=1}^{n} (v_i \cdot v) v_i, \quad \text{per ogni } v \in \mathbb{R}^n \qquad \text{[identità di Parseval]}.$$

Se non si vuole usare l'identificazione col duale, si può lo stesso verificare l'uguaglianza dei due membri osservando che la differenza tra v e la somma è ortogonale a tutti i vettori della base $\mathcal{V}$ e perciò è nulla.

Inoltre, dati due vettori $v = a_1 v_1 + \cdots + a_n v_n$ e $w = b_1 v_1 + \cdots + b_n v_n$, si ha

$$v \cdot w = \sum_{1 \leq i,j \leq n} a_i b_j (v_i \cdot v_j) = \sum_{i=1}^{n} a_i b_i,$$

sempre perché la base $\mathcal{V} = \{v_1, \ldots, v_n\}$ è ortonormale e $v_i \cdot v_j = \delta_{ij}$ (simbolo di Kronecker).

Se $\mathcal{V} = \{v_1, \ldots, v_n\}$ e $\mathcal{W} = \{w_1, \ldots, w_n\}$ sono due basi ortonormali di $\mathbb{R}^n$ e $A = \alpha_{\mathcal{W},\mathcal{V}}(\mathrm{id}_{\mathbb{R}^n})$ è la matrice di cambiamento di base, allora deve aversi ${}^tAA = \mathbf{1}_n$ (A è una matrice ortogonale) e quindi $\det A = \pm 1$.

Infatti, per quanto osservato sopra sul calcolo del prodotto scalare tramite coordinate relative ad una base ortonormale, l'entrata di posto (i, j) della matrice ${}^tAA$ coincide col valore di $w_i \cdot w_j = \delta_{ij}$. Di conseguenza $1 = \det \mathbf{1} = \det({}^tAA) = (\det A)^2$, da cui si deduce la condizione sul determinante.

Isometrie

Definizione [isometria]

Un'applicazione lineare $\phi : \mathbb{R}^n \to \mathbb{R}^n$ è un'*isometria* (per la metrica euclidea) se $\phi(v) \cdot \phi(w) = v \cdot w$ per ogni coppia di vettori v e w in $\mathbb{R}^n$.

Se $\phi : \mathbb{R}^n \to \mathbb{R}^n$ è un'isometria e $\mathcal{V} = \{v_1, \dots, v_n\}$ una base ortonormale allora, per ogni coppia dei vettori della base si ha $\phi(v_i) \cdot \phi(v_j) = v_i \cdot v_j = \delta_{ij}$. Ciò significa precisamente che $A = \alpha_{\mathcal{V},\mathcal{V}}(\phi)$ è una matrice ortogonale, ovvero soddisfa alla condizione ${}^tAA = \mathbf{1}$. Per questo le isometrie sono anche dette *trasformazioni ortogonali*. Si noti che, $\|\phi(v)\| = \|v\|$ per ogni $v \in \mathbb{R}^n$, per cui gli unici autovalori reali possibili per un'isometria sono ± 1.

Isometrie

Definizione [isometria]

Un'applicazione lineare $\phi : \mathbb{R}^n \to \mathbb{R}^n$ è un'*isometria* (per la metrica euclidea) se $\phi(v) \cdot \phi(w) = v \cdot w$ per ogni coppia di vettori v e w in $\mathbb{R}^n$.

Se $\phi : \mathbb{R}^n \to \mathbb{R}^n$ è un'isometria e $\mathcal{V} = \{v_1, \dots, v_n\}$ una base ortonormale allora, per ogni coppia dei vettori della base si ha $\phi(v_i) \cdot \phi(v_j) = v_i \cdot v_j = \delta_{ij}$. Ciò significa precisamente che $A = \alpha_{\mathcal{V},\mathcal{V}}(\phi)$ è una matrice ortogonale, ovvero soddisfa alla condizione ${}^tAA = \mathbf{1}$. Per questo le isometrie sono anche dette *trasformazioni ortogonali*. Si noti che, $\|\phi(v)\| = \|v\|$ per ogni $v \in \mathbb{R}^n$, per cui gli unici autovalori reali possibili per un'isometria sono ± 1.

Le isometrie di $\mathbb{R}^n$ formano un gruppo che si indica col simbolo O_n. Per quanto osservato sulle matrici ortogonali, un'isometria ha determinante ± 1 e le isometrie di determinante 1 formano un sottogruppo di indice 2 di O_n, che si denota con SO_n (gruppo ortogonale speciale, di ordine n).

N.B. Così come un elemento del gruppo $\mathrm{GL}(n, K)$ delle matrici invertibili può essere pensato sia come la matrice di un'applicazione lineare invertibile che come la matrice di un cambiamento di base in K^n, allo stesso modo una matrice ortogonale, elemento del gruppo $O_n(\mathbb{R}) \subset \mathrm{GL}(n, \mathbb{R})$, può essere pensata sia come la matrice di un'isometria di $\mathbb{R}^n$ (rispetto a una base ortonormale) che come la matrice di un cambiamento di base tra basi ortonormali di $\mathbb{R}^n$.

Trasformazioni ortogonali di $\mathbb{R}^2$: Sia $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ una matrice reale quadrata di ordine 2 tale che ${}^tAA = \mathbf{1}_2$. Calcolando il prodotto, si trovano le condizioni $a^2 + c^2 = 1 = b^2 + d^2$ e $ab + cd = 0$; da cui si deduce:

(a) se $\det A = 1$ si ha matrice $A = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$, $(0 \leq \alpha < 2\pi)$, ovvero la rotazione antioraria di angolo α, che ha autovalori reali solo per $\alpha \in \{0, \pi\}$;

(b) se $\det A = -1$ si ha matrice $A = \begin{pmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix}$, $(0 \leq \alpha < 2\pi)$, ovvero la simmetria ortogonale di asse la retta per l'origine che forma un angolo di $\alpha/2$ con l'asse $\langle e_1 \rangle$ (ossia la retta di equazione cartesiana $\sin(\alpha/2)x - \cos(\alpha/2)y = 0$). Ha ± 1 come autovalori e gli autovettori di autovalore 1 sono paralleli all'asse di simmetria mentre gli autovettori di autovalore -1 sono ortogonali all'asse.

Trasformazioni ortogonali di $\mathbb{R}^2$: Sia $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ una matrice reale quadrata di ordine 2 tale che ${}^tAA = \mathbf{1}_2$. Calcolando il prodotto, si trovano le condizioni $a^2 + c^2 = 1 = b^2 + d^2$ e $ab + cd = 0$; da cui si deduce:

(a) se $\det A = 1$ si ha matrice $A = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$, $(0 \le \alpha < 2\pi)$, ovvero la rotazione antioraria di angolo α, che ha autovalori reali solo per $\alpha \in \{0, \pi\}$;

(b) se $\det A = -1$ si ha matrice $A = \begin{pmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix}$, $(0 \le \alpha < 2\pi)$, ovvero la simmetria ortogonale di asse la retta per l'origine che forma un angolo di $\alpha/2$ con l'asse $\langle e_1 \rangle$ (ossia la retta di equazione cartesiana $\sin(\alpha/2)x - \cos(\alpha/2)y = 0$). Ha ± 1 come autovalori e gli autovettori di autovalore 1 sono paralleli all'asse di simmetria mentre gli autovettori di autovalore -1 sono ortogonali all'asse.

