Parte 8. Prodotto scalare, teorema spettrale

A. Savo − Appunti del Corso di Geometria 2013-14

Indice delle sezioni

1 Prodotto scalare in $\mathbf{R}^n$

1.1 Definizione

Dati i vettori $u = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ e $v = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ di $\mathbf{R}^n$, definiamo *prodotto scalare di u e v* il numero reale:

$$\langle u, v \rangle = x_1 y_1 + \cdots + x_n y_n.$$

Il risultato del prodotto scalare è dunque un numero.

Esempio Se $u = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ e $v = \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix}$ allora $\langle u, v \rangle = 3$.

Notiamo che il prodotto scalare di due vettori può risultare nullo anche se nessuno dei due fattori è il vettore nullo.

Esempio Se $u = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ e $v = \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}$ allora $\langle u, v \rangle = 0$.

- I vettori u e v si dicono *ortogonali* se il loro prodotto scalare è nullo: $\langle u, v\rangle = 0$.

Notazione:

$$u \perp v.$$

Dunque i vettori dell'esempio precedente sono ortogonali. È evidente che, se O è il vettore nullo, si ha $\langle v, O\rangle = 0$ per ogni $v \in \mathbf{R}^n$: dunque il vettore nullo è ortogonale a tutti i vettori.

La denominazione di *vettori ortogonali* legata alla condizione $\langle u, v\rangle = 0$ (che è puramente algebrica) sarà giustificata quando studieremo la geometria analitica, e introdurremo i vettori geometrici del piano e dello spazio. Infatti, l'introduzione del prodotto scalare permette di definire, in modo puramente algebrico, la *norma* di un vettore (che va intepretata come la distanza del vettore stesso dal vettore nullo) e l'*angolo* fra due vettori non nulli. Per il momento, ci proponiamo di studiare le proprietà algebriche dell'operazione di prodotto scalare.

Proposizione *Siano u, v, w vettori arbitrari di $\mathbf{R}^n$ e sia $k \in \mathbf{R}$ un qualunque scalare. Allora si hanno le seguenti proprietà.*

1) $\langle u, v\rangle = \langle v, u\rangle$.

2) $\langle u + v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$.

3) $\langle ku, v\rangle = k\langle u, v\rangle$.

4) $\langle u, u\rangle \geq 0$.

5) $\langle u, u\rangle = 0$ *se e solo se* $u = O$.

La 1) dice che il prodotto scalare è commutativo. Le proprietà 2), 3) esprimono la cosiddetta proprietà di *bilinearità.* Le proprietà 4) e 5) esprimono il fatto che il prodotto scalare è *definito positivo.*

Dimostrazione. La dimostrazione di 1),2),3) si riduce a una semplice verifica. Osserviamo che, se $u = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ allora

$$\langle u, u\rangle = x_1^2 + \cdots + x_n^2,$$

che è un numero sempre positivo o nullo: questo dimostra la 4). Se $\langle u, u\rangle = 0$ evidentemente $x_1 = \cdots = x_n = 0$, e quindi $u = O$. $\square$

Dalle proprietà di bilinearità osserviamo che il prodotto scalare si comporta in modo naturale rispetto alle combinazioni lineari. Per ogni scelta dei vettori $v_1, \ldots, v_k, u, w \in \mathbf{R}^n$ e degli scalari $a_1, \ldots, a_k \in \mathbf{R}$ si ha:

$$\langle a_1 v_1 + \cdots + a_k v_k, w\rangle = a_1\langle v_1, w\rangle + \cdots + a_k\langle v_k, w\rangle.$$

Di conseguenza, poiché il prodotto scalare è commutativo, si ha anche

$$\langle u, a_1 v_1 + \cdots + a_k v_k \rangle = a_1 \langle u, v_1 \rangle + \cdots + a_k \langle u, v_k \rangle$$

1.2 Norma e disuguaglianza di Schwarz

Per definizione, la *norma* di un vettore $u \in \mathbf{R}^n$ è il numero positivo o nullo

$$\|u\| = \sqrt{\langle u, u \rangle}.$$

Esplicitamente $\|u\| = \sqrt{x_1^2 + \cdots + x_n^2}$ ovvero

$$\|u\|^2 = x_1^2 + \cdots + x_n^2.$$

In particolare, $\|u\| \geq 0$ e si ha l'uguaglianza solo quando $u = O$: la norma di un vettore non nullo è sempre positiva.

Esempio Se $u = \begin{pmatrix} 1 \\ 3 \\ -2 \end{pmatrix}$ allora $\|u\| = \sqrt{1 + 9 + 4} = \sqrt{14}$.

Teorema (Disuguaglianza di Schwarz). *Dati $u, v \in \mathbf{R}^n$ si ha sempre:*

$$|\langle u, v \rangle| \leq \|u\| \|v\|.$$

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

Dimostrazione. Omessa. □

Esempio Dati n numeri reali $a_1, \ldots, a_n$ si ha sempre:

$$(a_1 + \cdots + a_n)^2 \leq n(a_1^2 + \cdots + a_n^2).$$

Infatti, basta applicare la disuguaglianza di Schwarz ai vettori $u = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$ e $v = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$.

Notiamo che si ha l'uguaglianza solo quando $a_1, \ldots, a_n$ sono tutti uguali tra loro.

1.3 Angolo tra due vettori

Supponiamo che u e v siano due vettori non nulli. Per la disuguaglianza di Schwarz, si ha

$$\frac{|\langle u, v\rangle|}{\|u\|\|v\|} \leq 1,$$

dunque esiste un unico valore di $\theta \in [0, \pi]$ tale che

$$\cos\theta = \frac{\langle u, v\rangle}{\|u\|\|v\|}.$$

Per definizione, θ è detto l'*angolo* tra u e v.

Esempio Dati $u = \begin{pmatrix} 1 \\ 2 \\ 1 \\ 0 \end{pmatrix}, v = \begin{pmatrix} 1 \\ 0 \\ 2 \\ 1 \end{pmatrix}$ si ha:

$$\|u\| = \sqrt{6}, \quad \|v\| = \sqrt{6}, \quad \langle u, v\rangle = 3.$$

Dunque $\cos\theta = \frac{1}{2}$ cioè $\theta = \frac{\pi}{3}$.

