Capitolo 7. Forme quadratiche[1].

1. Forme bilineari simmetriche.

Fissiamo uno spazio vettoriale V di dimensione finita. In tutto questo capitolo supporremo che gli scalari siano in $\mathbf{R}$. Denotiamo con $V \times V$ l'insieme delle coppie di vettori di V, cioe' l'insieme dei simboli del tipo $(\mathbf{u}, \mathbf{v})$, con $\mathbf{u}$ e $\mathbf{v}$ vettori di V. Ricordiamo che due coppie $(\mathbf{u}, \mathbf{v})$ e $(\mathbf{u}', \mathbf{v}')$ sono *uguali* se e sole se per definizione $\mathbf{u} = \mathbf{u}'$ e $\mathbf{v} = \mathbf{v}'$.

Una *forma bilineare simmetrica su* V e' una funzione del tipo

$$\phi : V \times V \to \mathbf{R}$$

soddisfacente le seguenti due proprieta':

- *per ogni* $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ *ed ogni* $a, b \in \mathbf{R}$ *si ha* $\phi(a\mathbf{u}+b\mathbf{v}, \mathbf{w}) = a\phi(\mathbf{u}, \mathbf{w})+b\phi(\mathbf{v}, \mathbf{w})$;
- *per ogni* $\mathbf{u}, \mathbf{v} \in V$ *si ha* $\phi(\mathbf{u}, \mathbf{v}) = \phi(\mathbf{v}, \mathbf{u})$.

La prima proprieta' si chiama *proprieta' di linearita' a sinistra*, mentre la seconda si chiama *proprieta' di simmetria*. Il termine *bilineare* deriva dal fatto che se valgono queste due proprieta' allora ϕ e' anche *lineare a destra*, cioe' e' vero anche che $\phi(\mathbf{u}, a\mathbf{v} + b\mathbf{w}) = a\phi(\mathbf{u}, \mathbf{v}) + b\phi(\mathbf{u}, \mathbf{w})$ per ogni $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ ed ogni $a, b \in \mathbf{R}$. Infatti possiamo scrivere:

$$\phi(\mathbf{u}, a\mathbf{v} + b\mathbf{w}) = \phi(a\mathbf{v} + b\mathbf{w}, \mathbf{u}) = a\phi(\mathbf{v}, \mathbf{u}) + b\phi(\mathbf{w}, \mathbf{u}) = a\phi(\mathbf{u}, \mathbf{v}) + b\phi(\mathbf{u}, \mathbf{w}).$$

La prima uguaglianza e' consentita dalla simmetria, la seconda dalla linearita' a sinistra, e l'ultima ancora dalla simmetria.

Esempio 1. *Consideriamo la seguente funzione* $\phi : \mathbf{R}^2 \times \mathbf{R}^2 \to \mathbf{R}$ *(quindi in questo caso* $V = \mathbf{R}^2$*) definita ponendo per ogni* $\mathbf{u} = (x_1, x_2), \mathbf{v} = (y_1, y_2) \in \mathbf{R}^2$

$$\phi(\mathbf{u}, \mathbf{v}) := x_1y_1 + 3x_1y_2 + 3x_2y_1 + 2x_2y_2.$$

Ad esempio $\phi(\mathbf{e}_1, \mathbf{e}_1) = 1$, $\phi(\mathbf{e}_1, \mathbf{e}_2) = 3$, $\phi(\mathbf{e}_2, \mathbf{e}_1) = 3$, $\phi(\mathbf{e}_2, \mathbf{e}_2) = 2$. *Ora andiamo a provare che questa funzione e' una forma bilineare simmetrica. Faremo due dimostrazioni diverse. La prima consiste in una verifica diretta. Poi useremo un argomento diverso, che ci sara' utile per motivare le definizioni che seguiranno.*

Cominciamo con una verifica diretta. Posto $\mathbf{w} = (z_1, z_2)$, *abbiamo:*

$$\phi(a\mathbf{u} + b\mathbf{v}, \mathbf{w}) = \phi((ax_1 + by_1, ax_2 + by_2), (z_1, z_2))$$

$$= (ax_1 + by_1)z_1 + 3(ax_1 + by_1)z_2 + 3(ax_2 + by_2)z_1 + 2(ax_2 + by_2)z_2$$

$$= ax_1z_1 + by_1z_1 + 3ax_1z_2 + 3by_1z_2 + 3ax_2z_1 + 3by_2z_1 + 2ax_2z_2 + 2by_2z_2$$

$$= a(x_1z_1+3x_1z_2+3x_2z_1+2x_2z_2)+b(y_1z_1+3y_1z_2+3y_2z_1+2y_2z_2) = a\phi(\mathbf{u}, \mathbf{w})+b\phi(\mathbf{v}, \mathbf{w}).$$

[1] *Ultimo aggiornamento: 9 gennaio 2013*

Cio' prova la linearita' a sinistra. Poi abbiamo

$$\phi(\mathbf{u},\mathbf{v}) = x_1y_1 + 3x_1y_2 + 3x_2y_1 + 2x_2y_2 = y_1x_1 + 3y_2x_1 + 3y_1x_2 + 2y_2x_2$$

$$= y_1x_1 + 3y_1x_2 + 3y_2x_1 + 2y_2x_2 = \phi(\mathbf{v},\mathbf{u}),$$

e cio' prova che ϕ *e' anche simmetrica. In conclusione abbiamo provato che* ϕ *e' una forma bilineare simmetrica.*

Ora invece dimostreremo che ϕ *e' una forma bilineare simmetrica con un argomento che utilizza le proprieta' di calcolo delle matrici. Punto di partenza consiste nel riscrivere* $\phi(\mathbf{u},\mathbf{v})$ *nel seguente modo:*

$$\phi(\mathbf{u},\mathbf{v}) = x_1y_1 + 3x_1y_2 + 3x_2y_1 + 2x_2y_2 = x_1(y_1 + 3y_2) + x_2(3y_1 + 2y_2)$$

$$= (x_1, x_2) \cdot \begin{pmatrix} y_1 + 3y_2 \\ 3y_1 + 2y_2 \end{pmatrix} = (x_1, x_2) \cdot \begin{bmatrix} 1 & 3 \\ 3 & 2 \end{bmatrix} \cdot \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \mathbf{x}^T \begin{bmatrix} 1 & 3 \\ 3 & 2 \end{bmatrix} \mathbf{y},$$

dove $\mathbf{x} = [\mathbf{u}]_{\mathcal{E}}$ *ed* $\mathbf{y} = [\mathbf{v}]_{\mathcal{E}}$ *denotano le coordinate di* $\mathbf{u}$ *e* $\mathbf{v}$ *rispetto alla base canonica* $\mathcal{E}$*. Cioe' abbiamo*

$$\phi(\mathbf{u},\mathbf{v}) = \mathbf{x}^T \begin{bmatrix} 1 & 3 \\ 3 & 2 \end{bmatrix} \mathbf{y} = [\mathbf{u}]_{\mathcal{E}}^T \begin{bmatrix} 1 & 3 \\ 3 & 2 \end{bmatrix} [\mathbf{v}]_{\mathcal{E}}.$$

Si osservi che la matrice $\begin{bmatrix} 1 & 3 \\ 3 & 2 \end{bmatrix}$ *altro non e' che la matrice*

$$\begin{bmatrix} 1 & 3 \\ 3 & 2 \end{bmatrix} = \begin{bmatrix} \phi(\mathbf{e}_1,\mathbf{e}_1) & \phi(\mathbf{e}_1,\mathbf{e}_2) \\ \phi(\mathbf{e}_2,\mathbf{e}_1) & \phi(\mathbf{e}_2,\mathbf{e}_2) \end{bmatrix}.$$

Tale matrice usualmente si denota con la lettera

$$G := \begin{bmatrix} 1 & 3 \\ 3 & 2 \end{bmatrix}$$

e prende il nome di matrice di Gram di ϕ *relativa alla base canonica. E' una sorta di matrice rappresentativa. Una volta che abbiamo riscritto nel modo precedente la funzione* ϕ *possiamo dimostrare che e' bilineare e simmetrica ricorrendo alle proprieta' di calcolo delle matrici. Infatti per la proprieta' distributiva del prodotto righe per colonna, e per le proprieta' della matrice trasposta, abbiamo:*

$$\phi(a\mathbf{u} + b\mathbf{v},\mathbf{w}) = (a\mathbf{x} + b\mathbf{y})^T \begin{bmatrix} 1 & 3 \\ 3 & 2 \end{bmatrix} \mathbf{z} = (a\mathbf{x}^T + b\mathbf{y}^T) \begin{bmatrix} 1 & 3 \\ 3 & 2 \end{bmatrix} \mathbf{z}$$

$$= a(\mathbf{x}^T \begin{bmatrix} 1 & 3 \\ 3 & 2 \end{bmatrix} \mathbf{z}) + b(\mathbf{y}^T \begin{bmatrix} 1 & 3 \\ 3 & 2 \end{bmatrix} \mathbf{z}) = a\phi(\mathbf{u},\mathbf{w}) + b\phi(\mathbf{v},\mathbf{w}).$$

Cio' prova la linearita' a sinistra. Poi abbiamo

$$\phi(\mathbf{u},\mathbf{v}) = \mathbf{x}^T \begin{bmatrix} 1 & 3 \\ 3 & 2 \end{bmatrix} \mathbf{y} = \left(\mathbf{x}^T \begin{bmatrix} 1 & 3 \\ 3 & 2 \end{bmatrix} \mathbf{y} \right)^T$$

$$= \mathbf{y}^T \begin{bmatrix} 1 & 3 \\ 3 & 2 \end{bmatrix}^T \mathbf{x} = \mathbf{y}^T \begin{bmatrix} 1 & 3 \\ 3 & 2 \end{bmatrix} \mathbf{x} = \phi(\mathbf{v},\mathbf{u}),$$

e cio' prova anche la simmetria. Si osservi che nel calcolo precedente la seconda uguaglianza e' lecita in quanto $\mathbf{x}^T \begin{bmatrix} 1 & 3 \\ 3 & 2 \end{bmatrix} \mathbf{y}$ *e' un numero, cioe' una matrice* 1×1 *che e' simmetrica e quindi coincide con la propria trasposta, la terza uguaglianza e' una conseguenza della proprieta' secondo cui* $(AB)^T = B^T A^T$*, e la quarta e' dovuta al fatto che la matrice* $\begin{bmatrix} 1 & 3 \\ 3 & 2 \end{bmatrix}$ *e' simmetrica.* ■

Esempio 2. *(i) La funzione identicamente nulla* $\phi(\mathbf{u},\mathbf{v}) \equiv 0$ *e' una forma bilineare simmetrica;*

(ii) La funzione $\phi : \mathbf{R}^2 \times \mathbf{R}^2 \to \mathbf{R}$ *definita ponendo per ogni* $\mathbf{u} = (x_1, x_2), \mathbf{v} = (y_1, y_2) \in \mathbf{R}^2$

$$\phi(\mathbf{u},\mathbf{v}) := x_1y_1 + 3x_1y_2 - 5x_2y_1 + 2x_2y_2$$

non e' una forma bilineare simmetrica. Infatti $\phi(\mathbf{e}_1,\mathbf{e}_2) = 3$ *mentre* $\phi(\mathbf{e}_2,\mathbf{e}_1) = -5$*. Quindi viene meno la simmetria. Osserviamo tuttavia che e' bilineare. Cio' si puo' provare come nell'esempio precedente, osservando che*

$$\phi(\mathbf{u},\mathbf{v}) = \mathbf{x}^T \begin{bmatrix} 1 & 3 \\ -5 & 2 \end{bmatrix} \mathbf{y}. \quad ■$$

Esempio 3. *Un esempio di grande importanza e' il prodotto punto. Cioe' la funzione* $\phi : \mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}$ *definita ponendo per ogni* $\mathbf{u} = (x_1, x_2, \dots, x_n), \mathbf{v} = (y_1, y_2, \dots, y_n) \in \mathbf{R}^n$

$$\phi(\mathbf{u},\mathbf{v}) := x_1y_1 + x_2y_2 + \cdots + x_ny_n = \mathbf{x}^T \cdot \mathbf{y}$$

e' una forma bilineare simmetrica. In questo caso la matrice G *e' la matrice unitaria* I*.* ■

Esempio 4. *Un altro esempio importante e' il prodotto scalare tra vettori geometrici. Denotiamo con* $\mathcal{V}_O$ *lo spazio dei vettori geometrici applicati in un fissato punto* O *dello spazio. Siano* $\mathbf{u} := \overrightarrow{OP}$*,* $\mathbf{v} := \overrightarrow{OQ}$ *vettori in* $\mathcal{V}_O$*. Si definisce prodotto scalare geometrico tra il vettore* $\mathbf{u}$ *ed il vettore* $\mathbf{v}$ *il numero:*

$$\phi(\mathbf{u},\mathbf{v}) := \|\mathbf{u}\| \cdot \|\mathbf{v}\| \cdot \cos \widehat{\mathbf{uv}}$$

dove $\|\mathbf{u}\|$ *e* $\|\mathbf{v}\|$ *denotano la lunghezza di* $\mathbf{u}$ *e* $\mathbf{v}$*, e* $\cos\widehat{\mathbf{uv}}$ *il coseno dell'angolo* $\widehat{\mathbf{uv}}$ *formato dai vettori* $\mathbf{u}$ *e* $\mathbf{v}$*. Per provare che il prodotto scalare e' una forma bilineare simmetrica si puo' ragionare come segue.*

Innanzitutto ricordiamo che, in base alla Regola del coseno applicata al triangolo O, P, Q*, si ha*

$$\|\mathbf{u}-\mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - 2\|\mathbf{u}\|\|\mathbf{v}\|\cos\widehat{\mathbf{uv}},$$

cioe'

$$(*) \qquad \|\mathbf{u}\| \cdot \|\mathbf{v}\| \cdot \cos\widehat{\mathbf{uv}} = \frac{1}{2}\left(\|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - \|\mathbf{u}-\mathbf{v}\|^2\right).$$

Poi introduciamo delle coordinate fissando una base $\mathbf{e}_1$*,* $\mathbf{e}_2$*,* $\mathbf{e}_3$ *di* $\mathcal{V}_O$ *formata da tre vettori di lunghezza* 1 *mutualmente ortogonali tra loro. Denotate con* $\mathbf{x}$ *ed* $\mathbf{y}$ *le coordinate di* $\mathbf{u}$ *e* $\mathbf{v}$ *rispetto a tale base, dal Teorema di Pitagora deduciamo*

$$\|\mathbf{u}\| = \sqrt{x_1^2 + x_2^2 + x_3^2}, \quad \|\mathbf{v}\| = \sqrt{y_1^2 + y_2^2 + y_3^2}, \quad e$$

$$\|\mathbf{u}-\mathbf{v}\| = \sqrt{(x_1-y_1)^2 + (x_2-y_2)^2 + (x_3-y_3)^2}.$$

Sostituendo nella formula $(*)$ *abbiamo*

$$\phi(\mathbf{u},\mathbf{v}) := \|\mathbf{u}\| \cdot \|\mathbf{v}\| \cdot \cos\widehat{\mathbf{uv}} = x_1y_1 + x_2y_2 + x_3y_3 = \mathbf{x}^T \cdot \mathbf{y}.$$

Quindi utilizzando tali coordinate il prodotto scalare si identifica con il prodotto punto dell'esempio precedente. ■

La matrice G che abbiamo incontrato nell'esempio precedente puo' essere definita in generale. A tale proposito sia $\phi : V \times V \to \mathbf{R}$ una forma bilineare simmetrica, e fissiamo una base $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ di V. Si definisce *matrice di Gram di* ϕ *rispetto alla base* $\mathcal{B}$ la matrice

$$G_{\mathcal{B}}^{\mathcal{B}}(\phi) := (g_{ij}), \quad \text{dove} \quad g_{ij} := \phi(\mathbf{b}_i, \mathbf{b}_j).$$

Si osservi che la matrice $G_{\mathcal{B}}^{\mathcal{B}}(\phi)$ e' una matrice quadrata $n \times n$ e simmetrica. Valgono le seguenti proprieta':

Proposizione 1. *1) Per ogni* $\mathbf{u}, \mathbf{v} \in V$ *si ha*

$$\phi(\mathbf{u},\mathbf{v}) = [\mathbf{u}]_{\mathcal{B}}^T \cdot G_{\mathcal{B}}^{\mathcal{B}}(\phi) \cdot [\mathbf{v}]_{\mathcal{B}};$$

2) per ogni altra base $\mathcal{B}'$ *si ha*

$$G_{\mathcal{B}'}^{\mathcal{B}'}(\phi) = M_{\mathcal{B}}^{\mathcal{B}'}(id_V)^T \cdot G_{\mathcal{B}}^{\mathcal{B}}(\phi) \cdot M_{\mathcal{B}}^{\mathcal{B}'}(id_V).$$

Ponendo $\mathbf{x} = [\mathbf{u}]_{\mathcal{B}}$, $\mathbf{y} = [\mathbf{v}]_{\mathcal{B}}$, $P = M_{\mathcal{B}}^{\mathcal{B}'}(id_V)$, $G = G_{\mathcal{B}}^{\mathcal{B}}(\phi)$ e $G' = G_{\mathcal{B}'}^{\mathcal{B}'}(\phi)$, le proprieta' precedenti si possono anche scrivere piu' semplicemente sotto la forma:

$$\phi(\mathbf{u}, \mathbf{v}) = \mathbf{x}^T \cdot G \cdot \mathbf{y}, \quad \text{e} \quad G' = P^T G P.$$