Trasformazioni ortogonali di $\mathbb{R}^3$: Sia $\phi\colon \mathbb{R}^3 \to \mathbb{R}^3$ una trasformazione ortogonale. Il polinomio caratteristico di ϕ ha coefficienti reali e grado 3, pertanto ha almeno uno zero reale, $\lambda = \pm 1$. Se v è autovettore relativo a λ, $\mathbb{R}^3 = \langle v \rangle \oplus \langle v \rangle^\perp$ e ϕ induce un'isometria nel sottospazio $\langle v \rangle^\perp$. Si dimostra che si può scegliere $\lambda = \det\phi$ e quindi che $\phi_{|\langle v \rangle^\perp}$ è una rotazione nel piano $\langle v \rangle^\perp$. Dunque,

(a) se $\det\phi = 1$ esiste una base ortonormale $\mathcal{V} = \{v_1, v_2, v_3\}$ tale che $\alpha_{\mathcal{V},\mathcal{V}}(\phi) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{pmatrix}$ e ϕ è la rotazione di asse $\langle v_1 \rangle$ e angolo $\alpha \in [0, \pi]$ (**n.b.** $\operatorname{tr}\phi = 1 + 2\cos\alpha$);

(b) se $\det\phi = -1$ esiste una base ortonormale $\mathcal{V} = \{v_1, v_2, v_3\}$ tale che $\alpha_{\mathcal{V},\mathcal{V}}(\phi) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{pmatrix}$ e ϕ è una rotoriflessione, ovvero la rotazione di angolo α attorno al sottospazio $\langle v_1 \rangle$, seguita dalla simmetria ortogonale che ha come asse il piano di rotazione $\langle v_1 \rangle^\perp$. La trasformazione è una simmetria (rispetto a un piano o rispetto a un punto) se $\alpha \in \{0, \pi\}$.

Esercizio : Si consideri la matrice $A = \begin{pmatrix} 1/3 & 2/3 & 2/3 \\ 2/3 & 1/3 & -2/3 \\ -2/3 & 2/3 & -1/3 \end{pmatrix}$.
Si verifichi che si tratta di una matrice ortogonale e si classifichi la trasformazione ortogonale corrispondente. Si determini la matrice canonica e una base ortonormale rispetto a cui la trasformazione ha matrice canonica.

Svolg.: Si ha ${}^tAA = \mathbf{1}_3$ e $\det A = 1$. Dunque si tratta di una rotazione di angolo ϑ, ove
$\cos\vartheta = \frac{\mathrm{tr}\,A-1}{2} = -\frac{1}{3}$.
$A - \mathbf{1}_3 = \begin{pmatrix} -2/3 & 2/3 & 2/3 \\ 2/3 & -2/3 & -2/3 \\ -2/3 & 2/3 & -4/3 \end{pmatrix}$ e l'asse di rotazione è generato da $v_1 = \frac{1}{\sqrt{2}}(e_1 + e_2)$.
Il piano di rotazione è $\langle v_1 \rangle^\perp$ e ha quindi equazione cartesiana $x + y = 0$. Una base ortonormale di questo sottospazio è data dai vettori $v_2 = \frac{1}{\sqrt{2}}(e_1 - e_2)$ e $v_3 = e_3$. Rispetto a questa base la rotazione indotta ha matrice $\begin{pmatrix} -1/3 & 4/3\sqrt{2} \\ -4/3\sqrt{2} & -1/3 \end{pmatrix}$ da cui si ricava la forma canonica della matrice di rotazione.
Si noti che il seno dell'angolo di rotazione, $4/3\sqrt{2}$, cambia di segno se si cambia il verso di uno dei generatori di $\langle v_1 \rangle^\perp$ o se se ne cambia l'ordine (ovvero se si cambia l'orientamento del sottospazio). $\square$

Esercizio : Si consideri la matrice $A = \begin{pmatrix} 1/3 & 2/3 & 2/3 \\ 2/3 & 1/3 & -2/3 \\ -2/3 & 2/3 & -1/3 \end{pmatrix}$.
Si verifichi che si tratta di una matrice ortogonale e si classifichi la trasformazione ortogonale corrispondente. Si determini la matrice canonica e una base ortonormale rispetto a cui la trasformazione ha matrice canonica.

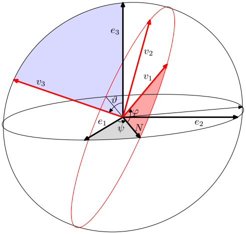
Svolg.: Si ha ${}^tAA = \mathbf{1}_3$ e $\det A = 1$. Dunque si tratta di una rotazione di angolo ϑ, ove $\cos\vartheta = \frac{\operatorname{tr} A - 1}{2} = -\frac{1}{3}$.
$A - \mathbf{1}_3 = \begin{pmatrix} -2/3 & 2/3 & 2/3 \\ 2/3 & -2/3 & -2/3 \\ -2/3 & 2/3 & -4/3 \end{pmatrix}$ e l'asse di rotazione è generato da $v_1 = \frac{1}{\sqrt{2}}(e_1 + e_2)$.
Il piano di rotazione è $\langle v_1 \rangle^\perp$ e ha quindi equazione cartesiana $x + y = 0$. Una base ortonormale di questo sottospazio è data dai vettori $v_2 = \frac{1}{\sqrt{2}}(e_1 - e_2)$ e $v_3 = e_3$. Rispetto a questa base la rotazione indotta ha matrice $\begin{pmatrix} -1/3 & 4/3\sqrt{2} \\ -4/3\sqrt{2} & -1/3 \end{pmatrix}$ da cui si ricava la forma canonica della matrice di rotazione.
Si noti che il seno dell'angolo di rotazione, $4/3\sqrt{2}$, cambia di segno se si cambia il verso di uno dei generatori di $\langle v_1 \rangle^\perp$ o se se ne cambia l'ordine (ovvero se si cambia l'orientamento del sottospazio). □

Esercizio : Si consideri la matrice $A = \begin{pmatrix} -2/7 & 3/7 & 6/7 \\ 3/7 & 6/7 & -2/7 \\ 6/7 & -2/7 & 3/7 \end{pmatrix}$.
Si verifichi che si tratta di una matrice ortogonale e si classifichi la trasformazione ortogonale corrispondente. Si determini la matrice canonica e una base ortonormale rispetto a cui la trasformazione ha matrice canonica.

Svolg.: Si ha ${}^tAA = \mathbf{1}_3$ e $\det A = -1$. Dunque si tratta di una rotoriflessione o di una simmetria. L'angolo di rotazione ϑ, soddisfa alla condizione $\cos\vartheta = \frac{\operatorname{tr} A + 1}{2} = 1$ e si tratta quindi di una simmetria rispetto a un piano.
$A + \mathbf{1}_3 = \begin{pmatrix} 5/7 & 3/7 & 6/7 \\ 3/7 & 13/7 & -2/7 \\ 6/7 & -2/7 & 10/7 \end{pmatrix}$ e il piano di riflessione è ortogonale all'autovettore $3e_1 - e_2 - 2e_3$, ovvero ha equazione cartesiana $3x - y - 2z = 0$. La matrice canonica è la matrice diagonale $\begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ e una base ortonormale di autovettori è $v_1 = \frac{1}{\dots}(3e_1 - e_2 - 2e_3)$, $v_2 = \frac{1}{\dots}(2e_2 - e_3)$

Angoli di Eulero: Ogni trasformazione ortogonale $\Phi \in SO_3(\mathbb{R})$ è individuata da 3 rotazioni. I tre angoli di rotazione sono gli *Angoli di Eulero* e la loro conoscenza permette di scrivere facilmente la matrice di Φ.