2 Basi ortonormali

2.1 Ortogonalità e indipendenza lineare

Proposizione *Siano $v_1, \dots, v_k$ vettori non nulli di $\mathbf{R}^n$, a due a due ortogonali. Allora $v_1, \dots, v_k$ sono linearmente indipendenti. In particolare, $k \leq n$.*

Dimostrazione. Supponiamo che

$$a_1 v_1 + \cdots + a_k v_k = O. \tag{1}$$

Prendendo il prodotto scalare dei due membri della (1) per v_1 otteniamo

$$\begin{aligned} 0 &= \langle a_1 v_1 + a_2 v_2 + \cdots + a_k v_k, v_1\rangle \\ &= a_1\langle v_1, v_1\rangle + a_2\langle v_2, v_1\rangle + \cdots + a_k\langle v_k, v_1\rangle \\ &= a_1\|v_1\|^2 \end{aligned}$$

perché per ipotesi $\langle v_j, v_1\rangle = 0$ per ogni $j = 2, \dots, k$. Per ipotesi, v_1 è non nullo, dunque $\|v_1\|^2 > 0$ e ne segue che $a_1 = 0$. Prendendo successivamente il prodotto scalare dei due membri della (1), ordinatamente per $v_2, \ldots, v_k$, si dimostra in modo analogo che $a_j = 0$ per ogni j. $\Box$

- n vettori non nulli, a due a due ortogonali formano una base di $\mathbf{R}^n$ (che sarà chiamata *base ortogonale*).

Esempio I vettori $v_1 = \begin{pmatrix} 1 \\ -2 \end{pmatrix}, v_2 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ formano una base ortogonale di $\mathbf{R}^2$ perchè $\langle v_1, v_2 \rangle = 0$.

I vettori $w_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, w_2 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, w_3 = \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}$ sono non nulli e a due a due ortogonali:

$$\langle w_1, w_2 \rangle = \langle w_1, w_3 \rangle = \langle w_2, w_3 \rangle = 0.$$

Dunque (w_1, w_2, w_3) è una base ortogonale di $\mathbf{R}^3$.

Esempio La matrice

$$A = \begin{pmatrix} 1 & 1 & 1 & -1 \\ 1 & 1 & -1 & 1 \\ 1 & -1 & 1 & 1 \\ 1 & -1 & -1 & -1 \end{pmatrix}$$

ha rango 4. Infatti i suoi vettori colonna sono a due a due ortogonali, e quindi sono linearmente indipendenti.

- Il numero *massimo* di vettori di $\mathbf{R}^n$, non nulli e ortogonali a due a due, è n.

In modo analogo, possiamo definire la nozione di *base ortogonale* di un qualunque sottospazio E di $\mathbf{R}^n$: se $\dim E = k$ allora i vettori $v_1, \ldots, v_k$ formano una base ortogonale di E se sono non nulli e $\langle v_i, v_j \rangle = 0$ per ogni $i \neq j$.

Esempio Il sottospazio $E : x + y + z = 0$ di $\mathbf{R}^3$ ha dimensione 2. I due vettori $v_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, v_2 = \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}$ appartengono a E e sono ortogonali tra loro, dunque formano una base ortogonale di E.

2.2 Basi ortonormali

Diremo che una base $(v_1, \ldots, v_k)$ di un sottospazio E di $\mathbf{R}^n$ è una *base ortonormale* di E se:

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

Dunque una base ortonormale è formata da vettori a due a due ortogonali, tutti di norma unitaria. Una base ortonormale è, in particolare, anche ortogonale.

Esempio La base canonica di $\mathbf{R}^n$ è una base ortonormale.

Fare i conti con le basi ortonormali è più semplice. Ad esempio, trovare le coordinate di un vettore rispetto a una base implica, normalmente, la risoluzione di un certo numero di sistemi lineari. Se la base è ortonormale, è sufficiente calcolare un certo numero di prodotti scalari.

Proposizione *Sia $\mathcal{B} = (v_1, \dots, v_k)$ una base ortonormale di un sottospazio E di $\mathbf{R}^n$. Allora le coordinate del vettore $v \in E$ rispetto a $\mathcal{B}$ sono date da*

$$\begin{pmatrix} \langle v, v_1 \rangle \\ \vdots \\ \langle v, v_k \rangle \end{pmatrix},$$

e sono dette coefficienti di Fourier *di v.*

Dimostrazione. Se $v \in E$ possiamo scrivere

$$v = a_1 v_1 + \cdots + a_k v_k$$

e per definizione le coordinate di v sono $a_1, \dots, a_k$. Ora, prendendo il prodotto scalare dei due membri successivamente per $v_1, \dots, v_k$, otteniamo facilmente

$$a_j = \langle v, v_j \rangle$$

per ogni $j = 1, \dots, k$. $\square$

Esempio I vettori:

$$v_1 = \frac{1}{2}\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, v_2 = \frac{1}{2}\begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix}, v_3 = \frac{1}{2}\begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}, v_4 = \frac{1}{2}\begin{pmatrix} -1 \\ 1 \\ 1 \\ -1 \end{pmatrix},$$

sono a due a due ortogonali e hanno tutti norma 1. Dunque tali vettori formano una base ortonormale $\mathcal{B}$ di $\mathbf{R}^4$. Calcoliamo le coordinate del vettore $v = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}$ rispetto a $\mathcal{B}$. I coefficienti di Fourier sono

$$\langle v, v_1 \rangle = 5,\ \langle v, v_2 \rangle = -2,\ \langle v, v_3 \rangle = -1,\ \langle v, v_4 \rangle = 0.$$

Dunque v ha coordinate

$$\begin{pmatrix} 5 \\ -2 \\ -1 \\ 0 \end{pmatrix},$$

rispetto a $\mathcal{B}$. In altre parole $v = 5v_1 - 2v_2 - v_3$. $\square$

3 Algoritmo di Gram-Schmidt

Lo scopo di questa sezione è quello di dimostrare che ogni sottospazio di $\mathbf{R}^n$ ammette almeno una base ortonormale.

3.1 Vettore normalizzato

Proposizione 1) *Dato un vettore v e uno scalare $a \in \mathbf{R}$ si ha:* $\|av\| = |a|\|v\|$.