Esempio 5. *Consideriamo di nuovo la forma bilineare dell'Esempio 1, cioe'* $\phi : \mathbf{R}^2 \times \mathbf{R}^2 \to \mathbf{R}$ *definita ponendo per ogni* $\mathbf{u} = (x_1, x_2), \mathbf{v} = (y_1, y_2) \in \mathbf{R}^2$

$$\phi(\mathbf{u}, \mathbf{v}) := x_1 y_1 + 3x_1 y_2 + 3x_2 y_1 + 2x_2 y_2.$$

Denotiamo con $\mathcal{E}$ *la base canonica di* $\mathbf{R}^2$*, e con* $\mathcal{B}$ *la base formata dai vettori* $(1,1), (1,2)$*. Allora si ha*

$$G_{\mathcal{E}}^{\mathcal{E}}(\phi) = \begin{bmatrix} 1 & 3 \\ 3 & 2 \end{bmatrix}, \quad e \quad G_{\mathcal{B}}^{\mathcal{B}}(\phi) = \begin{bmatrix} 9 & 14 \\ 14 & 21 \end{bmatrix}.$$

Osserviamo che se si adoperano le coordinate $\mathbf{x}'$ *rispetto alla base* $\mathcal{B}$ *allora possiamo scrivere*

$$\phi(\mathbf{u}, \mathbf{v}) := 9x_1' y_1' + 14x_1' y_2' + 14x_2' y_1' + 21x_2' y_2'.$$

Supponiamo di voler calcolare

$$\phi((1,1),(2,0)).$$

Se usiamo le coordinate canoniche allora possiamo calcolare tale numero cosi':

$$\phi((1,1),(2,0)) = [1\,1] \begin{bmatrix} 1 & 3 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \end{bmatrix} = 8.$$

Se invece vogliamo usare le coordinate rispetto alla base $\mathcal{B}$ *allora, tenuto conto che* $[(1,1)]_{\mathcal{B}} = (1,0)^T$ *e che* $[(2,0)]_{\mathcal{B}} = (4,-2)^T$*, possiamo fare il calcolo nel modo seguente:*

$$\phi((1,1),(2,0)) = [1\,0] \begin{bmatrix} 9 & 14 \\ 14 & 21 \end{bmatrix} \begin{bmatrix} 4 \\ -2 \end{bmatrix} = 8.$$

Chiaramente il valore $\phi((1,1),(2,0))$ *non puo' dipendere dalla base scelta.* ■

Osservazione 1. *(i) La seconda proprieta' nella Proposizione precedente motiva la seguente definizione: date due matrici simmetriche reali* G *e* G'*, si dice che* G *e' congruente alla matrice* G' *se esiste una matrice invertibile* P *tale che*

$$G' = P^T G P.$$

La relazione di congruenza e' una relazione di equivalenza nell'insieme di tutte le matrici simmetriche di dato ordine n*. Inoltre si puo' provare che* G *e' congruente a* G' *se e solo se* G *e* G' *sono matrici di Gram di una stessa forma bilineare simmetrica* $\phi : \mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}$*, riferite a basi eventualmente distinte. Queste nozioni hanno una analogia con quelle che abbiamo gia' incontrato a proposito della relazione di similitudine. Ricordiamo che* *due matrici quadrate A ed A' dello stesso ordine n si dicono simili se esiste una matrice invertibile P tale che*

$$A' = P^{-1}AP.$$

La relazione di similitudine e' una relazione di equivalenza e due matrici A ed A' sono simili se e solo se rappresentano uno stesso operatore lineare $f : \mathbf{R}^n \to \mathbf{R}^n$, riferite a basi eventualmente diverse.

(ii) Abbiamo visto che, fissato uno spazio vettoriale V ed una base $\mathcal{B}$ di V, allora ogni forma bilineare simmetrica $\phi : V \times V \to \mathbf{R}$ determina una matrice simmetrica G, cioe' la matrice di Gram di ϕ associata alla base fissata $\mathcal{B}$. Viceversa, data una matrice simmetrica G qualsiasi (di ordine pari alla dimensione di V) possiamo costruire una forma bilineare simmetrica $\phi : V \times V \to \mathbf{R}$ ponendo per definizione $\phi(\mathbf{u}, \mathbf{v}) := \mathbf{x}^T G\mathbf{y}$, dove $\mathbf{x}$ ed $\mathbf{y}$ denotano le coordinate rispetto alla base fissata $\mathcal{B}$. In altre parole, una volta fissata una base di V, allora esiste una corrispondenza biiettiva tra l'insieme di tutte le forme bilineari simmetriche su V e l'insieme delle matrici simmetriche di ordine $\dim(V)$.

2. Forme quadratiche.

C'e' un altro modo per introdurre la nozione di forma bilineare simmetrica, e cioe' la nozione di forma quadratica. Si tratta di definizioni diverse ma che, in un senso che sara' piu' chiaro dopo, sono equivalenti.

Sia V uno spazio vettoriale reale. Una *forma quadratica su V* e' una funzione del tipo

$$q : V \to \mathbf{R}$$

soddisfacente le seguenti due proprieta':

- *per ogni $\mathbf{u} \in V$ ed ogni $c \in \mathbf{R}$ si ha* $q(c\mathbf{u}) = c^2 q(\mathbf{u})$;
- *la funzione $\phi_q : (\mathbf{u}, \mathbf{v}) \in V \times V \to \frac{1}{2}(q(\mathbf{u} + \mathbf{v}) - q(\mathbf{u}) - q(\mathbf{v})) \in \mathbf{R}$ e' una forma bilineare simmetrica su V (tale forma si dice che e' ottenuta polarizzando q).*

Esempio 6. *(i) La funzione identicamente nulla $q(\mathbf{u}) \equiv \mathbf{0}$ e' una forma quadratica.*

(ii) Consideriamo la funzione $q : \mathbf{R}^2 \to \mathbf{R}$ definita ponendo

$$q(\mathbf{u}) := x_1^2 + 6x_1x_2 + 2x_2^2.$$

Per provare che q e' una forma quadratica possiamo procedere con una verifica diretta come nell'Esempio 1, oppure osservare innanzitutto che

$$q(\mathbf{u}) = \mathbf{x}^T \begin{bmatrix} 1 & 3 \\ 3 & 2 \end{bmatrix} \mathbf{x}.$$

Dopodiche' la verifica che q e' una forma quadratica si puo' fare osservando che:

$$q(c\mathbf{u}) = (c\mathbf{x}^T) \begin{bmatrix} 1 & 3 \\ 3 & 2 \end{bmatrix} (c\mathbf{x}) = c^2 \left(\mathbf{x}^T \begin{bmatrix} 1 & 3 \\ 3 & 2 \end{bmatrix} \mathbf{x} \right) = c^2 q(\mathbf{u}).$$

Cio' prova la prima condizione richiesta. Poi abbiamo

$$\phi_q(\mathbf{u},\mathbf{v}) = \frac{1}{2}(q(\mathbf{u}+\mathbf{v}) - q(\mathbf{u}) - q(\mathbf{v}))$$

$$= \frac{1}{2}\left((\mathbf{x}+\mathbf{y})^T \begin{bmatrix} 1 & 3 \\ 3 & 2 \end{bmatrix} (\mathbf{x}+\mathbf{y}) - \mathbf{x}^T \begin{bmatrix} 1 & 3 \\ 3 & 2 \end{bmatrix} \mathbf{x} - \mathbf{y}^T \begin{bmatrix} 1 & 3 \\ 3 & 2 \end{bmatrix} \mathbf{y}\right) = \mathbf{x}^T \begin{bmatrix} 1 & 3 \\ 3 & 2 \end{bmatrix} \mathbf{y},$$

cioe'

$$\phi_q(\mathbf{u},\mathbf{v}) = \mathbf{x}^T \begin{bmatrix} 1 & 3 \\ 3 & 2 \end{bmatrix} \mathbf{y},$$

e sappiamo che tale funzione e' una forma bilineare simmetrica.

(iii) La funzione $q : \mathbf{R}^n \to \mathbf{R}$ definita ponendo

$$q(\mathbf{u}) := x_1^2 + x_2^2 + \cdots + x_n^2 = \mathbf{x}^T\mathbf{x}$$

e' una forma quadratica. In questo caso

$$\phi_q(\mathbf{u},\mathbf{v}) = \mathbf{x}^T\mathbf{y},$$

cioe' ϕ_q e' il prodotto punto. ■

Ora andiamo a vedere in che senso le definizioni di forma bilineare simmetrica e di forma quadratica sono equivalenti. Innanzitutto, per la stessa definizione, una forma quadratica q determina una forma bilineare simmetrica, quella che abbiamo denotato con ϕ_q. Viceversa data una forma bilineare simmetrica ϕ la funzione

$$q_\phi(\mathbf{u}) := \phi(\mathbf{u},\mathbf{u})$$

e' una forma quadratica. Le corrispondenze

$$q \to \phi_q \qquad \phi \to q_\phi$$

sono una l'inversa dell'altra e stabiliscono una corrispondenza biiettiva tra l'insieme delle forme quadratiche e l'insieme delle forme bilineari simmetriche. Da ora in poi quando assegneremo una forma quadratica q denoteremo con ϕ la forma bilineare simmetrica corrispondente. In particolare se e' assegnata una forma quadratica q e ϕ e' la forma bilineare simmetrica corrispondente a q, introdotte delle coordinate nello spazio V e denotata con G la matrice di Gram corrispondente a ϕ, si ha

$$\phi(\mathbf{u},\mathbf{v}) = \mathbf{x}^T G\mathbf{y} \qquad \text{e} \qquad q(\mathbf{u}) = \mathbf{x}^T G\mathbf{x}.$$

La matrice G si dice anche la matrice di Gram di q. Esplicitando l'espressione $\phi(\mathbf{u},\mathbf{v}) = \mathbf{x}^T G\mathbf{y}$ si ottiene (denotiamo con n la dimensione di V):

$$\phi(\mathbf{u},\mathbf{v}) = g_{11}x_1y_1 + g_{12}x_1y_2 + \cdots + g_{1n}x_1y_n$$

$$+g_{21}x_2y_1 + g_{22}x_2y_2 + \cdots + g_{2n}x_2y_n + \cdots + g_{n1}x_ny_1 + g_{n2}x_ny_2 + \cdots + g_{nn}x_ny_n$$
$$= \sum_{i,j=1}^{n} g_{ij}x_iy_j.$$

L'espressione precedente si chiama *l'espressione esplicita della forma bilineare ϕ in termini delle coordinate scelte.* L'espressione esplicita di ϕ ci presenta ϕ come un polinomio omogeneo di secondo grado nelle variabili x_i ed y_i, e tale polinomio e' bi-omogeneo nel senso che e' di primo grado nelle variabili x_i e di primo grado nelle variabili y_i. I coefficienti di tale polinomio sono le componenti (ordinate) della matrice di Gram. Invece esplicitando l'espressione $q(\mathbf{u}) = \mathbf{x}^T G\mathbf{x}$ si ottiene:

$$q(\mathbf{u}) = g_{11}x_1^2 + g_{12}x_1x_2 + \cdots + g_{1n}x_1x_n$$
$$+g_{21}x_2y_1 + g_{22}x_2^2 + \cdots + g_{2n}x_2x_n + \cdots + g_{n1}x_nx_1 + g_{n2}x_nx_2 + \cdots + g_{nn}x_n^2$$
$$= \sum_{i=1}^{n} g_{ii}x_i^2 + \sum_{1\leq i<j\leq n} 2g_{ij}x_ix_j.$$

L'espressione precedente si chiama *l'espressione esplicita della forma quadratica q in termini delle coordinate scelte.* L'espressione esplicita di q ci presenta q come un polinomio omogeneo di secondo grado nelle variabili x_i. In tale polinomio i monomi del tipo x_i^2 appaiono con coefficiente g_{ii} pari alla componente di posto i, i della matrice di Gram della forma bilineare associata ϕ, mentre i monomi del tipo x_ix_j (con $i \neq j$) appaiono con coefficiente $2g_{ij}$ pari al doppio della componente corrispondente nella matrice di Gram.

Esempio 7. *(i) Si consideri la seguente forma bilineare simmetrica su $\mathbf{R}^2$:*

$$\phi(\mathbf{u}, \mathbf{v}) = ax_1y_1 + bx_1y_2 + bx_2y_1 + cx_2y_2.$$

La matrice di Gram G di ϕ rispetto alla base canonica e':

$$G = \begin{bmatrix} a & b \\ b & c \end{bmatrix}.$$

(ii) Si consideri la seguente forma quadratica su $\mathbf{R}^2$:

$$q(\mathbf{u}) = ax_1^2 + bx_1x_2 + cx_2^2.$$

La matrice di Gram G di q rispetto alla base canonica e':

$$G = \begin{bmatrix} a & \frac{b}{2} \\ \frac{b}{2} & c \end{bmatrix}.$$

Infatti

$$ax_1^2 + bx_1x_2 + cx_2^2 = \mathbf{x}^T \begin{bmatrix} a & \frac{b}{2} \\ \frac{b}{2} & c \end{bmatrix} \mathbf{x}.$$

(iii) Si consideri la seguente forma bilineare simmetrica su $\mathbf{R}^3$*:*

$$\phi(\mathbf{u},\mathbf{v}) = ax_1y_1 + bx_1y_2 + cx_1y_3 + bx_2y_1 + dx_2y_2 + ex_2y_3 + cx_3y_1 + ex_3y_2 + fx_3y_3.$$

La matrice di Gram G *di* ϕ *rispetto alla base canonica e':*

$$G = \begin{bmatrix} a & b & c \\ b & d & e \\ c & e & f \end{bmatrix}.$$

(iv) Si consideri la seguente forma quadratica su $\mathbf{R}^3$*:*

$$q(\mathbf{u}) = ax_1^2 + bx_1x_2 + cx_1x_3 + dx_2^2 + ex_2x_3 + fx_3^2.$$

La matrice di Gram G *di* q *rispetto alla base canonica e':*

$$G = \begin{bmatrix} a & \frac{b}{2} & \frac{c}{2} \\ \frac{b}{2} & d & \frac{e}{2} \\ \frac{c}{2} & \frac{e}{2} & f \end{bmatrix}.$$

Infatti

$$ax_1^2 + bx_1x_2 + cx_1x_3 + dx_2^2 + ex_2x_3 + fx_3^2 = \mathbf{x}^T \begin{bmatrix} a & \frac{b}{2} & \frac{c}{2} \\ \frac{b}{2} & d & \frac{e}{2} \\ \frac{c}{2} & \frac{e}{2} & f \end{bmatrix} \mathbf{x}. \quad \blacksquare$$

3. Diagonalizzazione di una forma quadratica.

Ora proveremo che ogni forma bilineare simmetrica (quindi ogni forma quadratica) puo' essere diagonalizzata, cioe':

Teorema 1. *Sia* ϕ *una forma bilineare simmetrica su uno spazio* V*. Allora esiste una base* $\mathcal{B}$ *di* V *tale che* $G_{\mathcal{B}}^{\mathcal{B}}(\phi)$ *sia una matrice diagonale.*

La dimostrazione si basa su un algoritmo, detto Algoritmo di Gauss-Lagrange, e procede cosi': si consideri una base qualunque $\mathcal{E}$ di V, e sia $G_{\mathcal{E}}^{\mathcal{E}}(\phi)$ la relativa matrice di Gram che in generale sara' una matrice non diagonale. Affianchiamo a G la matrice I, formando la matrice $n \times 2n$ ($n = \dim V$)

$$M := [\quad G \quad | \quad I \quad].$$

Sulla matrice M applichiamo l'Algoritmo di Gauss-Lagrange che dice quanto segue:

Algoritmo di Gauss-Lagrange

Passo 1

1.1 Se la matrice M *ha una riga l'algoritmo termina; altrimenti:*

1.2 individuare la colonna non nulla con indice piu' basso, ed il suo pivot a_{ij}; se non esistono colonne non nulle la matrice e' nulla e l'algoritmo termina qui; altrimenti:

1.3 se $i = j$ passare ad 1.5; se $i \neq j$ sommare alla j-esima riga la i-esima riga, e sommare alla j-esima colonna la i-esima colonna;

1.4 se dopo queste operazioni l'elemento di posto j, j e' nullo, ripetere il passo 1.3; altrimenti:

1.5 rendere nulli tutti gli altri elementi della colonna j-esima sommando alle varie righe opportuni multipli della j-esima riga;

1.6 ripetere sulle colonne le stesse operazioni elementari che sono state effettuate in 1.5 sulle righe, nello stesso ordine.

Passo 2
Ripetere il Passo 1 sulla matrice ottenuta dal passo precedente schermandone le prime j righe e le prime j colonne.

Passo 3
Ripetere il Passo 2 sulla matrice schermata. ■

Al termine di questo algoritmo si perviene ad una matrice del tipo:

$$[\quad G \quad | \quad I \quad] \longrightarrow [\quad D \quad | \quad P^T \quad]$$

dove D e' una matrice diagonale e P^T una matrice invertibile tale che

$$D = P^T G P.$$

Quindi se $\mathcal{B}$ e' la base di V costituita da quei vettori le cui coordinate rispetto alla base $\mathcal{E}$ sono date dalle colonne della matrice P (cioe' se $\mathcal{B}$ e' la base di V tale che $M_{\mathcal{E}}^{\mathcal{B}}(id_V) = P$), si ha

$$G_{\mathcal{B}}^{\mathcal{B}}(\phi) = D,$$

e cio' conclude la dimostrazione del Teorema 1.