Siano $\Phi(e_1) = v_1$, $\Phi(e_2) = v_2$, $\Phi(e_3) = v_3$ e sia ϑ l'angolo che porta e_3 su v_3. Se $\vartheta = 0$ o $\vartheta = \pi$, allora Φ è una rotazione del piano $\langle e_1, e_2\rangle$, oppure è la composizione di una tale rotazione con un ribaltamento dell'asse e_3. Altrimenti $\langle e_1, e_2\rangle \cap \langle v_1, v_2\rangle$ è una retta detta la *linea dei nodi*. Sia N il versore di tale retta per cui $\vartheta \in [0, \pi]$ ($\langle N\rangle = \langle e_3, v_3\rangle^\perp$). Siano poi $\psi \in [0, 2\pi]$ l'ampiezza della rotazione del piano $\langle e_1, e_2\rangle$ che porta e_1 su N e $\varphi \in [0, 2\pi]$ l'ampiezza della rotazione del piano $\langle v_1, v_2\rangle$ che porta N su v_1. Allora possiamo portare i vettori della base e_1, e_2, e_3, ordinatamente sui vettori v_1, v_2, v_3 con le seguenti trasformazioni:



- rotazione di asse e_3 ed angolo ψ (precessione), seguita da
- rotazione di asse N ed angolo ϑ (nutazione), seguita da
- rotazione di asse v_3 ed angolo φ (rotazione propria).

Vediamo ora come sia possibile determinare la matrice $A = \alpha_{\mathcal{E},\mathcal{E}}(\Phi)$ a partire dai tre angoli di Eulero.

- Si ha $N = \cos\psi e_1 + \sin\psi e_2$, e si pone $N' = -\sin\psi e_1 + \cos\psi e_2$ in $\langle e_1, e_2\rangle \cap \langle N\rangle^\perp$ e $N'' = \cos\vartheta N' + \sin\vartheta e_3$ in $\langle v_1, v_2\rangle \cap \langle N\rangle^\perp$.
- Da ciò si ricava $v_1 = \cos\varphi N + \sin\varphi N''$, $v_2 = -\sin\varphi N + \cos\varphi N''$, $v_3 = -\sin\vartheta N' + \cos\vartheta e_3$.

Quindi $(v_1, v_2, v_3) = (e_1, e_2, e_3)A$, ove

$$A = \alpha_{\mathcal{E},\mathcal{E}}(\Phi) = \begin{pmatrix} \cos\psi & -\sin\psi & 0 \\ \sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\vartheta & -\sin\vartheta \\ 0 & \sin\vartheta & \cos\vartheta \end{pmatrix} \begin{pmatrix} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

A volte, si usa il simbolo $R_{(\psi,\vartheta,\varphi)}$ per indicare il prodotto scritto sopra.

Per ogni isometria $\phi : \mathbb{R}^n \to \mathbb{R}^n$, indichiamo con $F_\phi = \{\ v \in \mathbb{R}^n \mid \phi(v) = v\ \}$ il sottospazio dei vettori uniti.

Osservazione

Un'isometria non identica $\phi : \mathbb{R}^n \to \mathbb{R}^n$ che abbia un'iperpiano i vettori uniti è la simmetria ortogonale rispetto a un iperpiano (simmetria assiale).

Per ogni isometria $\phi : \mathbb{R}^n \to \mathbb{R}^n$, indichiamo con $F_\phi = \{\ v \in \mathbb{R}^n \mid \phi(v) = v\ \}$ il sottospazio dei vettori uniti.

Osservazione

Un'isometria non identica $\phi : \mathbb{R}^n \to \mathbb{R}^n$ che abbia un'iperpiano i vettori uniti è la simmetria ortogonale rispetto a un iperpiano (simmetria assiale).

dim. Sia $F = F_\phi$ il sottospazio dei vettori uniti. Poiché $\phi \neq \mathrm{id}$, F è un iperpiano e fissiamo un generatore v del sottospazio ortogonale, per cui $\mathbb{R}^n = F \oplus \langle v \rangle$. Per ogni $w \in F$, si ha $\phi(v) \cdot w = \phi(v) \cdot \phi(w) = v \cdot w = 0$, per cui $\phi(v) \in F^\perp = \langle v \rangle$, quindi v è un autovettore per ϕ e $\phi(v) = -v$. Dunque ϕ è la simmetria ortogonale di asse F, ovvero $\boxed{\phi(x) = x - 2\frac{v \cdot x}{v \cdot v} v, \text{ per ogni } x \in \mathbb{R}^n}$. □

Per ogni isometria $\phi : \mathbb{R}^n \to \mathbb{R}^n$, indichiamo con $F_\phi = \{ v \in \mathbb{R}^n \mid \phi(v) = v \}$ il sottospazio dei vettori uniti.

Osservazione

Un'isometria non identica $\phi : \mathbb{R}^n \to \mathbb{R}^n$ che abbia un'iperpiano i vettori uniti è la simmetria ortogonale rispetto a un iperpiano (simmetria assiale).

dim. Sia $F = F_\phi$ il sottospazio dei vettori uniti. Poiché $\phi \neq \mathrm{id}$, F è un iperpiano e fissiamo un generatore v del sottospazio ortogonale, per cui $\mathbb{R}^n = F \oplus \langle v \rangle$. Per ogni $w \in F$, si ha $\phi(v) \cdot w = \phi(v) \cdot \phi(w) = v \cdot w = 0$, per cui $\phi(v) \in F^\perp = \langle v \rangle$, quindi v è un autovettore per ϕ e $\phi(v) = -v$. Dunque ϕ è la simmetria ortogonale di asse F, ovvero $\boxed{\phi(x) = x - 2\frac{v \cdot x}{v \cdot v} v, \text{ per ogni } x \in \mathbb{R}^n}$. □

Proposizione

Ogni isometria non identica $\phi : \mathbb{R}^n \to \mathbb{R}^n$ è prodotto di al più n simmetrie assiali.

Per ogni isometria $\phi : \mathbb{R}^n \to \mathbb{R}^n$, indichiamo con $F_\phi = \{ v \in \mathbb{R}^n \mid \phi(v) = v \}$ il sottospazio dei vettori uniti.