2) *Se $v \neq O$ il vettore*

$$u = \frac{1}{\|v\|}v$$

ha norma 1.

Dimostrazione. Si ha, dalle proprietà del prodotto scalare:

$$\|av\|^2 = \langle av, av \rangle = a^2 \langle v, v \rangle = a^2 \|v\|^2,$$

e la 1) segue prendendo la radice quadrata ad ambo i membri. La 2) segue immediatamente dalla 1) prendendo $a = \dfrac{1}{\|v\|}$. $\square$

Il vettore $u = \dfrac{1}{\|v\|}v$ si dice *normalizzato* di v. Normalizzare un vettore significa semplicemente dividere il vettore per la propria norma.

Esempio Il vettore $v = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ ha norma $\sqrt{14}$. Il suo normalizzato è dunque

$$u = \frac{1}{\|v\|}v = \frac{1}{\sqrt{14}} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

e ha norma 1.

Corollario *Se* $(v_1, \dots, v_k)$ *è una base ortogonale del sottospazio* E*, allora i vettori normalizzati*

$$u_1 = \frac{1}{\|v_1\|} v_1, \quad \dots \quad , u_k = \frac{1}{\|v_k\|} v_k$$

formano una base ortonormale di E*.*

Dimostrazione. I vettori $u_1, \dots, u_k$ hanno tutti norma 1, ed evidentemente appartengono a E. Essi sono a due a due ortogonali, poiché

$$\langle u_i, u_j \rangle = \frac{1}{\|v_i\|} \frac{1}{\|v_j\|} \langle v_i, v_j \rangle = 0$$

per ogni $i \neq j$. $\square$

Esempio Il sottospazio di $\mathbf{R}^3$ definito dall'equazione $E : x + y + z = 0$ ha dimensione 2 e ha una base ortogonale formata dai vettori

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

Per ottenere una base ortonormale di E è sufficiente normalizzare i vettori v_1, v_2. Si ottiene la base ortonormale

$$u_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \quad u_2 = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}.$$

3.2 Procedimento di ortonormalizzazione

L'algoritmo di Gram-Schmidt è un procedimento che, applicato ad una base di un sottospazio di $\mathbf{R}^n$, permette di ottenere una base ortogonale del sottospazio stesso; normalizzando i vettori di tale base, otterremo una base ortonormale. Descriviamo l'algoritmo in dettaglio.
Sia E un sottospazio di $\mathbf{R}^n$ e sia $(v_1, \dots, v_k)$ una sua base. Dunque $\dim E = k$.
Notiamo che se $k = 1$ la base è formata dal solo vettore v_1. È sufficiente dunque normalizzare v_1 per ottenere la base ortonormale cercata.
1) Supponiamo che la dimensione di E sia 2, e sia (v_1, v_2) una base di E. Introduciamo nuovi vettori (w_1, w_2) nel modo seguente:

$$\begin{cases} w_1 = v_1 \\ w_2 = v_2 - a w_1 \end{cases}$$

con $a \in \mathbf{R}$ da determinare in modo opportuno. Notiamo che i vettori w_1, w_2 appartengono a E; inoltre w_2 non è nullo (altrimenti v_1 e v_2 sarebbero linearmente dipendenti). Ora scegliamo il coefficiente a in modo tale che w_2 risulti ortogonale a w_1. È facile vedere che cio' accade se solo se:

$$a = \frac{\langle v_2, w_1 \rangle}{\langle w_1, w_1 \rangle}.$$

Dunque, con tale scelta, otteniamo la base ortogonale (w_1, w_2) di E.

2) Supponiamo ora che $\dim E = 3$, con base (v_1, v_2, v_3) e poniamo:

$$\begin{cases} w_1 = v_1 \\ w_2 = v_2 - a w_1 \\ w_3 = v_3 - b w_1 - c w_2 \end{cases}$$

dove $a = \dfrac{\langle v_2, w_1 \rangle}{\langle w_1, w_1 \rangle}$ è stato già determinato, cosicché $\langle w_1, w_2 \rangle = 0$. Imponendo le condizioni

$$\langle w_3, w_1 \rangle = \langle w_3, w_2 \rangle = 0,$$

otteniamo:

$$b = \frac{\langle v_3, w_1 \rangle}{\langle w_1, w_1 \rangle}, \quad c = \frac{\langle v_3, w_2 \rangle}{\langle w_2, w_2 \rangle}.$$

Con tali scelte di a, b, c otteniamo quindi la base ortogonale (w_1, w_2, w_3) di E e quindi, normalizzando, una base ortonormale (notiamo che w_3 non è nullo, altrimenti v_1, v_2, v_3 sarebbero linearmente dipendenti).

Procedendo per induzione, possiamo enunciare il seguente teorema.

Teorema (Algoritmo di Gram-Schmidt) *Sia* $(v_1, \dots, v_k)$ *una base del sottospazio* E *di* $\mathbf{R}^n$. *Si introducano i vettori:*

$$\begin{cases} w_1 = v_1 \\ w_2 = v_2 - a_{21} w_1 \\ w_3 = v_3 - a_{31} w_1 - a_{32} w_2 \\ \quad \dots \\ w_k = v_k - a_{k1} w_1 - a_{k2} w_2 - \dots - a_{k,k-1} w_{k-1} \end{cases}$$

dove si è posto:

$$a_{ij} = \frac{\langle v_i, w_j \rangle}{\langle w_j, w_j \rangle}.$$

Allora $(w_1, \dots, w_k)$ *è una base ortogonale di* E*, e quindi i vettori normalizzati:*

$$u_1 = \frac{1}{\|w_1\|} w_1, \dots, u_k = \frac{1}{\|w_k\|} w_k,$$

formano una base ortonormale di E*.*

Esempio Trovare una base ortonormale del sottospazio E di $\mathbf{R}^3$ di equazione:

$$E : x - y - 2z = 0.$$

Soluzione. Determiniamo una base di E, e poi applichiamo l'algoritmo di Gram-Schmidt per ottenere una base ortonormale. Base di E:

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, v_2 = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}.$$

L'algoritmo consiste di due passi:

$$\begin{cases} w_1 = v_1 \\ w_2 = v_2 - a_{21} w_1 \end{cases}.$$

Si ha $w_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$, dunque:

$$a_{21} = \frac{\langle v_2, w_1 \rangle}{\langle w_1, w_1 \rangle} = \frac{2}{2} = 1.$$

Allora:

$$\begin{cases} w_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \\ w_2 = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} \end{cases}.$$

Si verifica che in effetti $\langle w_1, w_2 \rangle = 0$. Una base ortonormale di E è dunque:

$$u_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad u_2 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}.$$

- Ovviamente la base ortonormale ottenuta dipende dalla base di partenza. Per esercizio, vedere quale base ortonormale si ottiene scambiando i vettori della base di partenza.