Si osservi che, se denotiamo con $\mathbf{x}$ le coordinate rispetto ad una base $\mathcal{B}$ che diagonalizza ϕ, e se denotiamo con $d_1, d_2, \ldots, d_n$ le componenti sulla diagonale principale di $D = G_{\mathcal{B}}^{\mathcal{B}}(\phi)$, allora si ha

$$\phi(\mathbf{u}, \mathbf{v}) = d_1 x_1 y_1 + d_2 x_2 y_2 + \cdots + d_n x_n y_n.$$

Come conseguenze del Teorema precedente abbiamo i seguenti Corollari, che lo traducono in termini di forme quadratiche e di matrici simmetriche.

Corollario 1. *Sia q una forma quadratica su uno spazio V di dimensione n. Allora, per una opportuna scelta delle coordinate $\mathbf{x}$ in V, esistono delle costanti $d_1, d_2, \ldots, d_n$ tali che per ogni $\mathbf{u} \in V$ si ha*

$$q(\mathbf{u}) = d_1 x_1^2 + d_2 x_2^2 + \cdots + d_n x_n^2.$$

Corollario 2. *Ogni matrice reale simmetrica G e' congruente a qualche matrice diagonale. Cioe' esiste una matrice diagonale D ed una matrice invertibile P tale che*

$$D = P^T G P.$$

Esempio 8. *Consideriamo la forma bilineare simmetrica $\phi : \mathbf{R}^2 \times \mathbf{R}^2 \to \mathbf{R}$ definita ponendo*

$$\phi(\mathbf{u}, \mathbf{v}) := x_1y_1 + 3x_1y_2 + 3x_2y_1 + 2x_2y_2.$$

Andiamo a trovare una base che diagonalizzi ϕ.

A tale proposito, affianchiamo alla matrice di Gram G di ϕ rispetto alla base canonica la matrice I:

$$[\quad G \quad | \quad I \quad] = \begin{bmatrix} 1 & 3 & 1 & 0 \\ 3 & 2 & 0 & 1 \end{bmatrix}.$$

Applicando l'Algoritmo di Gauss-Lagrange (cioe' l'operazione elementare sulle righe $e_{21}(-3)$ e la stessa operazione sulle colonne $e^{21}(-3)$) perveniamo alla matrice

$$[\quad D \quad | \quad P^T \quad] = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & -7 & -3 & 1 \end{bmatrix}.$$

Posto allora $\mathcal{B} = \{(1,0), (-3,1)\}$ si ha

$$G_{\mathcal{B}}^{\mathcal{B}}(\phi) = \begin{bmatrix} 1 & 0 \\ 0 & -7 \end{bmatrix},$$

e $\mathcal{B}$ e' la base cercata. Se denotiamo con $\mathbf{x}'$ le coordinate rispetto a tale base abbiamo

$$\phi(\mathbf{u}, \mathbf{v}) = x_1'y_1' - 7x_2'y_2'$$

(quindi se nell'espressione $\phi(\mathbf{u}, \mathbf{v}) = x_1y_1 + 3x_1y_2 + 3x_2y_1 + 2x_2y_2$ effettuiamo la sostituzione

$$\begin{cases} x_1 = x_1' - 3x_2' \\ x_2 = x_2' \end{cases}$$

e semplifichiamo, otteniamo $\phi(\mathbf{u}, \mathbf{v}) = x_1'y_1' - 7x_2'y_2'$). ■

Esempio 9. *Trovare una base che diagonalizzi la forma quadratica*

$$q(\mathbf{u}) := x_1x_2.$$

Come prima affianchiamo alla matrice di Gram G di q rispetto alla base canonica la matrice I:

$$[\quad G \quad | \quad I \quad] = \begin{bmatrix} 0 & \frac{1}{2} & 1 & 0 \\ \frac{1}{2} & 0 & 0 & 1 \end{bmatrix}.$$

Applicando l'Algoritmo di Gauss-Lagrange (cioe' le operazioni elementari $e_{12}(1)$, $e^{12}(1)$, $e_{21}(-\frac{1}{2})$, $e^{21}(-\frac{1}{2})$) perveniamo alla matrice

$$\left[\quad D \quad | \quad P^T \quad\right] = \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & -\frac{1}{4} & -\frac{1}{2} & \frac{1}{2} \end{bmatrix}.$$

Posto allora $\mathcal{B} = \{(1,1), (-\frac{1}{2}, \frac{1}{2})\}$ si ha

$$G_{\mathcal{B}}^{\mathcal{B}}(\phi) = \begin{bmatrix} 1 & 0 \\ 0 & -\frac{1}{4} \end{bmatrix},$$

e $\mathcal{B}$ e' la base cercata. Se denotiamo con $\mathbf{x}'$ le coordinate rispetto a tale base abbiamo

$$q(\mathbf{u}) = {x'_1}^2 - \frac{1}{4}{x'_2}^2. \quad \blacksquare$$

Esempio 10. *Consideriamo la forma bilineare simmetrica $\phi : \mathbf{R}^3 \times \mathbf{R}^3 \to \mathbf{R}$ definita ponendo*

$$\phi(\mathbf{u}, \mathbf{v}) := x_1y_1 + x_1y_2 + x_1y_3 + x_2y_1 + x_2y_3 + x_3y_1 + x_3y_2 + x_3y_3.$$

Andiamo a trovare una base che diagonalizzi ϕ.

A tale proposito, affianchiamo alla matrice di Gram G di ϕ rispetto alla base canonica la matrice I:

$$\left[\quad G \quad | \quad I \quad\right] = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 \end{bmatrix}.$$

Applicando l'Algoritmo di Gauss-Lagrange (cioe' applicando $e_{21}(-1)$, $e^{21}(-1)$, $e_{31}(-1)$, $e^{31}(-1)$) perveniamo alla matrice

$$\left[\quad D \quad | \quad P^T \quad\right] = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & -1 & 0 & 1 \end{bmatrix}.$$

Posto allora $\mathcal{B} = \{(1,0,0), (-1,1,0), (-1,0,1)\}$ si ha

$$G_{\mathcal{B}}^{\mathcal{B}}(\phi) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

e $\mathcal{B}$ e' la base cercata. Se denotiamo con $\mathbf{x}'$ le coordinate rispetto a tale base abbiamo

$$\phi(\mathbf{u}, \mathbf{v}) = x'_1y'_1 - x'_2y'_2. \quad \blacksquare$$

4. Spazi pseudoeuclidei.

Dato uno spazio vettoriale V, e fissata una forma bilineare simmetrica ϕ su V, si dice che *la coppia* (V, ϕ) *e' uno spazio pseudoeuclideo.* In uno spazio pseudoeuclideo (V, ϕ) due vettori $\mathbf{u}$, $\mathbf{v}$ si dicono *ortogonali* se $\phi(\mathbf{u}, \mathbf{v}) = 0$. Una base $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n\}$ di (V, ϕ) si dice che e' *una base ortogonale* se i vettori di $\mathcal{B}$ sono mutualmente ortogonali, cioe' se per ogni $i \neq j$ si ha $\phi(\mathbf{b}_i, \mathbf{b}_j) = 0$. In altre parole $\mathcal{B}$ e' una base ortogonale se e solo se la matrice di Gram $G_{\mathcal{B}}^{\mathcal{B}}(\phi)$ e' una matrice diagonale. Abbiamo appena dimostrato (Teorema 1) che in ogni spazio pseudoeuclideo ci sono basi ortogonali. Fissiamo una tale base, sia D la matrice di Gram relativa (che e' diagonale) e siano $d_1, d_2, \dots, d_n$ le componenti di D sulla diagonale principale (le uniche eventualmente non nulle). Cioe'

$$d_i := \phi(\mathbf{b}_i, \mathbf{b}_i).$$

Ora a partire da (V, ϕ) e dalla base ortogonale $\mathcal{B}$ possiamo definire i seguenti numeri interi:

$$(n, p, r, s)$$

dove n denota la dimensione di V, p e' il rango di D (cioe' il numero di componenti diverse da 0 sulla diagonale principale di D), r e' *l'indice di* D, cioe' il numero di componenti positive sulla diagonale principale di D, ed s e' la *segnatura* di D, cioe' la differenza tra il numero di componenti positive ed il numero di componenti negative sulla diagonale principale di D. Si osservi che s e' funzione di p e di r in quanto $s = 2r - p$. L'interesse di questi numeri risiede nel fatto che essi *non dipendono dalla base ortogonale fissata.* Cio' si deduce dal celebre

Teorema di Sylvester (Legge di inerzia). *Due matrici reali diagonali congruenti hanno lo stesso rango, lo stesso numero di componenti positive, lo stesso numero di componenti negative, e lo stesso numero di componenti nulle.*

Andiamo a dimostrare il Teorema di Sylvester. Prima pero' abbiamo bisogno del seguente

Lemma 1. *Sia A una matrice $m \times n$, P una matrice invertibile $m \times m$ e Q un matrice invertibile $n \times n$. Allora il rango di A e' uguale al rango di PAQ. In altre parole, la moltiplicazione per una matrice invertibile (a destra o a sinistra) non altera il rango.*

Dimostrazione del Lemma 1. Per l'Algoritmo di Gauss-Jordan sappiamo che, dopo un certo numero finito di operazioni elementari sulle righe di P, la matrice P, che e' invertibile, diventa la matrice I. Ricordiamo che eseguire un'operazione elementare sulle righe equivale a moltiplicare a sinistra per la matrice elementare corrispondente. Per cui possiamo dire che poiche' P e' invertibile allora esistono opportune matrici elementari $E_1, E_2, \dots, E_h$ tali che $E_1 \cdot E_2 \cdot \ldots \cdot E_h \cdot P = I$. Ne consegue

$$(E_1 \cdot E_2 \cdot \ldots \cdot E_h) \cdot (P \cdot A) = (E_1 \cdot E_2 \cdot \ldots \cdot E_h \cdot P) \cdot A = I \cdot A = A.$$

Cio' significa che le stesse operazioni elementari che portano P in I, portano $P \cdot A$ in A. Poiche' le operazioni elementari non alterano il rango segue che il rango di $P \cdot A$ e' uguale al rango di A.

Per concludere sara' sufficiente provare che il rango di $P \cdot A$ e' uguale al rango di $P \cdot A \cdot Q$.

Per fare questo ragioniamo in modo simile al precedente, utiizzando questa volta le operazioni elementari sulle colonne di Q, e tenendo presente che eseguire un'operazione elementare sulle colonne equivale a moltiplicare a destra per la matrice elementare corrispondente. Quindi, poiche' Q e' invertibile allora dopo un opportuno numero di operazioni elementari sulle colonne di Q, la matrice Q diviene la matrice I. Cioe' esistono opportune matrici elementari $F_1, F_2, \dots, F_k$ tali che $Q \cdot F_1 \cdot F_2 \cdot \dots \cdot F_k = I$. Segue che

$$(P \cdot A \cdot Q) \cdot (F_1 \cdot F_2 \cdot \dots \cdot F_k) = (P \cdot A) \cdot (Q \cdot F_1 \cdot F_2 \cdot \dots \cdot F_k) = P \cdot A \cdot I = P \cdot A.$$

Cioe' $P \cdot A$ si ottiene da $P \cdot A \cdot Q$ tramite delle operazioni elementari sulle colonne di $P \cdot A \cdot Q$. Poiche' le operazioni elementari (sulle righe o sulle colonne) non alterano il rango segue che il rango di $P \cdot A \cdot Q$ e' uguale al rango di $P \cdot A$. Cio' conclude la dimostrazione del Lemma 1.

Possiamo cominciare la dimostrazione del Teorema di Sylvester.

A tale proposito siano D e D' due matrici reali congruenti. Allora per definizione esiste una matrice invertibile P tale che $D' = P^T DP$. Per il lemma appena dimostrato possiamo dire che D e D' sicuramente hanno lo stesso rango. Ne consegue che per dimostrare il Teorema di Sylvester e' sufficiente provare che l'indice r di D coincide con l'indice r' di D'.

Per provare cio' cominciamo col ricordare che poiche' D e D' sono congruenti allora esiste una forma bilineare simmetrica $\phi : \mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}$ e basi $\mathcal{B}$ e $\mathcal{B}'$ di $\mathbf{R}^n$ tali che $G_{\mathcal{B}}^{\mathcal{B}}(\phi) = D$ e $G_{\mathcal{B}'}^{\mathcal{B}'}(\phi) = D'$ (qui n denota l'ordine di D, che e' anche l'ordine di D').

Scambiando eventualmente l'ordine con cui appaiono i vettori della base $\mathcal{B}$, possiamo sempre supporre che la diagonale principale della matrice D sia un vettore numerico di lunghezza n della forma

$$(d_1, \dots, d_r, -d_{r+1}, \dots, -d_p, 0, \dots, 0)$$

dove i numeri d_i sono tutti > 0, cosi' che r e' l'indice di D e p ne e' il rango. Quindi se denotiamo con $\mathbf{b}_i$ i vettori della base $\mathcal{B}$ abbiamo

- $\phi(\mathbf{b}_i, \mathbf{b}_j) = 0$ per ogni $i \neq j$;
- $\phi(\mathbf{b}_i, \mathbf{b}_i) = d_i > 0$ per ogni $1 \leq i \leq r$;
- $\phi(\mathbf{b}_i, \mathbf{b}_i) = -d_i < 0$ per ogni $r + 1 \leq i \leq p$;
- $\phi(\mathbf{b}_i, \mathbf{b}_i) = 0$ per ogni $p + 1 \leq i \leq n$.

Analogamente possiamo supporre che la diagonale principale della matrice D' sia un vettore numerico di lunghezza n della forma

$$(d'_1, \dots, d'_{r'}, -d'_{r'+1}, \dots, -d'_p, 0, \dots, 0)$$

dove i numeri d'_i sono tutti > 0, cosi' che r' e' l'indice di D' e p ne e' il rango. Quindi se denotiamo con $\mathbf{b}'_i$ i vettori della base $\mathcal{B}'$ abbiamo

- $\phi(\mathbf{b}'_i, \mathbf{b}'_j) = 0$ per ogni $i \neq j$;
- $\phi(\mathbf{b}'_i, \mathbf{b}'_i) = d'_i > 0$ per ogni $1 \leq i \leq r'$;
- $\phi(\mathbf{b}'_i, \mathbf{b}'_i) = -d'_i < 0$ per ogni $r' + 1 \leq i \leq p$;
- $\phi(\mathbf{b}'_i, \mathbf{b}'_i) = 0$ per ogni $p + 1 \leq i \leq n$.

Ricordiamo che il nostro obiettivo consiste nel provare che $r = r'$.

Consideriamo il seguente sistema di vettori di $\mathbf{R}^n$:

$$\mathcal{S} := \{\mathbf{b}_1, \dots, \mathbf{b}_r, \mathbf{b}'_{r'+1}, \dots, \mathbf{b}'_n\}.$$

Andiamo a provare che $\mathcal{S}$ e' linearmente indipendente. Sia allora

$$h_1\mathbf{b}_1 + \cdots + h_r\mathbf{b}_r + k_{r'+1}\mathbf{b}'_{r'+1} + \cdots + k_n\mathbf{b}'_n = \mathbf{0} \tag{1}$$

una relazione per il sistema $\mathcal{S}$. Poniamo

$$\mathbf{u} := \sum_{i=1}^{r} h_i\mathbf{b}_i = -\sum_{i=r'+1}^{n} k_i\mathbf{b}'_i,$$

e calcoliamo

$$\phi(\mathbf{u}, \mathbf{u}) = \phi(\sum_{i=1}^{r} h_i\mathbf{b}_i, \sum_{i=1}^{r} h_i\mathbf{b}_i) = \sum_{i,j=1}^{r} h_ih_j\phi(\mathbf{b}_i, \mathbf{b}_j) = \sum_{i=1}^{r} h_i^2 d_i \geq 0.$$

Possiamo svolgere il calcolo precedente anche nel seguente modo:

$$\phi(\mathbf{u}, \mathbf{u}) = \phi(-\sum_{i=r'+1}^{n} k_i\mathbf{b}'_i, -\sum_{i=r'+1}^{n} k_i\mathbf{b}'_i) = \sum_{i,j=r'+1}^{n} k_ik_j\phi(\mathbf{b}'_i, \mathbf{b}'_j) = \sum_{i=r'+1}^{p} k_i^2(-d'_i) \leq 0.$$

Si deduce che

$$\phi(\mathbf{u}, \mathbf{u}) = 0.$$

In particolare

$$\sum_{i=1}^{r} h_i^2 d_i = 0.$$

Poiche' $d_i > 0$ per $1 \leq i \leq r$ allora tutti i numeri h_i sono nulli. Dalla relazione (1) segue che anche tutti i numeri k_i sono nulli e percio' la relazione (1) e' banale. Cio' prova che $\mathcal{S}$ e' linearmente indipendente. Poiche' in $\mathcal{S}$ ci sono $r + n - r'$ vettori, per il Lemma di Steinitz deve essere $r + n - r' \leq n$, cioe'

$$r \leq r'.$$

Ripetendo lo stesso ragionamento per il sistema di vettori

$$\mathcal{S}' := \{\mathbf{b}'_1, \dots, \mathbf{b}'_{r'}, \mathbf{b}_{r+1}, \dots, \mathbf{b}_n\}$$

proveremo che

$$r' \leq r,$$

e dunque che $r = r'$. Cio' conclude la dimostrazione del Teorema di Sylvester.