Osservazione

Un'isometria non identica $\phi : \mathbb{R}^n \to \mathbb{R}^n$ che abbia un'iperpiano i vettori uniti è la simmetria ortogonale rispetto a un iperpiano (simmetria assiale).

dim. Sia $F = F_\phi$ il sottospazio dei vettori uniti. Poiché $\phi \neq \mathrm{id}$, F è un iperpiano e fissiamo un generatore v del sottospazio ortogonale, per cui $\mathbb{R}^n = F \oplus \langle v \rangle$. Per ogni $w \in F$, si ha $\phi(v) \cdot w = \phi(v) \cdot \phi(w) = v \cdot w = 0$, per cui $\phi(v) \in F^\perp = \langle v \rangle$, quindi v è un autovettore per ϕ e $\phi(v) = -v$. Dunque ϕ è la simmetria ortogonale di asse F, ovvero $\boxed{\phi(x) = x - 2\frac{v \cdot x}{v \cdot v} v, \text{ per ogni } x \in \mathbb{R}^n}$. □

Proposizione

Ogni isometria non identica $\phi : \mathbb{R}^n \to \mathbb{R}^n$ è prodotto di al più n simmetrie assiali.

dim. Sia $F = F_\phi$ il sottospazio dei vettori uniti. Se $\dim F \geq n - 1$ non c'è niente da dimostrare. Altrimenti sia $0 \neq u \in F^\perp$ e poniamo $v = u - \phi(u) \neq 0$. Per quanto osservato nella dimostrazione precedente, $F \subset \langle v \rangle^\perp$, e possiamo considerare la simmetria di asse $\langle v \rangle^\perp$, ovvero $\sigma : \mathbb{R}^n \to \mathbb{R}^n$, definita ponendo $\sigma(x) = x - 2\frac{v \cdot x}{v \cdot v} v$, per ogni $x \in \mathbb{R}^n$.

Per ogni isometria $\phi : \mathbb{R}^n \to \mathbb{R}^n$, indichiamo con $F_\phi = \{ v \in \mathbb{R}^n \mid \phi(v) = v \}$ il sottospazio dei vettori uniti.

Osservazione

Un'isometria non identica $\phi : \mathbb{R}^n \to \mathbb{R}^n$ che abbia un'iperpiano i vettori uniti è la simmetria ortogonale rispetto a un iperpiano (simmetria assiale).

dim. Sia $F = F_\phi$ il sottospazio dei vettori uniti. Poiché $\phi \neq \mathrm{id}$, F è un iperpiano e fissiamo un generatore v del sottospazio ortogonale, per cui $\mathbb{R}^n = F \oplus \langle v \rangle$. Per ogni $w \in F$, si ha $\phi(v) \cdot w = \phi(v) \cdot \phi(w) = v \cdot w = 0$, per cui $\phi(v) \in F^\perp = \langle v \rangle$, quindi v è un autovettore per ϕ e $\phi(v) = -v$. Dunque ϕ è la simmetria ortogonale di asse F, ovvero $\boxed{\phi(x) = x - 2\frac{v \cdot x}{v \cdot v} v, \text{ per ogni } x \in \mathbb{R}^n}$. □

Proposizione

Ogni isometria non identica $\phi : \mathbb{R}^n \to \mathbb{R}^n$ è prodotto di al più n simmetrie assiali.

dim. Sia $F = F_\phi$ il sottospazio dei vettori uniti. Se $\dim F \geq n - 1$ non c'è niente da dimostrare. Altrimenti sia $0 \neq u \in F^\perp$ e poniamo $v = u - \phi(u) \neq 0$. Per quanto osservato nella dimostrazione precedente, $F \subset \langle v \rangle^\perp$, e possiamo considerare la simmetria di asse $\langle v \rangle^\perp$, ovvero $\sigma : \mathbb{R}^n \to \mathbb{R}^n$, definita ponendo $\sigma(x) = x - 2\frac{v \cdot x}{v \cdot v} v$, per ogni $x \in \mathbb{R}^n$. Dimostriamo che il sottospazio dei vettori uniti di $\phi \circ \sigma$ contiene propriamente F. Infatti $\sigma(w) = w$ per ogni $w \in F$ e inoltre $\sigma(\phi(u)) = \phi(u) - (u - \phi(u)) = u$, perché $\frac{2(u-\phi(u))\cdot\phi(u)}{(u-\phi(u))\cdot(u-\phi(u))} = 1$, essendo $u \cdot u = \phi(u) \cdot \phi(u)$. Ne consegue che $F \oplus \langle u \rangle \subseteq F_{\phi\circ\sigma}$.

Per ogni isometria $\phi : \mathbb{R}^n \to \mathbb{R}^n$, indichiamo con $F_\phi = \{\ v \in \mathbb{R}^n \mid \phi(v) = v\ \}$ il sottospazio dei vettori uniti.

Osservazione

Un'isometria non identica $\phi : \mathbb{R}^n \to \mathbb{R}^n$ che abbia un'iperpiano i vettori uniti è la simmetria ortogonale rispetto a un iperpiano (simmetria assiale).

dim. Sia $F = F_\phi$ il sottospazio dei vettori uniti. Poiché $\phi \neq \mathrm{id}$, F è un iperpiano e fissiamo un generatore v del sottospazio ortogonale, per cui $\mathbb{R}^n = F \oplus \langle v \rangle$. Per ogni $w \in F$, si ha $\phi(v) \cdot w = \phi(v) \cdot \phi(w) = v \cdot w = 0$, per cui $\phi(v) \in F^\perp = \langle v \rangle$, quindi v è un autovettore per ϕ e $\phi(v) = -v$. Dunque ϕ è la simmetria ortogonale di asse F, ovvero $\boxed{\phi(x) = x - 2\frac{v \cdot x}{v \cdot v} v, \text{ per ogni } x \in \mathbb{R}^n}$. □

Proposizione

Ogni isometria non identica $\phi : \mathbb{R}^n \to \mathbb{R}^n$ è prodotto di al più n simmetrie assiali.

dim. Sia $F = F_\phi$ il sottospazio dei vettori uniti. Se $\dim F \geq n - 1$ non c'è niente da dimostrare. Altrimenti sia $0 \neq u \in F^\perp$ e poniamo $v = u - \phi(u) \neq 0$. Per quanto osservato nella dimostrazione precedente, $F \subset \langle v \rangle^\perp$, e possiamo considerare la simmetria di asse $\langle v \rangle^\perp$, ovvero $\sigma : \mathbb{R}^n \to \mathbb{R}^n$, definita ponendo $\sigma(x) = x - 2\frac{v \cdot x}{v \cdot v} v$, per ogni $x \in \mathbb{R}^n$. Dimostriamo che il sottospazio dei vettori uniti di $\phi \circ \sigma$ contiene propriamente F. Infatti $\sigma(w) = w$ per ogni $w \in F$ e inoltre $\sigma(\phi(u)) = \phi(u) - (u - \phi(u)) = u$, perché $\frac{2(u-\phi(u)) \cdot \phi(u)}{(u-\phi(u)) \cdot (u-\phi(u))} = 1$, essendo $u \cdot u = \phi(u) \cdot \phi(u)$. Ne consegue che $F \oplus \langle u \rangle \subseteq F_{\phi \circ \sigma}$. Allora, componendo ϕ con al più n simmetrie si ottiene un sottospazio di vettori uniti di dimensione n, ovvero l'dentità e perciò ϕ è uguale alla composizione delle stesse simmetrie nell'ordine inverso. □

Teorema Spettrale (per endomorfismi simmetrici)

Definizione [endomorfismo simmetrico]

Un endomorfismo $\phi : \mathbb{R}^n \to \mathbb{R}^n$ si dice *simmetrico* se $\phi(v) \cdot w = v \cdot \phi(w)$ per ogni coppia di vettori v, w in $\mathbb{R}^n$.