Esempio Trovare una base ortonormale del sottospazio di $\mathbf{R}^4$ generato dai vettori:

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \ v_2 = \begin{pmatrix} 2 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \ v_3 = \begin{pmatrix} 1 \\ 3 \\ -1 \\ 2 \end{pmatrix}.$$

Soluzione. I tre vettori formano una base di E. Applichiamo l'algoritmo di Gram-Schmidt alla terna v_1, v_2, v_3:

$$\begin{cases} w_1 = v_1 \\ w_2 = v_2 - a_{21} w_1 \\ w_3 = v_3 - a_{31} w_2 - a_{32} w_2 \end{cases}$$

Abbiamo $w_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}$, e quindi $a_{21} = 1$. Dunque

$$w_2 = \begin{pmatrix} 2 \\ 0 \\ 1 \\ 0 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 1 \\ 0 \end{pmatrix}.$$

Ora:

$$\begin{cases} a_{31} = \dfrac{\langle v_3, w_1 \rangle}{\langle w_1, w_1 \rangle} = \dfrac{4}{2} = 2 \\ a_{32} = \dfrac{\langle v_3, w_2 \rangle}{\langle w_2, w_2 \rangle} = \dfrac{-3}{3} = -1 \end{cases}$$

dunque:

$$w_3 = \begin{pmatrix} 1 \\ 3 \\ -1 \\ 2 \end{pmatrix} - 2 \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 1 \\ -1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 2 \end{pmatrix}.$$

Otteniamo la base ortogonale:

$$w_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \ w_2 = \begin{pmatrix} 1 \\ -1 \\ 1 \\ 0 \end{pmatrix}, \ w_3 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 2 \end{pmatrix},$$

e, normalizzando, la base ortonormale:

$$u_1 = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\\0\\0\end{pmatrix}, \; u_2 = \frac{1}{\sqrt{3}}\begin{pmatrix}1\\-1\\1\\0\end{pmatrix}, \; u_3 = \begin{pmatrix}0\\0\\0\\1\end{pmatrix}.$$

4 Matrici ortogonali

Abbiamo visto che la matrice M di passaggio fra due basi $\mathcal{B}, \mathcal{B}'$ di uno spazio vettoriale è invertibile. Se le basi $\mathcal{B}, \mathcal{B}'$ sono ortonormali, la matrice di passaggio avrà delle proprietà particolari.

Definizione *Una matrice quadrata M si dice* ortogonale *se verifica $MM^t = I$. Quindi una matrice ortogonale M è invertibile e*

$$M^{-1} = M^t,$$

cioè l'inversa coincide con la trasposta.

Esempio La matrice $M = \frac{1}{\sqrt{5}}\begin{pmatrix}1 & -2\\2 & 1\end{pmatrix} = \begin{pmatrix}\frac{1}{\sqrt{5}} & \frac{-2}{\sqrt{5}}\\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}}\end{pmatrix}$ è ortogonale.

Esempio La matrice $M = \begin{pmatrix}1/\sqrt{2} & 1/\sqrt{3} & 1/\sqrt{6}\\ -1/\sqrt{2} & 1/\sqrt{3} & 1/\sqrt{6}\\ 0 & 1/\sqrt{3} & -2/\sqrt{6}\end{pmatrix}$ è ortogonale.

In entrambi i casi si verifica infatti che $MM^t = I$.

Osserviamo che, se $MM^t = I$ allora, prendendo il determinante di ambo i membri e applicando il teorema di Binet, si ha $(\det M)^2 = 1$. Dunque

- se M è una matrice ortogonale allora $\det M = 1$ oppure $\det M = -1$.

Il teorema seguente fornisce le proprietà importanti di una matrice ortogonale.

Teorema a) *La matrice di passaggio fra due basi ortonormali di $\mathbf{R}^n$ (o di un suo sottospazio) è ortogonale.*

b) *Una matrice $A \in \mathbf{Mat}(n \times n)$ è ortogonale se e solo se le colonne di A formano una base ortonormale di $\mathbf{R}^n$.*

Dimostrazione. La dimostrazione si riduce a una verifica, che omettiamo. $\square$

Osserviamo che le colonne delle matrici ortogonali dei due esempi precedenti formano, effettivamente, una base ortonormale di $\mathbf{R}^2$ (primo esempio), e di $\mathbf{R}^3$ (secondo esempio). Dalla parte b) del teorema abbiamo anche

- Incolonnando i vettori di una base ortonormale di $\mathbf{R}^n$ otteniamo una matrice ortogonale $n \times n$.

Infine, si può dimostrare che le matrici ortogonali di $\mathbf{R}^2$ sono di due tipi:

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

con $\theta \in \mathbf{R}$, oppure

$$\begin{pmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix},$$

con $\alpha \in \mathbf{R}$. Le matrici del primo tipo hanno determinante 1, mentre quelle del secondo tipo hanno determinante -1.

5 Complemento ortogonale di un sottospazio

Sia u_1 un vettore fissato di $\mathbf{R}^n$ e si consideri il sottoinsieme

$$E = \{v \in \mathbf{R}^n : \langle v, u_1 \rangle = 0\},$$

formato da tutti i vettori ortogonali a u_1. Per le proprietà di bilinearità del prodotto scalare, E risulta allora un sottospazio di $\mathbf{R}^n$. Più in generale, fissati k vettori di $\mathbf{R}^n$, diciamo $u_1, \dots, u_k$, l'insieme:

$$E = \{v \in \mathbf{R}^n : \langle v, u_1 \rangle = \cdots = \langle v, u_k \rangle = 0\},$$

formato dai vettori di $\mathbf{R}^n$ ortogonali a $u_1, \dots, u_k$ è un sottospazio di $\mathbf{R}^n$.