In base al Teorema di Sylvester la seguente definizione e' ben posta. Sia (V, ϕ) uno spazio pseudoeuclideo, $\mathcal{B}$ una base ortogonale di V, e D la matrice (diagonale) di Gram di ϕ rispetto alla base $\mathcal{B}$. Allora si definisce *rango di* (V, ϕ) il rango di D, *indice di* (V, ϕ) l'indice di D, *la segnatura di* (V, ϕ) la segnatura di D. Definizioni corrispondenti si hanno per le forme quadratiche, e le matrici simmetriche. E cioe' si definiscono rango, indice e segnatura di una forma quadratica q come il rango, l'indice e la segnatura della forma bilineare che si ottiene polarizzando q. E si definiscono rango, indice e segnatura di una matrice reale simmetrica G come il rango, l'indice e la segnatura di una qualunque matrice diagonale congruente a G.

Esempio 11. *(i) Calcolare rango, indice e segnatura della matrice*

$$G := \begin{bmatrix} -4 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 9 \end{bmatrix}.$$

Poiche' la matrice assegnata e' diagonale, possiamo leggere i dati richiesti sulla sua diagonale principale. Quindi il rango e' 2*, l'indice e'* 1 *e la segnatura e'* 0.

(ii) Calcolare rango, indice e segnatura della matrice

$$G := \begin{bmatrix} 1 & -2 & 3 \\ -2 & 6 & -10 \\ 3 & -10 & 8 \end{bmatrix}.$$

In questo caso occorre diagonalizzare G *rispetto alla congruenza. Possiamo fare cio' con l'Algoritmo di Gauss-Lagrange. Utilizzando le operazioni* $e_{21}(2)$, $e_{31}(-3)$, $e^{21}(2)$, $e^{31}(-3)$, $e_{32}(2)$, $e^{32}(2)$ *si ottiene la matrice*

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -9 \end{bmatrix}.$$

Dunque G *ha rango* 3*, indice* 2*, e segnatura* 1.

(iii) Si consideri lo spazio pseudoeuclideo $(\mathbf{R}^3, \phi)$ *dove* ϕ *e' la forma bilineare simmetrica definita ponendo*

$$\phi(\mathbf{u}, \mathbf{v}) := x_1y_1 - 2x_1y_2 + 3x_1y_3 - 2x_2y_1 + 6x_2y_2 - 10x_2y_3 + 3x_3y_1 - 10x_3y_2 + 8x_3y_3.$$

Calcolare una base ortogonale di $(\mathbf{R}^3, \phi)$*, il rango, l'indice e la segnatura.*

Osserviamo che la matrice di Gram rispetto alla base canonica e' la stessa matrice G *dell'esempio precedente. Quindi con le stesse operazioni di prima applicate alla matrice*

$$[\quad G \quad | \quad I \quad] = \begin{bmatrix} 1 & -2 & 3 & 1 & 0 & 0 \\ -2 & 6 & -10 & 0 & 1 & 0 \\ 3 & -10 & 8 & 0 & 0 & 1 \end{bmatrix}$$

perveniamo alla matrice

$$[\quad D \quad | \quad P^T \quad] = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 2 & 0 & 2 & 1 & 0 \\ 0 & 0 & -9 & 1 & 2 & 1 \end{bmatrix}.$$

Ne consegue che lo spazio pseudoeuclideo $(\mathbf{R}^3, \phi)$ *ha rango* 3*, indice* 2*, e segnatura* 1*, ed una sua base ortogonale* $\mathcal{B}$ *e' formata dai vettori* $\mathbf{e}_1$*,* $2\mathbf{e}_1 + \mathbf{e}_2$*,* $\mathbf{e}_1 + 2\mathbf{e}_2 + \mathbf{e}_3$*. In particolare se ci riferiamo alle coordinate* $\mathbf{x}'$ *rispetto alla base* $\mathcal{B}$ *allora abbiamo*

$$\phi(\mathbf{u}, \mathbf{v}) := x_1'y_1' + 2x_2'y_2' - 9x_3'y_3'.$$

Cioe' se sostituiamo le coordinate canoniche $\mathbf{x}$ *con* $\mathbf{x} = P\mathbf{x}'$*, cioe' se poniamo*

$$\begin{cases} x_1 = x_1' + 2x_2' + x_3' \\ x_2 = x_2' + 2x_3' \\ x_3 = x_3' \end{cases}$$

nell'espressione

$$\phi(\mathbf{u}, \mathbf{v}) := x_1y_1 - 2x_1y_2 + 3x_1y_3 - 2x_2y_1 + 6x_2y_2 - 10x_2y_3 + 3x_3y_1 - 10x_3y_2 + 8x_3y_3$$

e semplifichiamo, otteniamo

$$\phi(\mathbf{u}, \mathbf{v}) := x_1'y_1' + 2x_2'y_2' - 9x_3'y_3'. \quad \blacksquare$$

Consideriamo di nuovo uno spazio pseudoeuclideo (V, ϕ) di dimensione n, rango p ed indice r. Sia $\mathcal{B} := \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ una base ortogonale di V, e D la matrice (diagonale) di Gram di ϕ rispetto alla base $\mathcal{B}$. Riordinando eventualmente i vettori nella base $\mathcal{B}$, possiamo supporre che la diagonale principale di D sia del tipo

$$d_1, \dots, d_r, -d_{r+1}, \dots, -d_p, 0, \dots, 0$$

con $d_i > 0$ per ogni $i = 1, \dots, p$. Ricordiamo che $\phi(\mathbf{b}_i, \mathbf{b}_i)$ coincide con la componente di posto i della diagonale di D. Adesso, deformando opportunamente i vettori di $\mathcal{B}$, definiamo la seguente base $\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_n\}$ ponendo

$$\mathbf{c}_i := \frac{1}{\sqrt{d_i}}\mathbf{b}_i \quad \textit{per} \quad i = 1, \dots, p, \quad \text{e ponendo} \quad \mathbf{c}_i := \mathbf{b}_i \quad \textit{per} \quad i = p+1, \dots, n.$$

La base $\mathcal{C}$ e' ancora ortogonale, e la diagonale principale di $G_{\mathcal{C}}^{\mathcal{C}}(\phi)$ ha le seguenti componenti:

$\phi(\mathbf{c}_i, \mathbf{c}_i) = 1$ se $i = 1, \ldots, r$
$\phi(\mathbf{c}_i, \mathbf{c}_i) = -1$ se $i = r+1, \ldots, p$
$\phi(\mathbf{c}_i, \mathbf{c}_i) = 0$ se $i = p+1, \ldots, n$.

Cioe' la matrice $G_{\mathcal{C}}^{\mathcal{C}}(\phi)$ e' la matrice diagonale a blocchi

$$\begin{bmatrix} I_r & & \\ & -I_{p-r} & \\ & & \mathbf{0}_{n-p} \end{bmatrix}$$

dove I_r e' la matrice unitaria di ordine r, $-I_{p-r}$ e' l'opposto della matrice unitaria di ordine $p-r$, e $\mathbf{0}_{n-p}$ e' la matrice nulla di ordine $n-p$. Una base $\mathcal{C}$ siffatta si dice *base ortonormale per lo spazio pseudoeuclideo* (V, ϕ).

Possiamo riassumere quanto appena detto nella seguente

Proposizione 2. *Sia (V, ϕ) uno spazio pseudoeuclideo di dimensione n, rango p ed indice r. Allora esiste una base ortonormale $\mathcal{B}$ per (V, ϕ), cioe' una base $\mathcal{B}$ tale che*

$$G_{\mathcal{B}}^{\mathcal{B}}(\phi) = \begin{bmatrix} I_r & & \\ & -I_{p-r} & \\ & & \mathbf{0}_{n-p} \end{bmatrix}.$$

Se $\mathbf{x}$ denotano le coordinate rispetto alla base $\mathcal{B}$ allora

$$\phi(\mathbf{u}, \mathbf{v}) = x_1 y_1 + \cdots + x_r y_r - x_{r+1} y_{r+1} - \cdots - x_p y_p,$$

e se q e' la forma quadratica associata a ϕ si ha

$$q(\mathbf{u}) = x_1^2 + \cdots + x_r^2 - x_{r+1}^2 - \cdots - x_p^2.$$

Corollario 3. *Sia G una matrice reale simmetrica $n \times n$, di rango p ed indice r. Allora G e' congruente alla matrice*

$$\begin{bmatrix} I_r & & \\ & -I_{p-r} & \\ & & \mathbf{0}_{n-p} \end{bmatrix}.$$

La matrice che appare nel corollario precedente si dice *la forma canonica di G rispetto alla congruenza*. Si osservi che, per il Teorema di Sylvester, la forma canonica rispetto alla congruenza di una matrice G *e' unica*.

Corollario 4. *Siano G e G' due matrici reali simmetriche $n \times n$. Allora G e G' sono congruenti se e solo se hanno lo stesso rango e lo stesso indice.*

Infatti se hanno lo stesso rango e lo stesso indice allora hanno la stessa forma canonica rispetto alla congruenza, e per transitivita' sono congruenti.

Esempio 12. *Si consideri lo spazio pseudoeuclideo $(\mathbf{R}^3, \phi)$ dove ϕ e' la forma bilineare simmetrica definita ponendo*

$$\phi(\mathbf{u}, \mathbf{v}) := x_1y_1 - 2x_1y_2 + 3x_1y_3 - 2x_2y_1 + 6x_2y_2 - 10x_2y_3 + 3x_3y_1 - 10x_3y_2 + 8x_3y_3.$$

Calcolare una base ortonormale di $(\mathbf{R}^3, \phi)$.

Svolgimento. Nell'Esempio 11 abbiamo gia' visto che la base $\mathcal{B} = \{\mathbf{e}_1, 2\mathbf{e}_1 + \mathbf{e}_2, \mathbf{e}_1 + 2\mathbf{e}_2 + \mathbf{e}_3\}$ e' una base ortogonale, con $\phi(\mathbf{e}_1, \mathbf{e}_1) = 1$, $\phi(2\mathbf{e}_1 + \mathbf{e}_2, 2\mathbf{e}_1 + \mathbf{e}_2) = 2$, $\phi(\mathbf{e}_1 + 2\mathbf{e}_2 + \mathbf{e}_3, \mathbf{e}_1 + 2\mathbf{e}_2 + \mathbf{e}_3) = -9$. Quindi $\mathcal{B}$ non e' una base ortonormale. Possiamo "normalizzare" $\mathcal{B}$ ponendo:

$$\mathbf{c}_1 := \frac{\mathbf{e}_1}{\sqrt{\phi(\mathbf{e}_1, \mathbf{e}_1)}}, \quad \mathbf{c}_2 := \frac{2\mathbf{e}_1 + \mathbf{e}_2}{\sqrt{\phi(2\mathbf{e}_1 + \mathbf{e}_2, 2\mathbf{e}_1 + \mathbf{e}_2)}}, \quad e$$

$$\mathbf{c}_3 := \frac{\mathbf{e}_1 + 2\mathbf{e}_2 + \mathbf{e}_3}{\sqrt{-\phi(\mathbf{e}_1 + 2\mathbf{e}_2 + \mathbf{e}_3, \mathbf{e}_1 + 2\mathbf{e}_2 + \mathbf{e}_3)}}.$$

Cioe' ponendo

$$\mathbf{c}_1 := \mathbf{e}_1, \quad \mathbf{c}_2 := \frac{2\mathbf{e}_1 + \mathbf{e}_2}{\sqrt{2}}, \quad \mathbf{c}_3 := \frac{\mathbf{e}_1 + 2\mathbf{e}_2 + \mathbf{e}_3}{\sqrt{9}}.$$

La base ortonormale cercata e' dunque:

$$\mathcal{C} := \{(1,0,0), (\sqrt{2}, \frac{1}{\sqrt{2}}, 0), (\frac{1}{3}, \frac{2}{3}, \frac{1}{3})\}.$$

Si osservi che

$$G_{\mathcal{C}}^{\mathcal{C}}(\phi) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

e che se denotiamo con $\mathbf{x}'$ le coordinate rispetto a tale base allora abbiamo

$$\phi(\mathbf{u}, \mathbf{v}) = x_1'y_1' + x_2'y_2' - x_3'y_3' \quad e \quad q(\mathbf{u}) = {x_1'}^2 + {x_2'}^2 - {x_3'}^2. \quad \blacksquare$$

Esempio 13. *Dire se le matrici*

$$\begin{bmatrix} 1 & -2 & 3 \\ -2 & 6 & -10 \\ 3 & -10 & 8 \end{bmatrix}, \quad \begin{bmatrix} -167 & 0 & 0 \\ 0 & 16 & 0 \\ 0 & 0 & 8\pi \end{bmatrix}$$

sono congruenti oppure no.

Svolgimento. Abbiamo visto negli esempi precedenti che la prima matrice ha rango 3 *ed indice* 2*. Poiche' anche la seconda matrice ha gli stessi invarianti allora le due matrici sono congruenti.* ■

Esempio 14. *Classificare le forme quadratiche* $q : \mathbf{R} \to \mathbf{R}$ *su* $\mathbf{R}$*, a meno di un cambiamento delle coordinate.*

Svolgimento. Sia $q : \mathbf{R} \to \mathbf{R}$ *una forma quadratica, e denotiamone con* (p, r, s) *il rango, l'indice e la segnatura. Allora, pur di effettuare un opportuno cambiamento di base,* q *si puo' rappresentare in uno ed uno solo dei seguenti modi:*

se $(p, r, s) = (1, 1, 1)$ *allora* $q(\mathbf{u}) = x_1^2$*;*

se $(p, r, s) = (1, 0, -1)$ *allora* $q(\mathbf{u}) = -x_1^2$*;*

se $(p, r, s) = (0, 0, 0)$ *allora* $q(\mathbf{u}) = 0$. ■

Esempio 15. *Classificare le forme quadratiche* $q : \mathbf{R}^2 \to \mathbf{R}$ *su* $\mathbf{R}^2$*, a meno di un cambiamento delle coordinate.*

Svolgimento. Sia $q : \mathbf{R}^2 \to \mathbf{R}$ *una forma quadratica, e denotiamone con* (p, r, s) *il rango, l'indice e la segnatura. Allora, pur di effettuare un opportuno cambiamento di base,* q *si puo' rappresentare in uno ed uno solo dei seguenti modi:*

se $(p, r, s) = (2, 2, 2)$ *allora* $q(\mathbf{u}) = x_1^2 + x_2^2$*;*

se $(p, r, s) = (2, 1, 0)$ *allora* $q(\mathbf{u}) = x_1^2 - x_2^2$*;*

se $(p, r, s) = (2, 0, -2)$ *allora* $q(\mathbf{u}) = -x_1^2 - x_2^2$.

se $(p, r, s) = (1, 1, 1)$ *allora* $q(\mathbf{u}) = x_1^2$*;*

se $(p, r, s) = (1, 0, -1)$ *allora* $q(\mathbf{u}) = -x_1^2$*;*

se $(p, r, s) = (0, 0, 0)$ *allora* $q(\mathbf{u}) = 0$. ■

Esempio 16. *Classificare le matrici reali simmetriche* 2×2 *rispetto alla congruenza.*

Svolgimento. Sia G *una matrice reale simmetrica* 2×2*, e denotiamone con* (p, r, s) *il rango, l'indice e la segnatura. Allora* G *e' congruente ad una ed una sola delle seguenti matrici:*

$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ *se* $(p, r, s) = (2, 2, 2)$*;*

$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ *se* $(p, r, s) = (2, 1, 0)$*;*

$\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$ *se* $(p, r, s) = (2, 0, -2)$*;*

$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ *se* $(p, r, s) = (1, 1, 1)$*;*

$\begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix}$ *se* $(p, r, s) = (1, 0, -1)$*;*

$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ *se* $(p, r, s) = (0, 0, 0)$. ■

Esempio 17. *Classificare le forme quadratiche* $q : \mathbf{R}^3 \to \mathbf{R}$ *su* $\mathbf{R}^3$*, a meno di un cambiamento delle coordinate.*

Svolgimento. Sia $q : \mathbf{R}^3 \to \mathbf{R}$ *una forma quadratica, e denotiamone con* (p, r, s) *il rango, l'indice e la segnatura. Allora, pur di effettuare un opportuno cambiamento di base,* q *si puo' rappresentare in uno ed uno solo dei seguenti modi:*

se $(p, r, s) = (3, 3, 3)$ *allora* $q(\mathbf{u}) = x_1^2 + x_2^2 + x_3^2$*;*

se $(p, r, s) = (3, 2, 1)$ *allora* $q(\mathbf{u}) = x_1^2 + x_2^2 - x_3^2$*;*

se $(p, r, s) = (3, 1, -1)$ *allora* $q(\mathbf{u}) = x_1^2 - x_2^2 - x_3^2$*;*

se $(p, r, s) = (3, 0, -3)$ *allora* $q(\mathbf{u}) = -x_1^2 - x_2^2 - x_3^2$*;*

se $(p, r, s) = (2, 2, 2)$ *allora* $q(\mathbf{u}) = x_1^2 + x_2^2$*;*

se $(p, r, s) = (2, 1, 0)$ *allora* $q(\mathbf{u}) = x_1^2 - x_2^2$*;*

se $(p, r, s) = (2, 0, -2)$ *allora* $q(\mathbf{u}) = -x_1^2 - x_2^2$.

se $(p, r, s) = (1, 1, 1)$ *allora* $q(\mathbf{u}) = x_1^2$*;*

se $(p, r, s) = (1, 0, -1)$ *allora* $q(\mathbf{u}) = -x_1^2$*;*

se $(p, r, s) = (0, 0, 0)$ *allora* $q(\mathbf{u}) = 0$. ■

5. Classificazione degli spazi pseudoeuclidei.

Una classificazione generale degli spazi pseudoeuclidei e' la seguente. Sia (V, ϕ) uno spazio pseudoeuclideo, di dimensione n, rango p, ed indice r.