Teorema Spettrale (per endomorfismi simmetrici)

Definizione [endomorfismo simmetrico]

Un endomorfismo $\phi : \mathbb{R}^n \to \mathbb{R}^n$ si dice *simmetrico* se $\phi(v) \cdot w = v \cdot \phi(w)$ per ogni coppia di vettori v, w in $\mathbb{R}^n$.

Nell'identificazione tra $\mathbb{R}^n$ e il suo spazio duale data dal prodotto scalare, ciò significa che ϕ coincide con l'endomorfismo trasposto. Più semplicemente, se $\mathcal{V} = \{v_1, \dots, v_n\}$ ùna base ortonormale di $\mathbb{R}^n$, allora $\phi : \mathbb{R}^n \to \mathbb{R}^n$ è un endomorfismo simmetrico se, e solo se, la matrice $A = \alpha_{\mathcal{V},\mathcal{V}}(\phi)$ è simmetrica (${}^tA = A$). Infatti, basta osservare che l'elemento di posto (i, j) di A è $a_{ij} = v_i \cdot \phi(v_j)$.

Teorema Spettrale (per endomorfismi simmetrici)

Definizione [endomorfismo simmetrico]

Un endomorfismo $\phi : \mathbb{R}^n \to \mathbb{R}^n$ si dice *simmetrico* se $\phi(v) \cdot w = v \cdot \phi(w)$ per ogni coppia di vettori v, w in $\mathbb{R}^n$.

Nell'identificazione tra $\mathbb{R}^n$ e il suo spazio duale data dal prodotto scalare, ciò significa che ϕ coincide con l'endomorfismo trasposto. Più semplicemente, se $\mathcal{V} = \{v_1, \dots, v_n\}$ ùna base ortonormale di $\mathbb{R}^n$, allora $\phi : \mathbb{R}^n \to \mathbb{R}^n$ è un endomorfismo simmetrico se, e solo se, la matrice $A = \alpha_{\mathcal{V},\mathcal{V}}(\phi)$ è simmetrica (${}^tA = A$). Infatti, basta osservare che l'elemento di posto (i, j) di A è $a_{ij} = v_i \cdot \phi(v_j)$.

Definizione [diagonalizzabilità ortogonale]

Un endomorfismo $\phi : \mathbb{R}^n \to \mathbb{R}^n$ si dice *ortogonalmente diagonalizzabile* se ammette una base ortonormale di autovettori.
Una matrice $A \in M_n(\mathbb{R})$ è *ortogonalmente diagonalizzabile* se esiste una matrice ortogonale R tale che ${}^tRAR$ sia una matrice diagonale.

Teorema Spettrale (per endomorfismi simmetrici)

Definizione [endomorfismo simmetrico]

Un endomorfismo $\phi : \mathbb{R}^n \to \mathbb{R}^n$ si dice *simmetrico* se $\phi(v) \cdot w = v \cdot \phi(w)$ per ogni coppia di vettori v, w in $\mathbb{R}^n$.

Nell'identificazione tra $\mathbb{R}^n$ e il suo spazio duale data dal prodotto scalare, ciò significa che ϕ coincide con l'endomorfismo trasposto. Più semplicemente, se $\mathcal{V} = \{v_1, \dots, v_n\}$ ùna base ortonormale di $\mathbb{R}^n$, allora $\phi : \mathbb{R}^n \to \mathbb{R}^n$ è un endomorfismo simmetrico se, e solo se, la matrice $A = \alpha_{\mathcal{V},\mathcal{V}}(\phi)$ è simmetrica (${}^tA = A$). Infatti, basta osservare che l'elemento di posto (i, j) di A è $a_{ij} = v_i \cdot \phi(v_j)$.

Definizione [diagonalizzabilità ortogonale]

Un endomorfismo $\phi : \mathbb{R}^n \to \mathbb{R}^n$ si dice *ortogonalmente diagonalizzabile* se ammette una base ortonormale di autovettori.
Una matrice $A \in M_n(\mathbb{R})$ è *ortogonalmente diagonalizzabile* se esiste una matrice ortogonale R tale che ${}^tRAR$ sia una matrice diagonale.

Il collegamento tra le due definizioni è il contenuto del seguente

Teorema (spettrale per endomorfismi simmetrici)

Un endomorfismo $\phi : \mathbb{R}^n \to \mathbb{R}^n$ ammette una base ortonormale di autovettori se, e solo se, è simmetrico.

dim. Se ϕ ammette una base ortonormale di autovettori $\mathcal{V} = \{v_1, \dots, v_n\}$ la matrice $D = \alpha_{\mathcal{V},\mathcal{V}}(\phi)$ è diagonale e quindi simmetrica (${}^tD = D$) per cui ϕ è un endomorfismo simmetrico.

dim. Se ϕ ammette una base ortonormale di autovettori $\mathcal{V} = \{v_1, \dots, v_n\}$ la matrice $D = \alpha_{\mathcal{V},\mathcal{V}}(\phi)$ è diagonale e quindi simmetrica (${}^tD = D$) per cui ϕ è un endomorfismo simmetrico.
Viceversa, sia $\phi : \mathbb{R}^n \to \mathbb{R}^n$ un endomorfismo simmetrico. Allora

- gli autovalori di ϕ sono tutti reali. Sia $c \in \mathbb{C}$ un autovalore di ϕ (ovvero dell'endomorfismo di $\mathbb{C}^n$ che, in base canonica, ha la stessa matrice $A \in M_n(\mathbb{R})$ di ϕ) e sia $v \in \mathbb{C}^n$ un corrispondente autovettore. Allora
$$c\,{}^t\bar{v}v = {}^t\bar{v}(Av) = ({}^t\bar{v}A)v = {}^t(\bar{A}\bar{v})v = \bar{c}\,{}^t\bar{v}v,$$
da cui si deduce $c = \bar{c}$, essendo ${}^t\bar{v}v > 0$ per ogni $v \neq 0$ in $\mathbb{C}^n$.

dim. Se ϕ ammette una base ortonormale di autovettori $\mathcal{V} = \{v_1, \dots, v_n\}$ la matrice $D = \alpha_{\mathcal{V},\mathcal{V}}(\phi)$ è diagonale e quindi simmetrica (${}^tD = D$) per cui ϕ è un endomorfismo simmetrico.
Viceversa, sia $\phi : \mathbb{R}^n \to \mathbb{R}^n$ un endomorfismo simmetrico. Allora