Esempio Sia $u_1 = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$. Trovare una base di $E = \{v \in \mathbf{R}^3 : \langle v, u_1 \rangle = 0\}$.

Soluzione. Sia $v = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ il vettore generico di $\mathbf{R}^3$. Imponendo l'ortogonalità al vettore u_1 otteniamo l'unica condizione

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

Dunque E è il sottospazio delle soluzioni dell'equazione, e una sua base è, ad esempio, $\left(\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}\right)$. E ha dimensione 2. $\square$

Esempio Dati i vettori $u_1 = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}, u_2 = \begin{pmatrix} 0 \\ 1 \\ -2 \end{pmatrix}$ si consideri il sottospazio

$$F = \{v \in \mathbf{R}^3 : \langle v, u_1 \rangle = \langle v, u_2 \rangle = 0\}.$$

a) Trovare una base di F e calcolare la sua dimensione.

b) Trovare un vettore di F avente norma 1.

Soluzione. a) Imponendo al vettore generico $v = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ l'ortogonalità a u_1 e u_2 vediamo che F è descritto dalle equazioni

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

La matrice dei coefficienti del sistema che definisce F è $A = \begin{pmatrix} 1 & 1 & -1 \\ 0 & 1 & -2 \end{pmatrix}$. Notiamo che le righe di A sono proprio i vettori u_1, u_2 (piú precisamente, u_1^t, u_2^t): siccome u_1, u_2 sono linearmente indipendenti il rango vale 2 e l'insieme delle soluzioni F ha dimensione:

$$\dim F = 3 - \mathrm{rk} A = 1.$$

Una base si ottiene risolvendo il sistema. Si ottiene ad esempio la base $\begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix}$.

b) Un vettore di E di norma 1 si ottiene normalizzando il vettore della base trovata, dunque $w = \frac{1}{\sqrt{6}} \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix}$. Un altro vettore possibile è $-w = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ -2 \\ -1 \end{pmatrix}$. Verificare che non ce ne sono altri. $\square$

Generalizzando, otteniamo il seguente risultato.

Proposizione *Se i vettori $u_1, \dots, u_k \in \mathbf{R}^n$ sono linearmente indipendenti, allora*

$$E = \{v \in \mathbf{R}^n : \langle v, u_1 \rangle = \dots = \langle v, u_k \rangle = 0\},$$

è un sottospazio di $\mathbf{R}^n$ *di dimensione* $n-k$.

Dimostrazione. Abbiamo già osservato che E è un sottospazio. Sia $v = (x_1, \dots, x_n)^t$ il vettore generico di $\mathbf{R}^n$. Imponendo l'ortogonalità di v a ciascuno dei vettori $u_1, \dots, u_k$ otteniamo un sistema lineare omogeneo di k equazioni nelle n incognite $x_1, \dots, x_n$. Dunque E ha dimensione $n - \mathrm{rk}A$, dove A è la matrice dei coefficienti. Ora, si verifica che le righe di A sono i vettori trasposti di $u_1, \dots, u_k$. Poiché per ipotesi $u_1, \dots, u_k$ sono linearmente indipendenti il rango di A vale k e dunque

$$\dim E = n - k.$$

$\square$

5.1 Complemento ortogonale di un sottospazio

Sia E un sottospazio di $\mathbf{R}^n$. Definiamo *complemento ortogonale di* E l'insieme $E^\perp$ costituito dai vettori di $\mathbf{R}^n$ ortogonali a *tutti* i vettori di E:

$$E^\perp = \{v \in \mathbf{R}^n : \langle v, w \rangle = 0 \quad \text{per ogni } w \in E\}.$$

Dalle proprietà del prodotto scalare risulta che $E^\perp$ è chiuso rispetto alla somma e al prodotto per uno scalare, dunque è un sottospazio di $\mathbf{R}^n$.

Risulta che $v \in E^\perp$ se e solo se v è ortogonale a tutti i vettori di una base di E. Infatti:

Proposizione *Sia* $(v_1, \dots, v_k)$ *una base di* E. *Allora* $v \in E^\perp$ *se e solo se* $\langle v, v_i \rangle = 0$ *per ogni* $i = 1, \dots, k$.

Dimostrazione. Supponiamo che $\langle v, v_i \rangle = 0$ per ogni $i = 1, \dots, k$. Se w è un qualunque vettore di E, allora w è combinazione lineare dei vettori della base: $w = a_1 v_1 + \dots + a_k v_k$. Quindi

$$\langle v, w \rangle = a_1 \langle v, v_1 \rangle + \dots + a_h \langle v, v_k \rangle = 0,$$

che dimostra che v è ortogonale a w. Siccome $w \in E$ è arbitrario, $v \in E^\perp$. Il viceversa è immediato. $\square$

Esempio Determinare una base di $E^\perp$, complemento ortogonale del sottospazio E di $\mathbf{R}^4$ generato dai vettori $v_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, v_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}$.

Soluzione. Imponiamo al vettore generico $v = (x, y, z, w)^t \in \mathbf{R}^4$ l'ortogonalità ai vettori della base (v_1, v_2) di E, ottenendo il sistema omogeneo:

$$\begin{cases} x + y + z + w = 0 \\ x + z = 0 \end{cases}.$$

Risolvendo il sistema, otteniamo la base di $E^\perp$: $\begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ -1 \end{pmatrix}$. $\square$

Le proprietà importanti del complemento ortogonale sono espresse nel seguente teorema.

Teorema *Sia E un sottospazio di $\mathbf{R}^n$ e $E^\perp$ il suo complemento ortogonale. Allora*
a) $E \cap E^\perp = \{O\}$.
b) $\dim E^\perp = n - \dim E$.
c) $\mathbf{R}^n = E \oplus E^\perp$.