• Lo spazio (V, ϕ) si dice *definito positivo* (o anche che ϕ *e' definita positiva*) se $n = p = r$. Cio' equivale a dire che per un'opportuna scelta delle coordinate si ha $q(\mathbf{u}) = x_1^2 + x_2^2 + \cdots + x_n^2$ (come sempre, q denota la forma quadratica associata a ϕ, cioe' $q(\mathbf{u}) = \phi(\mathbf{u}, \mathbf{u})$). Il che equivale anche a dire che $q(\mathbf{u}) > 0$ per ogni $\mathbf{u} \neq \mathbf{0}$.

• Lo spazio (V, ϕ) si dice *semidefinito positivo* se $n > p = r$. Cio' equivale a dire che per un'opportuna scelta delle coordinate si ha $q(\mathbf{u}) = x_1^2 + x_2^2 + \cdots + x_r^2$ con $r < n$. Il che equivale anche a dire che $q(\mathbf{u}) \geq 0$ per ogni $\mathbf{u}$ ed esiste un $\mathbf{u} \neq \mathbf{0}$ tale che $q(\mathbf{u}) = 0$.

• Lo spazio (V, ϕ) si dice *definito negativo* se $(V, -\phi)$ e' definito positivo. Cio' equivale a dire che $n = p$ ed $r = 0$. Cio' equivale anche a dire che per un'opportuna scelta delle coordinate si ha $q(\mathbf{u}) = -x_1^2 - x_2^2 - \cdots - x_n^2$. Il che equivale anche a dire che $q(\mathbf{u}) < 0$ per ogni $\mathbf{u} \neq \mathbf{0}$.

• Lo spazio (V, ϕ) si dice *semidefinito negativo* se $(V, -\phi)$ e' semidefinito positivo. Cio' equivale a dire che $n > p$ ed $r = 0$. Cio' equivale anche a dire che per un'opportuna scelta delle coordinate si ha $q(\mathbf{u}) = -x_1^2 - x_2^2 - \cdots - x_p^2$ con $p < n$. Il che equivale anche a dire che $q(\mathbf{u}) \leq 0$ per ogni $\mathbf{u}$ ed esiste un $\mathbf{u} \neq \mathbf{0}$ tale che $q(\mathbf{u}) = 0$.

• Lo spazio (V, ϕ) si dice *indefinito* se esistono vettori $\mathbf{u}$, $\mathbf{v}$ non nulli tali che $q(\mathbf{u}) > 0$ e $q(\mathbf{v}) < 0$. Cio' equivale a dire che $0 < r < p$.

Esistono definizioni corrispondenti per le matrici reali simmetriche. Una matrice simmetrica G $n \times n$ si dice *definita positiva* se la forma quadratica $\mathbf{x} \in \mathbf{R}^n \to \mathbf{x}^T G\mathbf{x} \in \mathbf{R}$ e' definita positiva (in tal caso si puo' anche scrivere $G > 0$). Analogamente possiamo definire una matrice *semidefinita positiva, definita negativa, semidefinita negativa, indefinita*. Si osservi che per la stessa definizione, una matrice reale e simmetrica G e' definita positiva se e solo se e' congruente alla matrice I. Il che equivale a dire che esiste una matrice invertibile P tale che

$$G = P^T P.$$

Quindi facendo variare P nell'insieme delle matrici invertibili, le matrici $P^T P$ descrivono tutte le matrici definite positive.

Esempio 18. *La forma quadratica su* $\mathbf{R}^2$ $q(\mathbf{u}) = x_1^2 + x_2^2$ *e' definita positiva; invece* $q(\mathbf{u}) = x_1^2$ *e' semidefinita positiva,* $q(\mathbf{u}) = -x_1^2 - x_2^2$ *e' definita negativa,* $q(\mathbf{u}) = -x_1^2$ *e' semidefinita negativa,* $q(\mathbf{u}) = x_1^2 - x_2^2$ *e' indefinita.* ■

Esempio 19. *La forma quadratica su* $\mathbf{R}^3$ $q(\mathbf{u}) = x_1^2 + x_2^2 + x_3^2$ *e' definita positiva; la forma* $q(\mathbf{u}) = -x_1^2 - x_2^2 - x_3^2$ *e' definita negativa; le forme* $q(\mathbf{u}) = x_1^2 + x_2^2$ *e* $q(\mathbf{u}) = x_1^2$ *sono semidefinite positive; le forme* $q(\mathbf{u}) = -x_1^2 - x_2^2$ *e* $q(\mathbf{u}) = -x_1^2$ *sono semidefinite negative; le forme* $q(\mathbf{u}) = x_1^2 + x_2^2 - x_3^2$, $q(\mathbf{u}) = x_1^2 - x_2^2 - x_3^2$, $q(\mathbf{u}) = x_1^2 - x_2^2$ *sono indefinite.* ■

Esempio 20. *(i) Lo spazio* $(\mathbf{R}^n, \textit{prodotto punto})$ *e' definito positivo.*

(ii) Lo spazio $(\mathcal{V}_O, \textit{prodotto scalare geometrico})$ *e' definito positivo.*

(iii) Lo spazio $(\mathbf{R}^4, \phi)$*, con coordinate* x_1, x_2, x_3, t *e* $q(\mathbf{u}) := x_1^2 + x_2^2 + x_3^2 - t^2$ *e' indefinito. Tale spazio e' detto spazio di Minkowski.* ■

Tra gli spazi pseudoeuclidei risultano di grande interesse gli spazi definiti positivi, perche' sono quelli che piu' si avvicinano agli spazi geometrici. Uno spazio siffatto dicesi *spazio euclideo*. Dedicheremo il prossimo paragrafo agli spazi euclidei. Per il momento ci limitiamo alle seguenti considerazioni. Per riconoscere se uno spazio pseudoeuclideo e' oppure no uno spazio euclideo possiamo applicare l'Algoritmo di Gauss-Lagrange. Ma ci sono altri criteri. Tra questi ricordiamo il seguente

Criterio dei minori principali. *Sia G una matrice reale simmetrica. Allora*

(1) G e' definita positiva se e solo se tutti i suoi minori principali hanno determinante positivo;

(2) G e' definita negativa se e solo se tutti i suoi minori principali di ordine dispari hanno determinante negativo, e quelli di ordine pari hanno determinante positivo.

Il *minore principale di G di ordine i* e' la sottomatrice di G formata dagli elementi comuni alle prime i righe ed alle prime i colonne. Per esempio i minori principali della matrice

$$\begin{bmatrix} g_{11} & g_{12} & g_{13} \\ g_{21} & g_{22} & g_{23} \\ g_{31} & g_{32} & g_{33} \end{bmatrix}$$

sono

$$[\,g_{11}\,]\,,\qquad \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix},\qquad \begin{bmatrix} g_{11} & g_{12} & g_{13} \\ g_{21} & g_{22} & g_{23} \\ g_{31} & g_{32} & g_{33} \end{bmatrix}.$$

Un altro criterio interessante e' il seguente

Criterio di Jacobi. *Sia G una matrice reale simmetrica. Allora G e' definita positiva se e solo se tutti i coefficienti del polinomio caratteristico di G sono non nulli ed hanno segni alterni.*

Esempio 21. *Si consideri la seguente forma quadratica $q : \mathbf{R}^3 \to \mathbf{R}$ definita ponendo*

$$q(\mathbf{u}) := 6x_1^2 - 2x_1x_2 + 6x_1x_3 + 2x_2^2 + 2x_2x_3 + 6x_3^2.$$

Provare che $(\mathbf{R}^3, q)$ e' uno spazio euclideo, e calcolarne una base ortonormale.

Svolgimento. Si osservi che la matrice di Gram di q rispetto alla base canonica e'

$$G_{\mathcal{E}}^{\mathcal{E}}(q) = \begin{bmatrix} 6 & -1 & 3 \\ -1 & 2 & 1 \\ 3 & 1 & 6 \end{bmatrix}.$$

Applicando l'Algoritmo di Gauss-Lagrange alla matrice

$$\begin{bmatrix} 6 & -1 & 3 & 1 & 0 & 0 \\ -1 & 2 & 1 & 0 & 1 & 0 \\ 3 & 1 & 6 & 0 & 0 & 1 \end{bmatrix}$$

si perviene alla seguente matrice

$$\begin{bmatrix} 6 & 0 & 0 & 1 & 0 & 0 \\ 0 & \frac{11}{6} & 0 & \frac{1}{6} & 1 & 0 \\ 0 & 0 & \frac{36}{11} & -\frac{7}{11} & -\frac{9}{11} & 1 \end{bmatrix}.$$

Cio' prova che $(\mathbf{R}^3, q)$ e' uno spazio euclideo in quanto $p = r = 3 = \dim \mathbf{R}^3$. Una base ortonormale di tale spazio e' data dai vettori:

$$\frac{1}{\sqrt{6}}\mathbf{e}_1,\quad \sqrt{\frac{6}{11}}\left(\frac{1}{6}\mathbf{e}_1 + \mathbf{e}_2\right),\quad \sqrt{\frac{11}{36}}\left(-\frac{7}{11}\mathbf{e}_1 - \frac{9}{11}\mathbf{e}_2 + \mathbf{e}_3\right).$$

Se $\mathbf{x}'$ denotano le coordinate rispetto a tale base si ha $q(\mathbf{u}) = {x'_1}^2 + {x'_2}^2 + {x'_3}^2$. ■

Esempio 22. *Provare che la matrice*

$$G := \begin{bmatrix} 6 & -1 & 3 \\ -1 & 2 & 1 \\ 3 & 1 & 6 \end{bmatrix}$$

e' definita positiva.

Svolgimento. *Applicando l'Algoritmo di Gauss-Lagrange a G come nell'esempio precedente, si arriva alla matrice*

$$\begin{bmatrix} 6 & 0 & 0 \\ 0 & \frac{11}{6} & 0 \\ 0 & 0 & \frac{36}{11} \end{bmatrix}.$$

Quindi l'indice di G e' 3, e cio' prova che $G > 0$.

Altrimenti avremmo potuto applicare il Criterio dei minori principali. Poiche'

$$\det[6] = 6 > 0, \quad \det \begin{bmatrix} 6 & -1 \\ -1 & 2 \end{bmatrix} = 11 > 0, \quad e \quad \det \begin{bmatrix} 6 & -1 & 3 \\ -1 & 2 & 1 \\ 3 & 1 & 6 \end{bmatrix} = 36 > 0$$

allora per il Criterio dei minori principali possiamo dire che $G > 0$.

Oppure possiamo calcolare il polinomio caratteristico di G, che e'

$$p_G(t) = -t^3 + 14t^2 - 49t + 36.$$

Da qui si deduce che $G > 0$ per il Criterio di Jacobi. ■

Esempio 23. *Provare che comunque si assegnino tre numeri reali x, y, z non tutti nulli allora*

$$6x^2 - 2xy + 6xz + 2y^2 + 2yz + 6z^2 > 0.$$

Svolgimento. E' sufficiente osservare che l'espressione

$$q(x, y, z) = 6x^2 - 2xy + 6xz + 2y^2 + 2yz + 6z^2$$

e' la forma quadratica dell'Esempio 21, che e' definita positiva. ■

6. Spazi euclidei.

In questo paragrafo studieremo in modo piu' approfondito gli spazi euclidei. Ricordiamo la definizione. Uno spazio pseudoeuclideo (V, ϕ) si dice *euclideo* se e' definito positivo, cioe' se $\phi(\mathbf{u}, \mathbf{u}) > 0$ per ogni $\mathbf{u} \neq \mathbf{0}$ in V. Per uno spazio euclideo cambieremo la notazione, cioe' scriveremo

$$\langle \mathbf{u}, \mathbf{v} \rangle := \phi(\mathbf{u}, \mathbf{v}),$$

e diremo che $\langle \mathbf{u}, \mathbf{v} \rangle$ e' il *prodotto scalare di* $\mathbf{u}$ *con* $\mathbf{v}$ *nello spazio euclideo* $(V, \langle\ ,\ \rangle)$. Con questa nuova notazione possiamo dire che in uno spazio euclideo valgono le seguenti regole di calcolo:

- $\langle a\mathbf{u} + b\mathbf{v}, \mathbf{w} \rangle = a\langle \mathbf{u}, \mathbf{w} \rangle + b\langle \mathbf{v}, \mathbf{w} \rangle$ per ogni $a, b \in \mathbf{R}$ ed ogni $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$;
- $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$ per ogni $\mathbf{u}, \mathbf{v} \in V$;
- $\langle \mathbf{u}, \mathbf{0} \rangle = 0$ per ogni $\mathbf{u} \in V$;

- $\langle \mathbf{u}, \mathbf{u} \rangle > 0$ per ogni $\mathbf{u} \neq \mathbf{0}$ in V;
- $\langle \mathbf{u} + \mathbf{v}, \mathbf{u} + \mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{u} \rangle + \langle \mathbf{v}, \mathbf{v} \rangle + 2\langle \mathbf{u}, \mathbf{v} \rangle$ per ogni $\mathbf{u}$ e $\mathbf{v}$ in V.

Grazie alla positivita', in uno spazio euclideo si puo' definire la *norma* $\|\mathbf{u}\|$ di un vettore $\mathbf{u}$ (detta anche *lunghezza* di $\mathbf{u}$) ponendo per definizione

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

Siano ora $\mathbf{u}, \mathbf{v}$ vettori in $(V, \langle\ ,\ \rangle)$. Diremo che $\mathbf{u}, \mathbf{v}$ sono *ortogonali tra loro* se $\langle \mathbf{u}, \mathbf{v} \rangle = 0$, ed in tal caso scriveremo

$$\mathbf{u} \perp \mathbf{v}.$$

Diremo che $\mathbf{v}$ e' *parallelo* ad $\mathbf{u}$ se $\mathbf{v}$ e' un multiplo di $\mathbf{u}$, ed in tal caso scriveremo

$$\mathbf{v} \parallel \mathbf{u}.$$

Si osservi che la definizione di parallelismo si puo' dare in un qualsiasi spazio vettoriale, quella di ortogonalita' in un qualunque spazio pseudoeuclideo, mentre per dare la definizione di norma e' necessario uno spazio euclideo.

Sia $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ una base di $(V, \langle\ ,\ \rangle)$. Utilizzando le definizioni precedenti possiamo dire che $\mathcal{B}$ e' ortogonale se e solo se i vettori di $\mathcal{B}$ sono mutualmente ortogonali tra loro, cioe' quando $\langle \mathbf{b}_i, \mathbf{b}_j \rangle = 0$ per ogni $i \neq j$. Invece $\mathcal{B}$ e' ortonormale se e solo se $\mathcal{B}$ e' ortogonale ed i vettori di $\mathcal{B}$ sono tutti di lunghezza 1. Sappiamo che in tal caso $G_{\mathcal{B}}^{\mathcal{B}}(\langle\quad\rangle) = I$, ed usando le coordinate rispetto a $\mathcal{B}$ abbiamo

$$\langle \mathbf{u}, \mathbf{v} \rangle = x_1 y_1 + \cdots + x_n y_n = \mathbf{x}^T \mathbf{y},$$

$$\langle \mathbf{u}, \mathbf{u} \rangle = x_1^2 + \cdots + x_n^2 = \mathbf{x}^T \mathbf{x},$$

$$\|\mathbf{u}\| = \sqrt{x_1^2 + \cdots + x_n^2} = \sqrt{\mathbf{x}^T \mathbf{x}}.$$

Esempio 24. *(i) Lo spazio numerico $\mathbf{R}^n$ con il prodotto punto e' uno spazio euclideo. In questo caso la base canonica e' una base ortonormale. A meno di un esplicito avviso, tutte le volte che useremo un prodotto scalare tra vettori numerici ci riferiremo al prodotto punto.*

(ii) Lo spazio $(\mathcal{V}_O, \text{prodotto scalare geometrico})$ e' uno spazio euclideo. In questo caso la norma coincide con la lunghezza "fisica"dei vettori geometrici. Due vettori geometrici sono ortogonali tra loro se e solo se formano un angolo retto (cioe' di 90^0). E sono paralleli se e solo se stanno sulla stessa retta passante per il punto comune di applicazione O.