- gli autovalori di ϕ sono tutti reali. Sia $c \in \mathbb{C}$ un autovalore di ϕ (ovvero dell'endomorfismo di $\mathbb{C}^n$ che, in base canonica, ha la stessa matrice $A \in M_n(\mathbb{R})$ di ϕ) e sia $v \in \mathbb{C}^n$ un corrispondente autovettore. Allora
$$c\,{}^t\bar{v}v = {}^t\bar{v}(Av) = ({}^t\bar{v}A)v = {}^t(\bar{A}\bar{v})v = \bar{c}\,{}^t\bar{v}v,$$
da cui si deduce $c = \bar{c}$, essendo ${}^t\bar{v}v > 0$ per ogni $v \neq 0$ in $\mathbb{C}^n$.
- autovettori relativi ad autovalori distinti sono ortogonali. Siano $a \neq b$ autovalori di ϕ e siano v e w in $\mathbb{R}^n$ autovettori ad essi relativi. Allora
$$a(v \cdot w) = (av) \cdot w = \phi(v) \cdot w = v \cdot \phi(w) = v \cdot (bw) = b(v \cdot w),$$
da cui si deduce $(a - b)(v \cdot w) = 0$ ed essendo $a \neq b$, si conclude che $v \cdot w = 0$.

dim. Se ϕ ammette una base ortonormale di autovettori $\mathcal{V} = \{v_1, \dots, v_n\}$ la matrice $D = \alpha_{\mathcal{V},\mathcal{V}}(\phi)$ è diagonale e quindi simmetrica (${}^tD = D$) per cui ϕ è un endomorfismo simmetrico.
Viceversa, sia $\phi : \mathbb{R}^n \to \mathbb{R}^n$ un endomorfismo simmetrico. Allora

- gli autovalori di ϕ sono tutti reali. Sia $c \in \mathbb{C}$ un autovalore di ϕ (ovvero dell'endomorfismo di $\mathbb{C}^n$ che, in base canonica, ha la stessa matrice $A \in M_n(\mathbb{R})$ di ϕ) e sia $v \in \mathbb{C}^n$ un corrispondente autovettore. Allora
$$c\,{}^t\bar{v}v = {}^t\bar{v}(Av) = ({}^t\bar{v}A)v = {}^t(\bar{A}\bar{v})v = \bar{c}\,{}^t\bar{v}v,$$
da cui si deduce $c = \bar{c}$, essendo ${}^t\bar{v}v > 0$ per ogni $v \neq 0$ in $\mathbb{C}^n$.
- autovettori relativi ad autovalori distinti sono ortogonali. Siano $a \neq b$ autovalori di ϕ e siano v e w in $\mathbb{R}^n$ autovettori ad essi relativi. Allora
$$a(v \cdot w) = (av) \cdot w = \phi(v) \cdot w = v \cdot \phi(w) = v \cdot (bw) = b(v \cdot w),$$
da cui si deduce $(a - b)(v \cdot w) = 0$ ed essendo $a \neq b$, si conclude che $v \cdot w = 0$.
- gli autovettori di ϕ generano tutto $\mathbb{R}^n$. Sia W il sottospazio di $\mathbb{R}^n$ generato dagli autovettori di ϕ. Allora $\mathbb{R}^n = W \oplus W^\perp$ e ϕ induce un endomorfismo in $W^\perp$. Infatti, se $v \in W^\perp$ e w è un autovettore, relativo all'autovalore c, allora $\phi(v) \cdot w = v \cdot \phi(w) = c(v \cdot w) = 0$ e quindi $\phi(v) \in W^\perp$. Se fosse $W^\perp \neq \langle 0 \rangle$ si avrebbe una contraddizione, perché il polinomio caratteristico di $\phi_{|W^\perp}$ divide il polinomio caratteristico di ϕ e quindi avrebbe radici reali ed esisterebbero autovettori per ϕ in $W^\perp$, contro l'ipotesi che W contenga tutti gli autovettori per ϕ.

Ciò conclude la dimostrazione. $\square$

Valori singolari

Mostriamo una conseguenza del Teorema Spettrale di largo uso nelle applicazioni.

Proposizione

Sia $\phi : \mathbb{R}^n \to \mathbb{R}^m$ un'applicazione lineare. Allora esistono basi ortonormali, $\mathcal{V} = \{v_1, \dots, v_n\}$ di $\mathbb{R}^n$, $\mathcal{W} = \{w_1, \dots, w_m\}$ di $\mathbb{R}^m$, e scalari $s_1 \geq s_2 \geq \cdots \geq s_r > 0$, ove $r = \operatorname{rk} \phi$, tali che $\phi(v_i) = s_i w_i$ per $i = 1, \dots, r$ e $\phi(v_i) = 0$ per $i > r$.
I numeri reali $s_1, s_2, \dots, s_r$ sono detti i *valori singolari* di ϕ.

Valori singolari

Mostriamo una conseguenza del Teorema Spettrale di largo uso nelle applicazioni.

Proposizione

Sia $\phi : \mathbb{R}^n \to \mathbb{R}^m$ un'applicazione lineare. Allora esistono basi ortonormali, $\mathcal{V} = \{v_1, \dots, v_n\}$ di $\mathbb{R}^n$, $\mathcal{W} = \{w_1, \dots, w_m\}$ di $\mathbb{R}^m$, e scalari $s_1 \geq s_2 \geq \cdots \geq s_r > 0$, ove $r = \operatorname{rk}\phi$, tali che $\phi(v_i) = s_i w_i$ per $i = 1, \dots, r$ e $\phi(v_i) = 0$ per $i > r$.
I numeri reali $s_1, s_2, \dots, s_r$ sono detti i *valori singolari* di ϕ.

Nelle notazioni della Proposizione precedente, siano $A = \alpha_{\mathcal{E}_n, \mathcal{E}_m}(\phi)$, $P = \alpha_{\mathcal{W}, \mathcal{E}_m}(\mathrm{id})$, $Q = \alpha_{\mathcal{V}, \mathcal{E}_n}(\mathrm{id})$ e sia $S \in M_{m \times n}(\mathbb{R})$ con tutte le entrate nulle eccetto lo scalare s_i nel posto (i, i), per $i = 1, \dots r$. Si ha $A = PS^tQ$ ove P e Q sono matrici ortogonali e questa scrittura è detta una decomposizione ai valori singolari della matrice A.

Valori singolari

Mostriamo una conseguenza del Teorema Spettrale di largo uso nelle applicazioni.

Proposizione

Sia $\phi : \mathbb{R}^n \to \mathbb{R}^m$ un'applicazione lineare. Allora esistono basi ortonormali, $\mathcal{V} = \{v_1, \dots, v_n\}$ di $\mathbb{R}^n$, $\mathcal{W} = \{w_1, \dots, w_m\}$ di $\mathbb{R}^m$, e scalari $s_1 \geq s_2 \geq \cdots \geq s_r > 0$, ove $r = \operatorname{rk} \phi$, tali che $\phi(v_i) = s_i w_i$ per $i = 1, \dots, r$ e $\phi(v_i) = 0$ per $i > r$.
I numeri reali $s_1, s_2, \dots, s_r$ sono detti i *valori singolari* di ϕ.