Dimostrazione. a) Se $v \in E \cap E^\perp$ allora v è ortogonale a tutti i vettori di E; in particolare v è ortogonale a sé stesso, e dunque $\langle v, v \rangle = 0$. Ma l'unico vettore con tale proprietà è il vettore nullo.
b) Sia $\dim E = k$ e sia $(u_1, \dots, u_k)$ una base di E. Sappiamo che $v \in E^\perp$ se e solo se v è ortogonale ai vettori di una base di E: dunque

$$E^\perp = \{v \in \mathbf{R}^n : \langle v, u_1 \rangle = \dots = \langle v, u_k \rangle = 0\}.$$

Siccome $u_1, \dots, u_k$ sono linearmente indipendenti, dalla proposizione del paragrafo precedente otteniamo che $\dim E^\perp = n - k = n - \dim E$.
c) Applichiamo la formula di Grassmann ai sottospazi $E, E^\perp$:

$$\dim(E + E^\perp) + \dim(E \cap E^\perp) = \dim E + \dim E^\perp.$$

Da a) e b) concludiamo che $\dim(E+E^\perp) = n$. Dunque $E+E^\perp = \mathbf{R}^n$. Poiché $E \cap E^\perp = \{O\}$ la somma è diretta: $\mathbf{R}^n = E \oplus E^\perp$. $\square$

5.2 Proiezione ortogonale su un sottospazio

Dal teorema precedente abbiamo che un vettore $v \in \mathbf{R}^n$ si spezza, in modo unico, come somma di un vettore $w \in E$ e di un vettore $w^\perp \in E^\perp$:

$$v = w + w^\perp.$$

In particolare, w e $w^\perp$ sono ortogonali.

- Il vettore w è detto la *proiezione ortogonale* di v sul sottospazio E. Denoteremo w con il simbolo $P_E(v)$.

Esempio È dato il sottospazio di $\mathbf{R}^3$ descritto dall'equazione $x + y - 2z = 0$. Il vettore $v = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} \in \mathbf{R}^3$ si spezza

$$\begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 1/2 \\ 3/2 \\ 1 \end{pmatrix} + \begin{pmatrix} 1/2 \\ 1/2 \\ -1 \end{pmatrix},$$

dove il primo vettore appartiene a E e il secondo a $E^\perp$. Quindi $P_E(v) = \begin{pmatrix} 1/2 \\ 3/2 \\ 1 \end{pmatrix}$.

In generale, se $(u_1 \dots, u_k)$ è una base *ortonormale* di E allora la proiezione ortogonale si calcola con la formula

$$P_E(v) = \langle v, u_1 \rangle u_1 + \cdots + \langle v, u_k \rangle u_k.$$

6 Endomorfismi simmetrici

In questa sezione studieremo una classe importante di endomorfismi di $\mathbf{R}^n$: gli endomorfismi detti *simmetrici.* Tali endomorfismi sono caratterizzati dalla proprietà di ammettere una base *ortonormale* di autovettori, e sono legati in modo naturale alle matrici simmetriche. In particolare, risulterà che ogni matrice simmetrica è diagonalizzabile.

Definizione *Un endomorfismo di $\mathbf{R}^n$ si dice* simmetrico *se la sua matrice associata rispetto alla base canonica è simmetrica.*

Esempio Sia $f : \mathbf{R}^2 \to \mathbf{R}^2$ definito da $f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -3x + 3y \\ 3x + 5y \end{pmatrix}$. La matrice canonica di f è:

$$A = \begin{pmatrix} -3 & 3 \\ 3 & 5 \end{pmatrix}.$$

Siccome A è simmetrica, f è simmetrico.

Esempio L'endomorfismo $f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x + 2y \\ 3y \end{pmatrix}$ ha matrice canonica $\begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}$ dunque non è simmetrico.

Teorema *Le seguenti condizioni sono equivalenti.*

a) *f è un endomorfismo simmetrico di $\mathbf{R}^n$.*

b) *La matrice associata a f rispetto ad una qualunque base ortonormale di $\mathbf{R}^n$ è simmetrica.*

c) *Per ogni coppia di vettori $u, v \in \mathbf{R}^n$ si ha*

$$\langle f(u), v\rangle = \langle u, f(v)\rangle.$$

Dimostrazione. a) $\Longrightarrow$ b) Supponiamo che f sia simmetrico, e sia A la sua matrice canonica. Per definizione, A è simmetrica. Se $\mathcal{B}$ è una base ortonormale di $\mathbf{R}^n$, e A' è la matrice associata a f rispetto a tale base, allora sappiamo che

$$A' = M^{-1}AM,$$

dove M è la matrice di passaggio dalla base canonica $\mathcal{BC}$ alla base $\mathcal{B}$. Poichè tali basi sono entrambe ortonormali, si ha che M è ortogonale, quindi $M^{-1} = M^t$. Dunque $A' = M^tAM$, ed è sufficiente dimostrare che M^tAM è simmetrica. Ma questo è immediato:

$$(M^tAM)^t = M^tA(M^t)^t = M^tAM.$$

b) $\Longrightarrow$ c) Si ha la seguente identità, valida per ogni matrice A e per ogni scelta di u, v, vettori colonna di $\mathbf{R}^n$:

$$\langle Au, v\rangle = \langle u, A^tv\rangle.$$

L'identità si verifica con un calcolo diretto, e fornisce un legame tra il prodotto scalare e la trasposta di una matrice. Supponiamo che la matrice A, associata ad f rispetto alla base canonica, sia simmetrica. Ora sappiamo che f si scrive

$$f(v) = Av.$$

Poichè A è simmetrica, si ha $A = A^t$ e dall'identità precedente:

$$\langle f(u), v\rangle = \langle Au, v\rangle = \langle u, Av\rangle = \langle u, f(v)\rangle.$$

c) $\Longrightarrow$ a) Premettiamo che, se A è una matrice $n \times n$ e $e_1, \dots, e_n$ sono i vettori della base canonica di $\mathbf{R}^n$, un calcolo mostra che

$$\langle Ae_i, e_j\rangle = a_{ji},$$

dove a_{ji} è l'elemento di posto (j, i) della matrice A.

Per ipotesi, si ha la proprietà c). Dunque, se A è la matrice canonica di f, l'identità

$$\langle Au, v\rangle = \langle u, Av\rangle$$

risulta vera per ogni scelta dei vettori colonna u, v. Prendendo $u = e_i$ e $v = e_j$ otteniamo

$$a_{ji} = a_{ij}$$

per ogni i, j, dunque la matrice canonica di f è simmetrica e f risulta simmetrico. $\square$

Isoliamo la seguente proprietà degli autospazi di un endomorfismo simmetrico.