(iii) Lo spazio $(\mathbf{R}^3, \langle\quad\rangle)$ con

$$\langle \mathbf{u}, \mathbf{u} \rangle = 6x_1^2 - 2x_1x_2 + 6x_1x_3 + 2x_2^2 + 2x_2x_3 + 6x_3^2$$

e' euclideo. Si osservi che in questo caso la base canonica non e' una base ortonormale. Infatti abbiamo, per esempio,

$$\langle \mathbf{e}_1, \mathbf{e}_2 \rangle = -1, \quad \|\mathbf{e}_1\| = \sqrt{6}.$$

In uno spazio euclideo, oltre alla norma, si puo' anche definire *l'angolo tra due vettori.* Cio' e' una conseguenza della seguente importante

Diseguaglianza di Cauchy-Schwarz. *Siano* $\mathbf{u}, \mathbf{v}$ *vettori in uno spazio euclideo. Allora*

$$|\langle \mathbf{u}, \mathbf{v} \rangle| \leq \|\mathbf{u}\| \|\mathbf{v}\|.$$

Dimostrazione. Possiamo assumere $\mathbf{u} \neq \mathbf{0}$. Ora poiche' siamo in uno spazio euclideo allora per ogni $x \in \mathbf{R}$ si ha

$$\langle x\mathbf{u} + \mathbf{v}, x\mathbf{u} + \mathbf{v} \rangle \geq 0.$$

Cioe' per ogni $x \in \mathbf{R}$ si ha

$$\|\mathbf{u}\|^2 x^2 + 2x\langle \mathbf{u}, \mathbf{v} \rangle + \|\mathbf{v}\|^2 \geq 0.$$

Possiamo riguardare l'espressione precedente come una disequazione di secondo grado in x con coefficiente direttore $\|\mathbf{u}\|^2 > 0$, che assume sempre valori ≥ 0. Cio' accade solo quando il discriminante Δ e' ≤ 0, cioe' quando

$$\Delta = 4\langle \mathbf{u}, \mathbf{v} \rangle^2 - 4\|\mathbf{u}\|^2 \|\mathbf{v}\|^2 \leq 0. \quad \blacksquare$$

Siano $\mathbf{u}, \mathbf{v}$ vettori non nulli in uno spazio euclideo. Dalla Diseguaglianza di Cauchy-Schwarz deduciamo che

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

Quindi esiste un unico angolo α di misura compresa tra 0 e π tale che

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

Tale angolo si chiama *l'angolo compreso tra* $\mathbf{u}$ *e* $\mathbf{v}$, e lo denoteremo con

$$\widehat{\mathbf{uv}} := \alpha.$$

Osservazione 2. *(i) Nel caso del prodotto scalare tra vettori geometrici la nozione di angolo appena data coincide con quella di angolo tra vettori geometrici.*

(ii) Si osservi che $\mathbf{u} \perp \mathbf{v}$ *se e solo se* $\alpha = \frac{\pi}{2}$.

(iii) Si ha anche $\mathbf{v} \parallel \mathbf{u}$ *se e solo se* $\alpha \in \{0, \pi\}$*. Infatti dire che* $\mathbf{v} \parallel \mathbf{u}$ *equivale a dire che esiste un* $x \in \mathbf{R}$ *tale che* $x\mathbf{u} + \mathbf{v} = \mathbf{0}$*. Cio' a sua volta equivale a dire che la disequazione che appare nella dimostrazione della Diseguaglianza di Cauchy-Schwarz ha discriminante nullo, cioe' che* $|\langle \mathbf{u}, \mathbf{v} \rangle| = \|\mathbf{u}\| \|\mathbf{v}\|$*.*

(iv) Se si introducono coordinate rispetto ad una base ortonormale la Diseguaglianza di Cauchy-Schwarz equivale a dire che per ogni $\mathbf{x}, \mathbf{y} \in \mathbf{R}^n$ *si ha:*

$$\left| \sum_{i=1}^{n} x_i y_i \right| \leq \left(\sum_{i=1}^{n} x_i^2 \right)^{\frac{1}{2}} \cdot \left(\sum_{i=1}^{n} y_i^2 \right)^{\frac{1}{2}},$$

che e' una diseguaglianza numerica (apparentemente) non banale.

Un'altra conseguenza della Diseguaglianza di Cauchy-Schwarz e' la *Diseguaglianza triangolare* che afferma che in uno spazio euclideo continua a valere una proprieta' ben nota della geometria dei triangoli, e cioe' che in un triangolo la lunghezza di un lato e' minore o uguale alla somma delle lunghezze degli altri due lati.

Diseguaglianza triangolare. *Siano* $\mathbf{u}, \mathbf{v}$ *vettori in uno spazio euclideo. Allora*

$$\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|.$$

Dimostrazione. Basta osservare che

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

Nell'espressione precedente la Diseguaglianza di Cauchy-Schwarz interviene per consentire il passaggio con il $\leq$. ■

Come corollario otteniamo le importanti

Proprieta' della norma in uno spazio euclideo. *Sia* $(V, \langle \ , \ \rangle)$ *uno spazio euclideo. Allora valgono le seguenti proprieta':*

1) per ogni $\mathbf{u} \in V$ *si ha* $\|\mathbf{u}\| \geq 0$*, e* $\|\mathbf{u}\| = 0$ *se e solo se* $\mathbf{u} = \mathbf{0}$*;*

2) per ogni $\mathbf{u} \in V$ *ed ogni* $c \in \mathbf{R}$ *si ha* $\|c\mathbf{u}\| = |c|\|\mathbf{u}\|$*;*

3) per ogni $\mathbf{u}, \mathbf{v} \in V$ *si ha* $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$*.*

Esempio 25. *Se* $\mathbf{u}$ *e' un vettore non nullo allora*

$$\frac{\mathbf{u}}{\|\mathbf{u}\|}$$

e' un vettore parallelo ad $\mathbf{u}$ *di lunghezza* 1*; inoltre in Span*$(\mathbf{u})$ *ci sono solo due vettori di lunghezza* 1*, e cioe'*

$$\frac{\mathbf{u}}{\|\mathbf{u}\|} \quad e \quad -\frac{\mathbf{u}}{\|\mathbf{u}\|}.$$

Tali vettori si chiamano anche versori di Span($\mathbf{u}$).

Esempio 26. *Nello spazio euclideo* ($\mathbf{R}^3$, *prodotto punto*) *calcolare l'angolo* α *tra i vettori* $(1,0,0)$ *e* $(3,0,3)$.

Svolgimento. Sappiamo che

$$\cos\alpha = \frac{\langle \mathbf{u}, \mathbf{v}\rangle}{\|\mathbf{u}\|\|\mathbf{v}\|} = \frac{3}{3\sqrt{2}} = \frac{\sqrt{2}}{2}.$$

Quindi $\alpha = \frac{\pi}{4}$. ∎

Esempio 27. *Fare un esempio di spazio euclideo* $(\mathbf{R}^2, \langle\ \ \rangle)$ *in cui i vettori canonici hanno lunghezza* 1 *ma formano un angolo di* 120^0.

Svolgimento. In questo spazio deve essere $\langle \mathbf{e}_1, \mathbf{e}_1\rangle = 1$ *e* $\langle \mathbf{e}_2, \mathbf{e}_2\rangle = 1$*; inoltre* $\cos 120^0 = \langle \mathbf{e}_1, \mathbf{e}_2\rangle$, *cioe'* $\langle \mathbf{e}_1, \mathbf{e}_2\rangle = -\frac{1}{2}$. *Dunque la matrice di Gram rispetto alla base canonica e' determinata, cioe'*

$$G^{\mathcal{E}}_{\mathcal{E}}(\langle\ \ \rangle) = \begin{bmatrix} \langle \mathbf{e}_1, \mathbf{e}_1\rangle & \langle \mathbf{e}_1, \mathbf{e}_2\rangle \\ \langle \mathbf{e}_2, \mathbf{e}_1\rangle & \langle \mathbf{e}_2, \mathbf{e}_2\rangle \end{bmatrix} = \begin{bmatrix} 1 & -\frac{1}{2} \\ -\frac{1}{2} & 1 \end{bmatrix}.$$

Tale matrice e' definita positiva per il Criterio dei minori principali, quindi l'esempio richiesto e':

$$\langle \mathbf{u}, \mathbf{v}\rangle = x_1y_1 - \frac{1}{2}x_1y_2 - \frac{1}{2}x_2y_1 + x_2y_2. \quad \blacksquare$$

7. Proiezione ortogonale su un vettore in uno spazio euclideo.

Una tipica costruzione geometrica consiste nel proiettare un vettore geometrico su una retta. Piu' precisamente, sia $\mathbf{v} := \overrightarrow{OP}$ un vettore geometrico nello spazio euclideo $\mathcal{V}_O$ dei vettori geometrici applicati in un fissato punto O, e sia r una retta per O. Allora $\overrightarrow{OP}$ ed r individuano un piano (a meno che P non stia su r), ed in tale piano esiste un'unica retta per P ortogonale ad r. Tale retta interseca r in un punto H, ed il vettore geometrico $\overrightarrow{OH}$ (che giace su r) dicesi *la proiezione ortogonale di* $\overrightarrow{OP}$ *su* r (se $P \in r$ la proiezione ortogonale di $\overrightarrow{OP}$ su r e' proprio $\overrightarrow{OP}$). Si osservi che il vettore $\overrightarrow{OP} - \overrightarrow{OH}$ e' ortogonale ad r, sicche' possiamo scrivere il vettore $\overrightarrow{OP}$ come somma di un vettore parallelo ad r e di un vettore ortogonale ad r:

$$\overrightarrow{OP} = \overrightarrow{OH} + (\overrightarrow{OP} - \overrightarrow{OH}).$$

Se $\mathbf{u} \in \mathcal{V}_O$ e' un qualsiasi vettore non nullo giacente su r, allora $\overrightarrow{OH}$ sara' un multiplo di $\mathbf{u}$, cioe' sara' del tipo

$$\overrightarrow{OH} = c\mathbf{u}.$$

Dalla definizione di funzione coseno segue che $c = \frac{\|\mathbf{v}\|}{\|\mathbf{u}\|} \cos \widehat{\mathbf{u}\mathbf{v}}$, cioe'

$$c = \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\langle \mathbf{u}, \mathbf{u} \rangle}.$$

Questa formula consente di dare la definizione di proiezione ortogonale in un qualsiasi spazio munito di prodotto scalare, cioe' in un qualsiasi spazio euclideo. E' l'argomento di cui ci occuperemo nel presente paragrafo.

Sia $(V, \langle\ ,\ \rangle)$ uno spazio euclideo, e siano $\mathbf{u} \neq \mathbf{0}$ e $\mathbf{v}$ vettori di V. Si definisce *proiezione ortogonale di* $\mathbf{v}$ *su* $\mathbf{u}$ il vettore

$$p_{\mathbf{u}}(\mathbf{v}) := \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\langle \mathbf{u}, \mathbf{u} \rangle}\mathbf{u}.$$

Lo scalare

$$c = \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\langle \mathbf{u}, \mathbf{u} \rangle}$$

che appare nella formula precedente si chiama *il coefficiente di Fourier di* $\mathbf{v}$ *rispetto al vettore* $\mathbf{u}$.

Osservazione 3. *(i) Per la stessa definizione di proiezione ortogonale si ha* $p_{\mathbf{u}}(\mathbf{v}) \in Span(\mathbf{u})$*, cioe'* $p_{\mathbf{u}}(\mathbf{v}) \parallel \mathbf{u}$*;*

(ii) $p_{\mathbf{u}}(\mathbf{v}) = \mathbf{v}$ *se e solo se* $\mathbf{v} \in Span(\mathbf{u})$*;*

(iii) $p_{\mathbf{u}}(\mathbf{v}) = \mathbf{0}$ *se e solo se* $\mathbf{v} \perp \mathbf{u}$*;*

(iv) possiamo scrivere

$$\mathbf{v} = p_{\mathbf{u}}(\mathbf{v}) + (\mathbf{v} - p_{\mathbf{u}}(\mathbf{v}));$$

l'interesse di tale decomposizione risiede nel fatto che $p_{\mathbf{u}}(\mathbf{v})$ *e' parallelo ad* $\mathbf{u}$*, mentre* $\mathbf{v} - p_{\mathbf{u}}(\mathbf{v})$ *e' ortogonale a* $\mathbf{u}$*; infatti abbiamo*

$$\langle \mathbf{v} - p_{\mathbf{u}}(\mathbf{v}), \mathbf{u} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle - \langle p_{\mathbf{u}}(\mathbf{v}), \mathbf{u} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle - \langle c\mathbf{u}, \mathbf{u} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle - c\langle \mathbf{u}, \mathbf{u} \rangle = 0.$$

Corollario 5. *Sia* $\mathbf{u}$ *un vettore non nullo di uno spazio euclideo* V*. Allora ogni vettore* $\mathbf{v} \in V$ *si puo' decomporre nella somma*

$$\mathbf{v} = \mathbf{a} + \mathbf{b}$$

con $\mathbf{a} \parallel \mathbf{u}$ *e* $\mathbf{b} \perp \mathbf{u}$*. Tale decomposizione e' unica, ed* $\mathbf{a} = p_{\mathbf{u}}(\mathbf{v})$ *e* $\mathbf{b} = \mathbf{v} - p_{\mathbf{u}}(\mathbf{v})$*.*

Dimostrazione. Per quanto detto nell'osservazione precedente rimane solo da provare l'unicita'. Supponiamo allora che

$$\mathbf{v} = \mathbf{a} + \mathbf{b} = \mathbf{a}' + \mathbf{b}'$$

con $\mathbf{a}$ e $\mathbf{a}'$ paralleli ad $\mathbf{u}$, e $\mathbf{b}$ e $\mathbf{b}'$ ortogonali a $\mathbf{u}$. Poniamo

$$\mathbf{z} := \mathbf{a} - \mathbf{a}' = \mathbf{b}' - \mathbf{b}.$$

Il vettore $\mathbf{z}$ in quanto uguale ad $\mathbf{a} - \mathbf{a}'$ e' parallelo ad $\mathbf{u}$. Quindi $\mathbf{z} = x\mathbf{u}$ per qualche $x \in \mathbf{R}$. Ma essendo anche $\mathbf{z} = \mathbf{b}' - \mathbf{b}$ allora $\mathbf{z}$ e' anche ortogonale ad $\mathbf{u}$. Infatti

$$\langle \mathbf{z}, \mathbf{u} \rangle = \langle \mathbf{b}' - \mathbf{b}, \mathbf{u} \rangle = \langle \mathbf{b}', \mathbf{u} \rangle - \langle \mathbf{b}, \mathbf{u} \rangle = 0 - 0 = 0.$$

Quindi

$$0 = \langle \mathbf{z}, \mathbf{u} \rangle = \langle x\mathbf{u}, \mathbf{u} \rangle = x\langle \mathbf{u}, \mathbf{u} \rangle.$$

Poiche' $\mathbf{u} \neq \mathbf{0}$, deduciamo che $x = 0$, quindi $\mathbf{z} = \mathbf{0}$, e cioe' $\mathbf{a} = \mathbf{a}'$ e $\mathbf{b} = \mathbf{b}'$. ■

Esempio 28. *Nello spazio euclideo* ($\mathbf{R}^3$, *prodotto punto*) *calcolare la proiezione ortogonale di* $(1, 1, 1)$ *sul vettore* $(-1, 0, 2)$.

Svolgimento. Abbiamo

$$p_{(-1,0,2)}(1, 1, 1) = \frac{\langle (1, 1, 1), (-1, 0, 2) \rangle}{\langle (-1, 0, 2), (-1, 0, 2) \rangle}(-1, 0, 2) = \frac{1}{5}(-1, 0, 2). \quad ■$$

Esempio 29. *Nello spazio euclideo* ($\mathbf{R}^3$, *prodotto punto*) *decomporre* $(1, 1, 1)$ *nella somma di un vettore parallelo a* $(-1, 0, 2)$ *e di un vettore ortogonale a* $(-1, 0, 2)$.

Svolgimento. Sappiamo che tale decomposizione e'

$$(1, 1, 1) = p_{(-1,0,2)}(1, 1, 1) + \left[(1, 1, 1) - p_{(-1,0,2)}(1, 1, 1)\right] = \frac{1}{5}(-1, 0, 2) + \frac{1}{5}(6, 5, 3). \quad ■$$

Esempio 30. *Nello spazio euclideo* $(\mathbf{R}^3, \langle\ \ \rangle)$ *definito dalla forma bilineare simmetrica*

$$\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{x}^T \begin{bmatrix} 6 & -1 & 3 \\ -1 & 2 & 1 \\ 3 & 1 & 6 \end{bmatrix} \mathbf{y}$$

calcolare la proiezione ortogonale di $(1, 1, 1)$ *sul vettore* $(-1, 0, 2)$.