Nelle notazioni della Proposizione precedente, siano $A = \alpha_{\mathcal{E}_n, \mathcal{E}_m}(\phi)$, $P = \alpha_{\mathcal{W}, \mathcal{E}_m}(\mathrm{id})$, $Q = \alpha_{\mathcal{V}, \mathcal{E}_n}(\mathrm{id})$ e sia $S \in M_{m \times n}(\mathbb{R})$ con tutte le entrate nulle eccetto lo scalare s_i nel posto (i, i), per $i = 1, \dots r$. Si ha $A = PS^tQ$ ove P e Q sono matrici ortogonali e questa scrittura è detta una decomposizione ai valori singolari della matrice A.

dim. Siano $V = (\ker \phi)^\perp$ e $W = \operatorname{im} \phi$. Si ha $\dim V = \dim W = r$ e $\phi_{|V} : V \to W$ è un isomorfismo, perché iniettivo. Definiamo l'endomorfismo simmetrico $\psi : V \to V$, ponendo $\psi(v) \cdot v' = \phi(v) \cdot \phi(v')$ per ogni v, v' in V; e e osserviamo che gli autovalori di ψ sono tutti positivi. Infatti, se $\psi(v) = cv$, con $v \neq 0$, allora $c(v \cdot v) = \psi(v) \cdot v = \phi(v) \cdot \phi(v) > 0$.

Valori singolari

Mostriamo una conseguenza del Teorema Spettrale di largo uso nelle applicazioni.

Proposizione

Sia $\phi : \mathbb{R}^n \to \mathbb{R}^m$ un'applicazione lineare. Allora esistono basi ortonormali, $\mathcal{V} = \{v_1, \dots, v_n\}$ di $\mathbb{R}^n$, $\mathcal{W} = \{w_1, \dots, w_m\}$ di $\mathbb{R}^m$, e scalari $s_1 \geq s_2 \geq \cdots \geq s_r > 0$, ove $r = \operatorname{rk} \phi$, tali che $\phi(v_i) = s_i w_i$ per $i = 1, \dots, r$ e $\phi(v_i) = 0$ per $i > r$.
I numeri reali $s_1, s_2, \dots, s_r$ sono detti i *valori singolari* di ϕ.

Nelle notazioni della Proposizione precedente, siano $A = \alpha_{\mathcal{E}_n,\mathcal{E}_m}(\phi)$, $P = \alpha_{\mathcal{W},\mathcal{E}_m}(\mathrm{id})$, $Q = \alpha_{\mathcal{V},\mathcal{E}_n}(\mathrm{id})$ e sia $S \in M_{m\times n}(\mathbb{R})$ con tutte le entrate nulle eccetto lo scalare s_i nel posto (i,i), per $i = 1, \dots r$. Si ha $A = PS^tQ$ ove P e Q sono matrici ortogonali e questa scrittura è detta una decomposizione ai valori singolari della matrice A.

dim. Siano $V = (\ker \phi)^\perp$ e $W = \operatorname{im} \phi$. Si ha $\dim V = \dim W = r$ e $\phi_{|V} : V \to W$ è un isomorfismo, perché iniettivo. Definiamo l'endomorfismo simmetrico $\psi : V \to V$, ponendo $\psi(v) \cdot v' = \phi(v) \cdot \phi(v')$ per ogni v, v' in V; e e osserviamo che gli autovalori di ψ sono tutti positivi. Infatti, se $\psi(v) = cv$, con $v \neq 0$, allora $c(v \cdot v) = \psi(v) \cdot v = \phi(v) \cdot \phi(v) > 0$. Per il Teorema spettrale, esiste una base ortonormale $\{v_1, \dots, v_r\}$ di V fatta di autovettori per ψ e, a meno di riordinare i vettori di base, siano $s_1 \geq s_2 \geq \cdots \geq s_r$ tali che $\psi(v_i) = s_i^2 v_i$, per $i = 1, \dots, r$. Dunque, $\phi(v_i) \cdot \phi(v_j) = \psi(v_i) \cdot v_j = s_i^2 \delta_{ij}$, per ogni coppia di indici i, j tra 1 e r; perciò i vettori $w_i = \frac{1}{s_i}\phi(v_i)$, per $i = 1, \dots, r$ sono una base ortonormale di $W = \operatorname{im} \phi$ e si ha $\phi(v_i) = s_i w_i$ per $i = 1, \dots, r$.

Valori singolari

Mostriamo una conseguenza del Teorema Spettrale di largo uso nelle applicazioni.

Proposizione

Sia $\phi : \mathbb{R}^n \to \mathbb{R}^m$ un'applicazione lineare. Allora esistono basi ortonormali, $\mathcal{V} = \{v_1, \dots, v_n\}$ di $\mathbb{R}^n$, $\mathcal{W} = \{w_1, \dots, w_m\}$ di $\mathbb{R}^m$, e scalari $s_1 \geq s_2 \geq \cdots \geq s_r > 0$, ove $r = \operatorname{rk}\phi$, tali che $\phi(v_i) = s_i w_i$ per $i = 1, \dots, r$ e $\phi(v_i) = 0$ per $i > r$.
I numeri reali $s_1, s_2, \dots, s_r$ sono detti i *valori singolari* di ϕ.

Nelle notazioni della Proposizione precedente, siano $A = \alpha_{\mathcal{E}_n,\mathcal{E}_m}(\phi)$, $P = \alpha_{\mathcal{W},\mathcal{E}_m}(\mathrm{id})$, $Q = \alpha_{\mathcal{V},\mathcal{E}_n}(\mathrm{id})$ e sia $S \in M_{m\times n}(\mathbb{R})$ con tutte le entrate nulle eccetto lo scalare s_i nel posto (i,i), per $i = 1, \dots r$. Si ha $A = PS^tQ$ ove P e Q sono matrici ortogonali e questa scrittura è detta una decomposizione ai valori singolari della matrice A.

dim. Siano $V = (\ker\phi)^\perp$ e $W = \operatorname{im}\phi$. Si ha $\dim V = \dim W = r$ e $\phi_{|V} : V \to W$ è un isomorfismo, perché iniettivo. Definiamo l'endomorfismo simmetrico $\psi : V \to V$, ponendo $\psi(v)\cdot v' = \phi(v)\cdot\phi(v')$ per ogni v, v' in V; e e osserviamo che gli autovalori di ψ sono tutti positivi. Infatti, se $\psi(v) = cv$, con $v \neq 0$, allora $c(v\cdot v) = \psi(v)\cdot v = \phi(v)\cdot\phi(v) > 0$. Per il Teorema spettrale, esiste una base ortonormale $\{v_1, \dots, v_r\}$ di V fatta di autovettori per ψ e, a meno di riordinare i vettori di base, siano $s_1 \geq s_2 \geq \cdots \geq s_r$ tali che $\psi(v_i) = s_i^2 v_i$, per $i = 1, \dots, r$. Dunque, $\phi(v_i)\cdot\phi(v_j) = \psi(v_i)\cdot v_j = s_i^2\delta_{ij}$, per ogni coppia di indici i, j tra 1 e r; perciò i vettori $w_i = \frac{1}{s_i}\phi(v_i)$, per $i = 1, \dots, r$ sono una base ortonormale di $W = \operatorname{im}\phi$ e si ha $\phi(v_i) = s_i w_i$ per $i = 1, \dots, r$. Unendo alla base di V una base ortonormale $\{v_{r+1}, \dots, v_n\}$ di $\ker\phi = V^\perp$, si ha una base ortonormale $\mathcal{V} = \{v_1, \dots, v_n\}$ di $\mathbb{R}^n$. Analogamente si possono completare i vettori $w_1, \dots, w_r$ a una base ortonormale $\mathcal{W} = \{w_1, \dots, w_m\}$ di $\mathbb{R}^m$. Si ha dunque $S = \alpha_{\mathcal{V},\mathcal{W}}(\phi)$, ove $S \in M_{m\times n}(\mathbb{R})$ ha tutte le entrate nulle eccetto i valori singolari s_i nei posti (i,i), per $i = 1, \dots r$. □

N.B. Se A è una matrice reale simmetrica (${}^tA = A$), allora i valori singolari di A coincidono con i valori assoluti degli autovalori di A, ma, per una generica matrice quadrata, c'è differenza tra autovalori e valori singolari. Ad esempio, la matrice $\begin{pmatrix} 0 & -2 \\ 0 & 0 \end{pmatrix}$ ha decomposizione ai valori singolari $\begin{pmatrix} 0 & -2 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ e quindi i suoi valori singolari sono 2 e 0, mentre vi è il solo autovalore 0, con molteplicità 2.