Proposizione *Gli autospazi di un endomorfismo simmetrico sono ortogonali fra loro. In altre parole, se λ_1 e λ_2 sono autovalori distinti di f, e se $u \in E(\lambda_1)$ e $v \in E(\lambda_2)$ allora:*

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

Dimostrazione. Per ipotesi $f(u) = \lambda_1 u$; dunque

$$\langle f(u), v\rangle = \lambda_1\langle u, v\rangle.$$

D'altra parte, per la c) del teorema, poiché $f(v) = \lambda_2 v$:

$$\langle f(u), v\rangle = \langle u, f(v)\rangle = \lambda_2\langle u, v\rangle.$$

Uguagliando otteniamo $\lambda_1\langle u, v\rangle = \lambda_2\langle u, v\rangle$ cioè

$$(\lambda_1 - \lambda_2)\langle u, v\rangle = 0,$$

e poiché $\lambda_1 - \lambda_2 \neq 0$ si ha necessariamente $\langle u, v\rangle = 0$. $\square$

Esempio Sia $f : \mathbf{R}^2 \to \mathbf{R}^2$ definito da $f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -3x + 3y \\ 3x + 5y \end{pmatrix}$. Verifichiamo che gli autospazi di f sono ortogonali. Matrice canonica $A = \begin{pmatrix} -3 & 3 \\ 3 & 5 \end{pmatrix}$ con polinomio caratteristico $x^2 - 2x - 24$. e abbiamo due autovalori distinti: $\lambda_1 = -4$ e $\lambda_2 = 6$ e due autospazi $E(-4), E(6)$, entrambi di dimensione 1. Si trova che $E(-4)$ ha equazione $x + 3y = 0$ con base $\begin{pmatrix} 3 \\ -1 \end{pmatrix}$, e $E(6)$ ha equazione $3x - y = 0$ con base $\begin{pmatrix} 1 \\ 3 \end{pmatrix}$.

Effettivamente, gli autospazi sono ortogonali tra loro, la coppia $\begin{pmatrix} 3 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \end{pmatrix}$ è una base ortogonale di autovettori, e la coppia

$$\frac{1}{\sqrt{10}}\begin{pmatrix} 3 \\ -1 \end{pmatrix}, \frac{1}{\sqrt{10}}\begin{pmatrix} 1 \\ 3 \end{pmatrix}$$

è una base ortonormale di $\mathbf{R}^2$ formata da autovettori di f. $\square$

Esempio L'endomorfismo $f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x+2y \\ 3y \end{pmatrix}$ ha matrice canonica $\begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}$ dunque non è simmetrico. Si osserva che f ha autovalori $\lambda_1 = 1, \lambda_2 = 3$ e autospazi:

$$E(1) : y = 0, \quad E(3) : x - y = 0.$$

Si vede subito che gli autospazi non sono ortogonali. Risulta che f è diagonalizzabile, con base di autovettori $\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ ma *non è possibile* trovare una base ortonormale di autovettori (se ortonormalizziamo la base, non otteniamo piú autovettori).

Esempio Sia E un sottospazio di $\mathbf{R}^n$. L'endomorfismo P_E che associa al vettore $v \in \mathbf{R}^n$ la sua proiezione ortogonale sul sottospazio E è simmetrico.
Infatti, se fissiamo una base ortonormale $(u_1 \dots, u_k)$ di E allora la proiezione ortogonale è data da

$$P_E(v) = \langle v, u_1 \rangle u_1 + \dots + \langle v, u_k \rangle u_k.$$

Se w è un secondo vettore di $\mathbf{R}^n$ si ha

$$\langle P_E(v), w \rangle = \langle v, u_1 \rangle \langle u_1, w \rangle + \dots + \langle v, u_k \rangle \langle u_k, w \rangle.$$

Poiché il secondo membro rimane uguale scambiando v con w, si ha $\langle P_E(v), w \rangle = \langle P_E(w), v \rangle = \langle v, P_E(w) \rangle$ e P_E è simmetrico.

7 Teorema spettrale

Veniamo al seguente importante teorema, di cui omettiamo la dimostrazione.

Teorema spettrale. *Sia f un endomorfismo simmetrico di $\mathbf{R}^n$. Allora f è diagonalizzabile; inoltre esiste una base* ortonormale *di $\mathbf{R}^n$ costituita da autovettori di f.*

Anche il viceversa è vero, ed è facile da dimostrare:

Teorema *Sia f un endomorfismo di $\mathbf{R}^n$, e supponiamo che esista una base ortonormale di $\mathbf{R}^n$ formata da autovettori di f. Allora f è simmetrico.*

Dimostrazione. La matrice associata alla base di autovettori (che è ortonormale per ipotesi) è diagonale, dunque simmetrica, e quindi f è simmetrico per il teorema della sezione precedente. $\square$

Dunque, la classe degli endomorfismi di $\mathbf{R}^n$ che ammettono una base ortonormale di autovettori coincide con la classe degli endomorfismi simmetrici.
Notiamo anche il fatto seguente.

Corollario *Ogni matrice simmetrica è diagonalizzabile, ed è ortogonalmente simile ad una matrice diagonale. Cioè, possiamo trovare una matrice diagonale D e una matrice ortogonale M tali che:*

$$D = M^{-1}AM = M^tAM.$$

Dimostrazione. Sia f l'endomorfismo di $\mathbf{R}^n$ rappresentato da A rispetto alla base canonica. Poiche' A è simmetrica, anche f è simmetrico. Per il teorema spettrale, possiamo trovare una base ortonormale $\mathcal{B}$ formata da autovettori di f. In questa base, f si rappresenta con una matrice diagonale D; inoltre si ha

$$D = M^{-1}AM,$$

dove M è la matrice di passaggio dalla base canonica alla base $\mathcal{B}$. Poiche' tali basi sono entrambe ortonormali, la matrice M è ortogonale, quindi $M^{-1} = M^t$. $\square$

Diamo ora il procedimento per determinare una base ortonormale di autovettori di un endomorfismo simmetrico.