Svolgimento. A differenza dell'esempio precedente, adesso il prodotto scalare non e' il prodotto punto, ma e' definito da una nuova forma bilineare. Quindi abbiamo

$$p_{(-1,0,2)}(1, 1, 1) = \frac{\langle (1, 1, 1), (-1, 0, 2) \rangle}{\langle (-1, 0, 2), (-1, 0, 2) \rangle}(-1, 0, 2) = \frac{12}{18}(-1, 0, 2) = \frac{2}{3}(-1, 0, 2). \quad ■$$

Esempio 31. *Sia* $(V, \langle\ ,\ \rangle)$ *uno spazio euclideo,* $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ *una base ortonormale di* V*,* $\mathbf{v}$ *un vettore in* V*, ed* $x_1, \dots, x_n$ *siano le coordinate di* $\mathbf{v}$ *rispetto alla base* $\mathcal{B}$*. Quindi abbiamo* $\mathbf{v} = x_1\mathbf{b}_1 + \dots + x_n\mathbf{b}_n$*. Provare che*

$$x_i\mathbf{b}_i = p_{\mathbf{b}_i}(\mathbf{v}),$$

cioe' che

$$\mathbf{v} = p_{\mathbf{b}_1}(\mathbf{v}) + \dots + p_{\mathbf{b}_n}(\mathbf{v}).$$

In particolare la coordinata x_i *e' il coefficiente di Fourier di* $\mathbf{v}$ *rispetto a* $\mathbf{b}_i$*. Inoltre se* $\mathbf{v}$ *e' un versore (cioe' un vettore di lunghezza* 1*), allora* $x_i = \cos\widehat{\mathbf{b}_i\mathbf{v}}$*, e percio' in tal caso le coordinate di* $\mathbf{v}$ *si chiamano anche i coseni direttori di* $\mathbf{v}$ *rispetto alla base ortonormale* $\mathcal{B}$*.*

Svolgimento. Tenuto conto che $\|\mathbf{b}_i\| = 1$*, e' sufficiente osservare che*

$$\langle \mathbf{v}, \mathbf{b}_i\rangle = \langle \sum_{j=1}^{n} x_j\mathbf{b}_j, \mathbf{b}_i\rangle = \sum_{j=1}^{n} x_j\langle \mathbf{b}_j, \mathbf{b}_i\rangle = x_i\langle \mathbf{b}_i, \mathbf{b}_i\rangle = x_i. \quad \blacksquare$$

(Questo esempio mostra che in uno spazio euclideo qualsiasi le coordinate rispetto ad una base ortonormale si possono interpretare come nel caso dei vettori geometrici: cioe' si ottengono effettuando le proiezioni ortogonali lungo gli assi della base.)

8. Il procedimento di ortogonalizzazione di Gram-Schmidt.

Sia (V, ϕ) uno spazio pseudoeuclideo, e $\mathcal{B}$ una base. Abbiamo visto che, grazie all'Algoritmo di Gauss-Lagrange, a partire da $\mathcal{B}$ si puo' calcolare una base ortonormale di (V, ϕ). Nel caso in cui lo spazio e' anche euclideo esiste un altro algoritmo per calcolare una base ortonormale, ed e' il cosiddetto *procedimento di ortogonalizzazione di Gram-Schmidt.* Si puo' dimostrare che il procedimento di ortogonalizzazione di Gram-Schmidt produce lo stesso risultato dell'Algoritmo di Gauss-Lagrange, che pero' e' piu' generale perche' si puo' applicare anche in uno spazio non definito positivo. Nonostante cio', il procedimento di ortogonalizzazione di Gram-Schmidt ha un grande interesse perche' si basa sulla nozione di proiezione ortogonale su un vettore.

Cominciamo a vedere come funziona nel caso di una base formata da due vettori. Poi vedremo come funziona in generale. Sia dunque $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ una base per uno spazio euclideo $(V, \langle\ ,\ \rangle)$ di dimensione 2. Ora poniamo

$$\mathbf{c}_1 := \mathbf{b}_1.$$

Poi abbiamo visto nel Corollario 5 che il vettore

$$\mathbf{c}_2 := \mathbf{b}_2 - p_{\mathbf{c}_1}(\mathbf{b}_2)$$

e' ortogonale a $\mathbf{c}_1$. Dunque il sistema di vettori

$$\mathbf{c}_1, \mathbf{c}_2$$

forma una base ortogonale per V (confrontare con il Lemma 3 dopo). Questo e' il procedimento di ortogonalizzazione di Gram-Schmidt. Se si vuole una base ortonormale occorre normalizzare i vettori ottenuti, cioe' una base ortonormale e' costituita dai vettori

$$\frac{\mathbf{c}_1}{\|\mathbf{c}_1\|}, \frac{\mathbf{c}_2}{\|\mathbf{c}_2\|}.$$

Nel caso generale di dimensione $n \geq 2$, il procedimento di ortogonalizzazione di Gram-Schmidt funziona cosi'. Si parte da una base qualsiasi $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$, dopodiche' a partire dai vettori di $\mathcal{B}$ si costruiscono i vettori $\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n$ nel seguente modo:

$$\begin{aligned}
&\mathbf{c}_1 := \mathbf{b}_1 \\
&\mathbf{c}_2 := \mathbf{b}_2 - p_{\mathbf{c}_1}(\mathbf{b}_2) \\
&\mathbf{c}_3 := \mathbf{b}_3 - p_{\mathbf{c}_1}(\mathbf{b}_3) - p_{\mathbf{c}_2}(\mathbf{b}_3) \\
&\ldots\ldots\ldots \\
&\mathbf{c}_n := \mathbf{b}_n - p_{\mathbf{c}_1}(\mathbf{b}_n) - p_{\mathbf{c}_2}(\mathbf{b}_n) - \cdots - p_{\mathbf{c}_{n-1}}(\mathbf{b}_n).
\end{aligned}$$

I vettori cosi' costruiti formano una base ortogonale di V. Normalizzando si ottiene una base ortonormale:

$$\frac{\mathbf{c}_1}{\|\mathbf{c}_1\|}, \frac{\mathbf{c}_2}{\|\mathbf{c}_2\|}, \ldots\ldots, \frac{\mathbf{c}_n}{\|\mathbf{c}_n\|}.$$

Esempio 32. *Nello spazio euclideo* $(\mathbf{R}^3, prodotto\, punto)$ *si consideri il sottospazio* U *generato dai vettori* $(1,1,0)$, $(0,1,-1)$. *Calcolare una base ortonormale di* U.

Svolgimento. Innanzitutto osserviamo una cosa implicita nel testo dell'esercizio. Cioe' che in generale un sottospazio di uno spazio pseudoeuclideo (V, ϕ) *eredita in modo naturale una struttura di spazio pseudoeuclideo: basta semplicemente considerare la forma* ϕ *ristretta su* U. *E' ovvio che se* ϕ *e' definita positiva allora tutti i sottospazi di* V *saranno definiti positivi. In altre parole sottospazi di uno spazio euclideo sono anch'essi euclidei. Quindi nel nostro esempio abbiamo uno spazio* U *che e' euclideo, che avra' una base ortonormale. Andiamo a calcolarla applicando il procedimento di ortogonalizzazione di Gram-Schmidt alla base* $\{(1,1,0),(0,1,-1)\}$. *Poniamo quindi* $\mathbf{c}_1 = (1,1,0)$ *e poi*

$$\mathbf{c}_2 = (0,1,-1) - p_{(1,1,0)}(0,1,-1) = \frac{1}{2}(-1,1,-2).$$

Quindi una base ortogonale per U *e' data dai vettori* $(1,1,0), \frac{1}{2}(-1,1,-2)$. *Mentre una base ortonormale di* U *e' data da:*

$$\frac{1}{\sqrt{2}}(1,1,0), \frac{1}{\sqrt{6}}(-1,1,-2).$$

Avremmo potuto svolgere l'esercizio anche con l'Algoritmo di Gauss-Lagrange, procedendo come segue. Innanzitutto calcoliamo la matrice di Gram G *del prodotto punto ristretto su* U *rispetto alla base* $\{(1,1,0),(0,1,-1)\}$. *Si ha:*

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

Poi applichiamo l'Algoritmo di Gauss-Lagrange alla matrice

$$\begin{bmatrix} 2 & 1 & 1 & 0 \\ 1 & 2 & 0 & 1 \end{bmatrix}.$$

Eseguendo le operazioni elementari $e_{21}(-\frac{1}{2})$ *e* $e^{21}(-\frac{1}{2})$ *otteniamo*

$$\begin{bmatrix} 2 & 0 & 1 & 0 \\ 0 & \frac{3}{2} & -\frac{1}{2} & 1 \end{bmatrix}.$$

Quindi i vettori di U *che hanno coordinate* $(1,0)^T$ *e* $(-\frac{1}{2},1)^T$ *rispetto alla base data formano una base ortogonale per* U. *Tali vettori sono esattamente i vettori* $\mathbf{c}_1$ *e* $\mathbf{c}_2$ *calcolati in precedenza.* ■

Esempio 33. *Nello spazio euclideo* $(\mathbf{R}^3, \textit{prodotto punto})$ *si considerino i vettori ortogonali* $(1,0,1)$ *e* $(0,1,0)$. *Completare il sistema* $(1,0,1)$, $(0,1,0)$ *a base ortogonale di* $(\mathbf{R}^3, \textit{prodotto punto})$.

Svolgimento. Aggiungendo il vettore $(0,0,1)$ *otteniamo una base di* $\mathbf{R}^3$ *che non e' ortogonale:*

$$(1,0,1), (0,1,0), (0,0,1),$$

ma applicando il procedimento di ortogonalizzazione di Gram-Schmidt otteniamo la base cercata:

$$(1,0,1), (0,1,0), (-1,0,1). \quad ■$$

9. Proiezione ortogonale su un sottospazio di uno spazio euclideo.

Nel paragrafo 7 abbiamo imparato come, in uno spazio euclideo, si proietta un vettore $\mathbf{v}$ su un vettore non nullo $\mathbf{u}$. Ora consideriamo un vettore $\mathbf{u}'$ non nullo multiplo di $\mathbf{u}$, cioe' $\mathbf{u}' = a\mathbf{u}$, con $a \neq 0$. La proiezione di $\mathbf{v}$ su $\mathbf{u}'$ coincide con la proiezione su $\mathbf{u}$. Infatti

$$p_{\mathbf{u}'}(\mathbf{v}) = \frac{\langle a\mathbf{u}, \mathbf{v}\rangle}{\langle a\mathbf{u}, a\mathbf{u}\rangle}(a\mathbf{u}) = \frac{\langle \mathbf{u}, \mathbf{v}\rangle}{\langle \mathbf{u}, \mathbf{u}\rangle}\mathbf{u} = p_{\mathbf{u}}(\mathbf{v}).$$

Cio' rende ben posta la seguente definizione.

Sia $(V, \langle\ ,\ \rangle)$ uno spazio euclideo, ed U un sottospazio di V di dimensione 1. Sia $\mathbf{u}$ un vettore non nullo di U (cioe' una base di U, che, in questo caso, e' anche una base ortogonale di U). Sia poi $\mathbf{v}$ un qualunque vettore di V. Allora si definisce *proiezione ortogonale di* $\mathbf{v}$ *su* U il seguente vettore:

$$p_U(\mathbf{v}) := p_{\mathbf{u}}(\mathbf{v}).$$

Piu' in generale si puo' dare la seguente definizione di *proiezione ortogonale* $p_U(\mathbf{v})$ *di un vettore* $\mathbf{v} \in V$ *su un sottospazio* U *di* V *di dimensione qualunque* $h \geq 1$. Si fissa una base **ortogonale** di U, diciamo $\{\mathbf{u}_1, \dots, \mathbf{u}_h\}$, e si pone per definizione

$$p_U(\mathbf{v}) := \sum_{i=1}^{h} p_{\mathbf{u}_i}(\mathbf{v}).$$

Proveremo tra poco che tale definizione e' ben posta, cioe' che il vettore $p_U(\mathbf{v})$ *non dipende* dalla base ortogonale di U scelta. Abbiamo bisogno di una nozione preliminare.

Dato un sottospazio U di V, si definisce *il complemento ortogonale* $U^\perp$ *di* U *in* V come quel sottoinsieme di V formato dai vettori $\mathbf{v}$ che sono ortogonali a tutti i vettori di U. Cioe'

$$U^\perp := \{\mathbf{v} \in V \,:\, \langle \mathbf{v}, \mathbf{u} \rangle = 0 \text{ per ogni } \mathbf{u} \in U\}\,.$$

Ora possiamo provare che

Proposizione 3. *Valgono le seguenti proprieta':*

1) $U^\perp$ *e' un sottospazio di* V*;*

2) se $U = Span(\mathbf{u}_1, \dots, \mathbf{u}_h)$ *allora*

$$U^\perp = \{\mathbf{v} \in V \,:\, \langle \mathbf{v}, \mathbf{u}_i \rangle = 0 \textit{ per ogni } i = 1, \dots, h\}\,;$$

3) $V = U \oplus U^\perp$*;*

4) la definizione di proiezione ortogonale su U *data in precedenza non dipende dalla base ortogonale di* U *scelta.*

Dimostrazione. Osserviamo che $\mathbf{0} \in U^\perp$ in quanto $\langle \mathbf{0}, \mathbf{u} \rangle = 0$ per ogni vettore $\mathbf{u} \in V$. Poi siano $\mathbf{v}$ e $\mathbf{w}$ vettori di $U^\perp$. Per ogni $\mathbf{u} \in U$ abbiamo

$$\langle \mathbf{v} + \mathbf{w}, \mathbf{u} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle + \langle \mathbf{w}, \mathbf{u} \rangle = 0 + 0 = 0.$$

Cio' prova la stabilita' di $U^\perp$ rispetto all'addizione. Analogamente si ha

$$\langle c\mathbf{v}, \mathbf{u} \rangle = c\langle \mathbf{v}, \mathbf{u} \rangle = c0 = 0$$

per ogni $c \in \mathbf{R}$, $\mathbf{v} \in U^\perp$ ed $\mathbf{u} \in U$. Cio' prova che $U^\perp$ e' stabile rispetto alla moltiplicazione esterna, ed avendo gia' osservato che $\mathbf{0} \in U^\perp$, possiamo dire di aver provato che $U^\perp$ e' un sottospazio di V.

Per provare la seconda proprieta' sara' sufficiente provare che *se* $\langle \mathbf{v}, \mathbf{u}_i \rangle = 0$ *per ogni* $i = 1, \dots, h$ *allora* $\langle \mathbf{v}, \mathbf{u} \rangle = 0$ *per ogni* $\mathbf{u} \in U$. Poiche' i vettori $\mathbf{u}_i$ generano U allora per opportuni pesi $a_i \in \mathbf{R}$ si deve avere $\mathbf{u} = a_1\mathbf{u}_1 + \cdots + a_h\mathbf{u}_h$. Per cui

$$\langle \mathbf{v}, \mathbf{u} \rangle = \langle \mathbf{v}, a_1\mathbf{u}_1 + \cdots + a_h\mathbf{u}_h \rangle = a_1\langle \mathbf{v}, \mathbf{u}_1 \rangle + \cdots + a_h\langle \mathbf{v}, \mathbf{u}_h \rangle = 0.$$

Cio' prova la proprieta' 2).

Ora dimostriamo la proprieta' 3). Per fare cio' fissiamo una base ortogonale di U, diciamo $\mathcal{B} = \{\mathbf{u}_1, \dots, \mathbf{u}_h\}$. Sia $\mathbf{v}$ un qualunque vettore di V. A partire dalla base ortogonale $\mathcal{B}$ abbiamo definito il vettore $p_U(\mathbf{v})$. Possiamo scrivere

$$(*) \qquad \mathbf{v} = p_U(\mathbf{v}) + (\mathbf{v} - p_U(\mathbf{v})).$$

Andiamo a provare che

$$\mathbf{v} - p_U(\mathbf{v}) \in U^\perp.$$

Per la proprieta' 2) sara' sufficiente provare che

$$\langle \mathbf{v} - p_U(\mathbf{v}), \mathbf{u}_i \rangle = 0$$

per ogni $i = 1, \dots, h$. A tale proposito cominciamo con l'osservare che

$$\langle \mathbf{v} - p_U(\mathbf{v}), \mathbf{u}_i \rangle = \langle \mathbf{v}, \mathbf{u}_i \rangle - \langle p_U(\mathbf{v}), \mathbf{u}_i \rangle = \langle \mathbf{v}, \mathbf{u}_i \rangle - \sum_{j=1}^{h} \langle p_{\mathbf{u}_j}(\mathbf{v}), \mathbf{u}_i \rangle.$$

Ora sappiamo che $p_{\mathbf{u}_j}(\mathbf{v})$ e' parallelo a $\mathbf{u}_j$, cioe' e' del tipo $p_{\mathbf{u}_j}(\mathbf{v}) = c_j \mathbf{u}_j$. Quindi se $j \neq i$ deve essere

$$\langle p_{\mathbf{u}_j}(\mathbf{v}), \mathbf{u}_i \rangle = c_j \langle \mathbf{u}_j, \mathbf{u}_i \rangle = 0$$

perche' $\mathcal{B}$ e' ortogonale. Riprendendo il calcolo precedente abbiamo allora

$$\langle \mathbf{v} - p_U(\mathbf{v}), \mathbf{u}_i \rangle = \langle \mathbf{v}, \mathbf{u}_i \rangle - \langle p_{\mathbf{u}_i}(\mathbf{v}), \mathbf{u}_i \rangle = \langle \mathbf{v} - p_{\mathbf{u}_i}(\mathbf{v}), \mathbf{u}_i \rangle = 0$$

per il Corollario 5. Cio' prova dunque che $\mathbf{v} - p_U(\mathbf{v}) \in U^\perp$. D'altra parte per definizione $p_U(\mathbf{v}) \in U$. Quindi potendo scrivere $\mathbf{v} = p_U(\mathbf{v}) + (\mathbf{v} - p_U(\mathbf{v}))$, possiamo dire di aver provato che ogni vettore di V e' somma di un vettore di U e di un vettore di $U^\perp$. Abbiamo dimostrato cioe' che

$$V = U + U^\perp.$$

Per concludere la dimostrazione della proprieta' 3) dobbiamo solo provare che la somma e' diretta. Ci sara' sufficiente provare che $U \cap U^\perp = \{\mathbf{0}\}$. Sia infatti $\mathbf{u}$ un vettore appartenente ad $U \cap U^\perp$. Allora $\langle \mathbf{u}, \mathbf{u} \rangle = 0$, e poiche' lo spazio e' euclideo allora $\mathbf{u} = \mathbf{0}$. Cio' conclude la dimostrazione della proprieta' 3).