Esempio: Sia $A = \begin{pmatrix} 1 & 0 & -2 & 0 \\ 0 & 1 & 0 & -1 \\ -3 & 0 & 6 & 0 \end{pmatrix}$. Si determini una decomposizione ai valori singolari di A.

Svolg.: La matrice $A^tA = \begin{pmatrix} 5 & 0 & -15 \\ 0 & 2 & 0 \\ -15 & 0 & 45 \end{pmatrix}$ ha gli autovalori $50, 2, 0$ a cui corrispondono i sottospazi di autovettori $\left\langle \begin{pmatrix} 1 \\ 0 \\ -3 \end{pmatrix} \right\rangle$, $\left\langle \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\rangle$, $\left\langle \begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix} \right\rangle$. Quindi i valori singolari di A sono $5\sqrt{2}$ e $\sqrt{2}$, e perciò

$$S = \begin{pmatrix} 5\sqrt{2} & 0 & 0 & 0 \\ 0 & \sqrt{2} & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \qquad \text{e} \qquad P = \begin{pmatrix} 1/\sqrt{10} & 0 & 3/\sqrt{10} \\ 0 & 1 & 0 \\ -3/\sqrt{10} & 0 & 1/\sqrt{10} \end{pmatrix};$$

ove P ha come colonne una base ortonormale di autovettori per A^tA. Il nucleo di A è il sottospazio $\left\langle \begin{pmatrix} 2 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} \right\rangle$ e una base ortonormale di questo sottospazio ci dà le ultime due colonne della matrice Q. Per trovare le prime due colonne, possiamo osservare che dalla relazione $A = PS{}^tQ$, si ricava $Q{}^tS = {}^tAP$, essendo ${}^tQQ = \mathbf{1}_4$ e ${}^tPP = \mathbf{1}_3$. Quindi le prime due colonne di Q, moltiplicate per i valori singolari, sono le prime due colonne del prodotto ${}^tAP$. Si ottiene così $Q = \begin{pmatrix} 1/\sqrt{5} & 0 & 2/\sqrt{5} & 0 \\ 0 & 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ -2/\sqrt{5} & 0 & 1/\sqrt{5} & 0 \\ 0 & -1/\sqrt{2} & 0 & 1/\sqrt{2} \end{pmatrix}$. Si verifichi che $PS = AQ$.

$\square$

Pseudoinversa - Minimi quadrati

Sia $A = PS\,{}^tQ$ una decomposizione a valori singolari della matrice $A \in M_{m\times n}(\mathbb{R})$ e indichiamo con $S_{-1} \in M_{n\times m}(\mathbb{R})$ la matrice che ha tutte le entrate nulle, eccetto gli inversi dei valori singolari s_i^{-1} nel posto (i,i), per $i = 1,\dots,r$. La matrice $A^+ = QS_{-1}\,{}^tP \in M_{n\times m}(\mathbb{R})$ è la pseudoinversa di Moore-Penrose della matrice A.
Nelle notazioni della dimostrazione precedente, siano $\phi : \mathbb{R}^n \to \mathbb{R}^m$ l'omomorfismo di matrice A nelle basi canoniche, $\mathcal{V} = \{v_1,\dots,v_n\}$, $\mathcal{W} = \{w_1,\dots,w_m\}$, le basi ortonormali rispetto a cui $S = \alpha_{\mathcal{V},\mathcal{W}}(\phi)$, e $P = \alpha_{\mathcal{W},\mathcal{E}_m}(\mathrm{id})$, $Q = \alpha_{\mathcal{V},\mathcal{E}_n}(\mathrm{id})$. Allora, A^+ è la matrice (in basi canoniche) di quell'unica applicazione lineare $\phi^+ : \mathbb{R}^m \to \mathbb{R}^n$ che si annulla su $(\mathrm{im}\ \phi)^\perp$ e manda ogni vettore di $\mathrm{im}\ \phi$ nella sua unica controimmagine appartenente a $(\ker \phi)^\perp$. Si ha $SS_{-1} = \left(\begin{array}{c|c} \mathbf{1}_r & 0 \\ \hline 0 & 0 \end{array}\right) \in M_m(\mathbb{R})$ e $S_{-1}S = \left(\begin{array}{c|c} \mathbf{1}_r & 0 \\ \hline 0 & 0 \end{array}\right) \in M_n(\mathbb{R})$, ove $r = \mathrm{rk}\ A = \mathrm{rk}\ A^+$; quindi

- $AA^+ = P(SS_{-1})\,{}^tP$ è la matrice della proiezione ortogonale su $\mathrm{im}\ \phi$;
- $A^+A = Q(S_{-1}S)\,{}^tQ$ è la matrice della proiezione ortogonale su $(\ker \phi)^\perp$.

In particolare, si ha $A^+AA^+ = A^+$ e $AA^+A = A$.

Dato il sistema lineare $Ax = b$, con $A \in M_{m\times n}(\mathbb{R})$ e $b \in \mathbb{R}^m$, sia $\phi : \mathbb{R}^n \to \mathbb{R}^m$ l'applicazione lineare di matrice A nelle basi canoniche. Sappiamo che il sistema non ha soluzione se $b \notin \mathrm{im}\ \phi$. Possiamo però cercare il vettore $x_0 \in \mathbb{R}^n$ per cui la distanza $\|Ax_0 - b\|$ sia minima. Questo vettore x_0 si chiama una *soluzione ai minimi quadrati del sistema lineare $Ax = b$*.
Nel punto (d) della Proposizione sulle proprietà degli ortogonali, abbiamo dimostrato che il punto di $\mathrm{im}\ \phi$ che ha minima distanza da b è la proiezione ortogonale $p_{\mathrm{im}\ \phi}(b)$. Dobbiamo quindi trovare un vettore $x_0 \in \mathbb{R}^n$ tale che $\phi(x_0) = p_{\mathrm{im}\ \phi}(b)$. Per quanto visto sulla pseudoinversa di ϕ, basta prendere $x_0 = \phi^+(b)$, ovvero calcolare $x_0 = A^+b$, perché $A^+b = A^+AA^+b$, ove AA^+b è la proiezione ortogonale di b su $\mathrm{im}\ \phi$ e, moltiplicando a