1. Calcoliamo il polinomio caratteristico e quindi troviamo gli autovalori di f, diciamo $\lambda_1, \dots, \lambda_k$.

2. Con l'algoritmo di Gram-Schmidt, troviamo una base *ortonormale* di ciascun autospazio, diciamo $\mathcal{B}_1, \dots, \mathcal{B}_k$.

3. Uniamo le basi ortonormali cosi' trovate per ottenere la base $\mathcal{B} = \mathcal{B}_1 \cup \dots \cup \mathcal{B}_k$ di $\mathbf{R}^n$. L'insieme di vettori cosi' ottenuto formera' una base ortonormale di autovettori.

Infatti, ogni vettore di $\mathcal{B}$ ha chiaramente norma 1. Inoltre, se prendiamo due vettori appartenenti alla stessa base $\mathcal{B}_i$ questi sono ortogonali per costruzione; se prendiamo due vettori appartenenti a basi diverse, questi appartengono ad autospazi diversi e quindi sono ortogonali grazie alla proposizione della sezione precedente. I vettori di $\mathcal{B}$ sono a due a due ortogonali e di norma 1, dunque $\mathcal{B}$ è una base ortonormale.

Infine, per diagonalizzare una matrice simmetrica A, procediamo cosi':

1. Troviamo una base ortonormale $\mathcal{B}$ formata da autovettori dell'endomorfismo di $\mathbf{R}^n$ rappresentato da A rispetto alla base canonica.

2. Incolonniamo la base $\mathcal{B}$ per ottenere una matrice ortogonale M.

3. Scriviamo la matrice diagonale D, i cui elementi diagonali sono gli autovalori di f, presi nello stesso ordine dei corrispondenti autovettori di $\mathcal{B}$.

4. Risultera' allora $D = M^tAM$.

7.1 Esempio

Sia f l'operatore di $\mathbf{R}^2$ rappresentato da $\begin{pmatrix} -3 & 3 \\ 3 & 5 \end{pmatrix}$ rispetto alla base canonica. Abbiamo già trovato una base ortonormale di autovettori:

$$\frac{1}{\sqrt{10}} \begin{pmatrix} 3 \\ -1 \end{pmatrix}, \frac{1}{\sqrt{10}} \begin{pmatrix} 1 \\ 3 \end{pmatrix},$$

associati rispettivamente a -4 e 6. Quindi se prendiamo

$$M = \frac{1}{\sqrt{10}} \begin{pmatrix} 3 & 1 \\ -1 & 3 \end{pmatrix}, \quad D = \begin{pmatrix} -4 & 0 \\ 0 & 6 \end{pmatrix}.$$

si avrà $D = M^t A M$.

7.2 Esempio

Sia f l'operatore di $\mathbf{R}^3$ rappresentato da $\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$ rispetto alla base canonica. f è simmetrico. Il polinomio caratteristico è $-x^3 + 3x^2$ e gli autovalori sono $0, 3$. $E(0)$ è il nucleo, di equazione $x + y + z = 0$ e base $\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$. Applicando l'algoritmo di Gram-Schmidt, otteniamo la base ortonormale di $E(0)$:

$$w_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, w_2 = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}.$$

$E(3)$ ha base $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$; si osserva che $E(3)$ è ortogonale a $E(0)$. Otteniamo la base ortonormale di $E(3)$:

$$w_3 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Ne segue che una base ortonormale di autovettori è (w_1, w_2, w_3) cioè:

$$\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}, \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Ponendo

$$M = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & -\frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{pmatrix}, \quad D = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 3 \end{pmatrix},$$

si ha $D = M^t AM$.

7.3 Esempio

Sia f l'endomorfismo di $\mathbf{R}^4$ rappresentato, rispetto alla base canonica, dalla matrice

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

Poiché A è simmetrica, f è un endomorfismo simmetrico. Un calcolo mostra che $p_A(x) = (x-1)^2(x-3)^2$, dunque f ammette due autovalori distinti: $\lambda_1 = 1$ e $\lambda_2 = 3$, entrambi di molteplicità algebrica 2. Già sappiamo che f è diagonalizzabile, dunque la molteplicità geometrica di entrambi gli autovalori sarà 2.

Descriviamo gli autospazi. $A - I = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & -1 \\ 1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{pmatrix}$ dunque $E(1)$ ha equazioni:

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

e quindi

$$E(1) = \{ \begin{pmatrix} t \\ s \\ -t \\ s \end{pmatrix} \in \mathbf{R}^4 : t, s \in \mathbf{R} \}.$$

Procedendo in modo analogo, si ha:

$$E(3) = \{ \begin{pmatrix} t' \\ s' \\ t' \\ -s' \end{pmatrix} \in \mathbf{R}^4 : t', s' \in \mathbf{R} \}.$$

Osserviamo che i due autospazi sono fra loro ortogonali, nel senso che:

$$\langle \begin{pmatrix} t \\ s \\ -t \\ s \end{pmatrix}, \begin{pmatrix} t' \\ s' \\ t' \\ -s' \end{pmatrix} \rangle = tt' + ss' - tt' - ss' = 0$$

per ogni $t, s, t', s' \in \mathbf{R}$.
Passiamo ora a costruire una base ortonormale di autovettori di f. Una base di $E(1)$ è data dalla coppia $((1, 0, -1, 0)^t, (0, 1, 0, 1)^t)$: i due vettori sono ortogonali, dunque una base ortonormale di $V(1)$ è:

$$w_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix}, \quad w_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix},$$

In modo analogo, dalla base $((1, 0, 1, 0)^t, (0, 1, 0, -1)^t)$ di $E(3)$ otteniamo la base ortonormale di $E(3)$:

$$w_3 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad w_4 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ 0 \\ -1 \end{pmatrix}$$

Dunque

$$\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2 = (w_1, w_2, w_3, w_4)$$

è una base ortonormale di $\mathbf{R}^4$, costituita da autovettori di f.
La matrice A è diagonalizzabile; se M è la matrice ottenuta incolonnando la base ortonormale di autovettori descritta in precedenza, cioè

$$M = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ -1 & 0 & 1 & 0 \\ 0 & 1 & 0 & -1 \end{pmatrix}$$

allora M è una matrice *ortogonale* che diagonalizza A, nel senso che

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