Infine andiamo a dimostrare che la definizione di proiezione ortogonale su un sottospazio e' ben posta. Siano allora $\mathbf{v} \in V$ un vettore qualsiasi, e $\mathcal{B}$ e $\mathcal{B}'$ due basi ortogonali per U. Denotiamo con $\mathbf{a}$ la proiezione ortogonale di $\mathbf{v}$ su U definita a partire dalla base $\mathcal{B}$, e con $\mathbf{a}'$ la proiezione ortogonale di $\mathbf{v}$ su U definita a partire dalla base $\mathcal{B}'$. Dobbiamo provare che $\mathbf{a} = \mathbf{a}'$. Per definizione abbiamo $\mathbf{a}, \mathbf{a}' \in U$. Possiamo scrivere

$$\mathbf{v} = \mathbf{a} + (\mathbf{v} - \mathbf{a}) = \mathbf{a}' + (\mathbf{v} - \mathbf{a}')$$

ed abbiamo appena dimostrato nel corso della dimostrazione della proprieta' 3) che i vettori $\mathbf{v} - \mathbf{a}$ e $\mathbf{v} - \mathbf{a}'$ stanno in $U^\perp$. Quindi il vettore

$$\mathbf{z} := \mathbf{a} - \mathbf{a}' = (\mathbf{v} - \mathbf{a}') - (\mathbf{v} - \mathbf{a})$$

sta in $U \cap U^\perp$, per cui $\mathbf{z} = \mathbf{0}$. Cio' implica che $\mathbf{a} = \mathbf{a}'$, che e' quello che volevamo provare. ■

Sia U un sottospazio di uno spazio euclideo V, e $\mathbf{v} \in V$ un vettore. Diremo che $\mathbf{v}$ *e' parallelo ad* U se $\mathbf{v} \in U$, e scriveremo $\mathbf{v} \parallel U$. Mentre diremo che $\mathbf{v}$ *e' ortogonale ad* U se $\mathbf{v} \in U^\perp$, e scriveremo $\mathbf{v} \perp U$.

Corollario 6. *Sia U un sottospazio di uno spazio euclideo V. Allora valgono le seguenti proprieta':*

1) $(U^{\perp})^{\perp} = U$;

2) ogni vettore $\mathbf{v}$ *di* V *si puo' scrivere, in modo unico, come somma di un vettore* $\mathbf{a}$ *parallelo ad* U, *e di un vettore* $\mathbf{b}$ *ortogonale ad* U, *e si ha* $\mathbf{a} = p_U(\mathbf{v})$, *e* $\mathbf{b} = p_{U^{\perp}}(\mathbf{v})$, *cioe'*

$$\mathbf{v} = p_U(\mathbf{v}) + p_{U^{\perp}}(\mathbf{v});$$

3) $p_U(\mathbf{v}) = \mathbf{v}$ *se e solo se* $\mathbf{v} \parallel U$;

4) $p_U(\mathbf{v}) = \mathbf{0}$ *se e solo se* $\mathbf{v} \perp U$.

Dimostrazione. Per la proposizione precedente applicata al sottospazio $U^{\perp}$ sappiamo che $V = U^{\perp} \oplus (U^{\perp})^{\perp}$. Quindi

$$\dim U = \dim V - \dim U^{\perp} = \dim(U^{\perp})^{\perp}.$$

Ne consegue $U = (U^{\perp})^{\perp}$ perche' e' evidente che U e' contenuto in $(U^{\perp})^{\perp}$.

Circa la seconda proprieta', poiche' $V = U \oplus U^{\perp}$ allora e' chiaro che $\mathbf{v} = \mathbf{a} + \mathbf{b}$, con $\mathbf{a} \in U$ e $\mathbf{b} \in U^{\perp}$ univocamente determinati. E sappiamo anche che $\mathbf{a} = p_U(\mathbf{v})$ (si veda (*) nella dimostrazione della proprieta' 3) della proposizione precedente). Poiche' si ha anche $V = U^{\perp} \oplus (U^{\perp})^{\perp}$, allora in corrispondenza di tale decomposizione potremo scrivere in modo unico $\mathbf{v} = \mathbf{a}' + \mathbf{b}'$ con $\mathbf{a}' = p_{U^{\perp}}(\mathbf{v}) \in U^{\perp}$ e $\mathbf{b}' \in (U^{\perp})^{\perp} = U$. Per l'unicita' deve essere $\mathbf{b} = \mathbf{a}' = p_{U^{\perp}}(\mathbf{v})$. Cio' prova la seconda proprieta'.

Se $\mathbf{v} = p_U(\mathbf{v})$ allora e' ovvio che $\mathbf{v} \in U$; viceversa, se $\mathbf{v} \in U$ allora $\mathbf{v} - p_U(\mathbf{v}) = p_{U^{\perp}}(\mathbf{v}) \in U \cap U^{\perp}$, quindi $\mathbf{v} - p_U(\mathbf{v}) = \mathbf{0}$, cioe' $\mathbf{v} = p_U(\mathbf{v})$.

Infine se $p_U(\mathbf{v}) = \mathbf{0}$ allora $\mathbf{v} = p_{U^{\perp}}(\mathbf{v}) \in U^{\perp}$. Viceversa se $\mathbf{v} \in U^{\perp}$, per la proprieta' 3) prima dimostrata allora $\mathbf{v} = p_{U^{\perp}}(\mathbf{v})$, d'altra parte $\mathbf{v} = p_U(\mathbf{v}) + p_{U^{\perp}}(\mathbf{v})$ e quindi $p_U(\mathbf{v}) = \mathbf{0}$. ■

Sia $(V, \langle\ ,\ \rangle)$ uno spazio euclideo, ed U un sottospazio di V. Abbiamo appena visto che V e' la somma diretta di U e di $U^{\perp}$. Questa proprieta' si esprime anche dicendo che V *e' la somma diretta ortogonale di* U *e* $U^{\perp}$, e si scrive

$$V = U \perp U^{\perp}.$$

Osservazione 4. *Piu' in generale se* $U_1, U_2, \dots, U_r$ *sono sottospazi tali che* $V = U_1 \oplus U_2 \oplus \cdots \oplus U_r$ *ed inoltre, per ogni* $i \neq j$, *ogni vettore di* U_i *e' ortogonale ad ogni vettore di* U_j *(cioe'* $U_i \subseteq {U_j}^{\perp}$*), allora si dice che* V *e' la somma diretta ortogonale di* $U_1, U_2, \dots, U_r$ *e si scrive*

$$V = U_1 \perp U_2 \perp \cdots \perp U_r.$$

In tal caso, poiche' la somma e' diretta, allora ogni vettore $\mathbf{v} \in V$ *si scrive in unico modo come somma* $\mathbf{v} = \mathbf{u}_1 + \mathbf{u}_2 + \cdots + \mathbf{u}_r$, *con* $\mathbf{u}_i \in U_i$. *D'altra parte la somma e'* *ortogonale e quindi i vettori componenti* $\mathbf{u}_i$ *si possono interpretare geometricamente nel senso che* $\mathbf{u}_i = p_{U_i}(\mathbf{v})$*, cioe'*

$$(*) \qquad \mathbf{v} = p_{U_1}(\mathbf{v}) + p_{U_2}(\mathbf{v}) + \cdots + p_{U_r}(\mathbf{v}).$$

Possiamo provare questa decomposizione nel seguente modo. Per ogni $i = 1, \dots, r$ *sia* $\mathcal{B}_i = \{\mathbf{b}_1^{(i)}, \dots, \mathbf{b}_{n_i}^{(i)}\}$ *una base ortonormale di* U_i*. Poiche' la somma diretta e' ortogonale allora l'unione*

$$\mathcal{B} := \bigcup_{i=1}^{r} \mathcal{B}_i$$

e' una base ortonormale di V*. Per quanto detto nell'Esempio 31 sappiamo che la coordinata* $x_j^{(i)}$ *di* $\mathbf{v}$ *rispetto al vettore* $\mathbf{b}_j^{(i)}$ *coincide con il coefficiente di Fourier di* $\mathbf{v}$ *rispetto a* $\mathbf{b}_j^{(i)}$*, quindi*

$$\mathbf{v} = \sum_{i=1}^{r} p_{\mathbf{b}_1^{(i)}}(\mathbf{v}) + p_{\mathbf{b}_2^{(i)}}(\mathbf{v}) + \cdots + p_{\mathbf{b}_{n_i}^{(i)}}(\mathbf{v}).$$

Cio' prova la decomposizione $(^*)$ *in quanto per definizione*

$$p_{U_i}(\mathbf{v}) = p_{\mathbf{b}_1^{(i)}}(\mathbf{v}) + p_{\mathbf{b}_2^{(i)}}(\mathbf{v}) + \cdots + p_{\mathbf{b}_{n_i}^{(i)}}(\mathbf{v}).$$

Oltre alla decomposizione $V = U \perp U^{\perp}$*, un altro esempio di decomposizione ortogonale si puo' ottenere a partire da una base ortogonale* $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$*: infatti in tal caso si ha*

$$V = Span(\mathbf{b}_1) \perp Span(\mathbf{b}_2) \perp \cdots \perp Span(\mathbf{b}_n).$$

Esempio 34. *Nello spazio euclideo* $(\mathbf{R}^3, prodotto\,punto)$ *si consideri il sottospazio* U *definito dall'equazione cartesiana* $x + y + 4z = 0$*. Calcolare una base ortogonale per* U*, una per* $U^{\perp}$*, e decomporre il vettore* $(1, 1, 2)$ *nella somma di un vettore* $\mathbf{a}$ *parallelo ad* U *e di un vettore* $\mathbf{b}$ *ortogonale ad* U*.*

Svolgimento. Risolvendo l'equazione cartesiana di U *vediamo che una base di* U *e' formata dai vettori* $(-1, 1, 0), (-4, 0, 1)$*. Si osservi che tale base non e' ortogonale. Per trovare una base ortogonale di* U *applichiamo il procedimento di Gram-Schmidt. Quindi poniamo* $\mathbf{c}_1 = (-1, 1, 0)$ *e calcoliamo* $\mathbf{c}_2 = (-4, 0, 1) - p_{(-1,1,0)}(-4, 0, 1) = (-2, -2, 1)$*. I vettori* $(-1, 1, 0), (-2, -2, 1)$ *formano una base ortogonale di* U*.*

Poiche' $\dim U^{\perp} = 1$*, allora una base ortogonale di* $U^{\perp}$ *e' una qualunque base per* $U^{\perp}$*, cioe' un qualunque vettore non nullo di* $U^{\perp}$*. Per trovare un tale vettore ricordiamo che* $(x, y, z) \in U^{\perp}$ *se e solo se*

$$\langle (x, y, z), (-1, 1, 0) \rangle = \langle (x, y, z), (-4, 0, 1) \rangle = 0,$$

cioe' se e solo se

$$\begin{cases} -x+y=0 \\ -4x+z=0. \end{cases}$$

Risolvendo tale sistema vediamo che $U^\perp = Span(\{(1,1,4)\})$. *Dunque* $(1,1,4)$ *e' una base ortogonale di* $U^\perp$. *Per calcolare tale base avremmo anche potuto ragionare diversamente. E cioe' osservando che possiamo scrivere l'equazione cartesiana di* U *nel seguente modo:*

$$x+y+4z = \langle\,(1,1,4),(x,y,z)\rangle = 0.$$

Cio' significa proprio che il vettore $(1,1,4)$ *e' un vettore di* $U^\perp$, *e percio' ne e' una base ortogonale.*

Infine andiamo a calcolare $\mathbf{a}$ *e* $\mathbf{b}$. *Sappiamo che* $\mathbf{a} = p_U(1,1,2)$. *Per calcolare* $p_U(1,1,2)$ *non possiamo utilizzare la base* $(-1,1,0),(-4,0,1)$ *di* U *perche' questa non e' una base ortogonale. Invece possiamo utilizzare la base* $(-1,1,0),(-2,-2,1)$ *e dunque*

$$\mathbf{a} = p_U(1,1,2) = p_{(-1,1,0)}(1,1,2) + p_{(-2,-2,1)}(1,1,2)$$

$$= \frac{\langle\,(1,1,2),(-1,1,0)\rangle}{\langle\,(-1,1,0),(-1,1,0)\rangle}(-1,1,0) + \frac{\langle\,(1,1,2),(-2,-2,1)\rangle}{\langle\,(-2,-2,1),(-2,-2,1)\rangle}(-2,-2,1) = \frac{2}{9}(2,2,-1).$$

Ne consegue che

$$\mathbf{b} = (1,1,2) - \mathbf{a} = \frac{5}{9}(1,1,4).$$

Anche in questo caso avremmo potuto ragionare in un altro modo. Poiche' sappiamo che $\mathbf{b} = p_{U^\perp}(1,1,2)$ *allora*

$$\mathbf{b} = p_{U^\perp}(1,1,2) = \frac{\langle\,(1,1,2),(1,1,4)\rangle}{\langle\,(1,1,4),(1,1,4)\rangle}(1,1,4) = \frac{5}{9}(1,1,4). \quad \blacksquare$$

Esempio 35. *Nello spazio euclideo* $(\mathbf{R}^4, prodotto\,punto)$ *si consideri il sottospazio* U *definito dalla rappresentazione cartesiana*

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

Calcolare una base ortogonale per U *ed una per* $U^\perp$. *Infine calcolare la proiezione ortogonale* $p_U(x,y,z,t)$ *su* U *del generico vettore* (x,y,z,t) *di* $\mathbf{R}^4$.

Svolgimento. Risolvendo il sistema lineare che rappresenta U *otteniamo la seguente base* $(0,-1,1,0),(0,-1,0,1)$ *di* U. *Non e' una base ortogonale, ma possiamo ortogonalizzarla ponendo* $\mathbf{c}_1 = (0,-1,1,0)$, *e*

$$\mathbf{c}_2 = (0,-1,0,1) - p_{(0,-1,1,0)}(0,-1,0,1) = (0,-\frac{1}{2},-\frac{1}{2},1).$$

Quindi una base ortogonale per U *e' formata dai vettori* $(0,-1,1,0),(0,1,1,-2)$.

Per calcolare una base ortogonale di $U^\perp$ innanzitutto andiamo a calcolarne una base, ricordando che poiche' $U = Span((0,-1,1,0),(0,-1,0,1))$ allora $U^\perp$ e' costituito dai quei vettori (x,y,z,t) tali che

$$\begin{cases} \langle\,(x,y,z,t),(0,-1,1,0)\rangle = 0 \\ \langle\,(x,y,z,t),(0,-1,0,1)\rangle = 0 \end{cases}$$

cioe' tali che

$$\begin{cases} -y+z=0 \\ -y+t=0. \end{cases}$$

Questa e' una rappresentazione cartesiana di $U^\perp$. Risolvendo il sistema deduciamo che una base per $U^\perp$ e' formata dai vettori $(1,0,0,0),(0,1,1,1)$. I due vettori sono ortogonali fra loro, dunque tale base e' anche una base ortogonale di $U^\perp$.

Infine calcoliamo

$$p_U(x,y,z,t) = p_{(0,-1,1,0)}(x,y,z,t) + p_{(0,1,1,-2)}(x,y,z,t)$$

$$= \frac{-y+z}{2}(0,-1,1,0) + \frac{y+z-2t}{6}(0,1,1,-2)$$
$$= (0, \frac{2y-z-t}{3}, \frac{-y+2z-t}{3}, \frac{-y-z+2t}{3}).$$

Si osservi che le formule precedenti ci dicono che $p_U(x,y,z,t)$ e' una funzione lineare di (x,y,z,t). Questo e' un fatto di carattere generale che studieremo nel prossimo paragrafo. La matrice rappresentativa di p_U rispetto alla base canonica e'

$$M = M_{\mathcal{E}}^{\mathcal{E}}(p_U) = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} \\ 0 & -\frac{1}{3} & \frac{2}{3} & -\frac{1}{3} \\ 0 & -\frac{1}{3} & -\frac{1}{3} & \frac{2}{3} \end{bmatrix}.$$

Infine osserviamo anche che M e' simmetrica, e che $M^2 = M$. ■

Esempio 36. *Siano $\mathbf{u}$, $\mathbf{v}$ vettori in uno spazio euclideo V. Provare che*

1) $\|\mathbf{v}-\mathbf{u}\|^2 = \|\mathbf{v}\|^2 + \|\mathbf{u}\|^2 - 2\|\mathbf{v}\|\|\mathbf{u}\|\cos\widehat{\mathbf{uv}}$ (Regola del coseno, o anche Teorema di Pitagora generalizzato);

2) $\|\mathbf{v}+\mathbf{u}\|^2 = \|\mathbf{v}\|^2 + \|\mathbf{u}\|^2 + 2\|\mathbf{v}\|\|\mathbf{u}\|\cos\widehat{\mathbf{uv}}$;

3) se $\mathbf{v} \perp \mathbf{u}$ allora $\|\mathbf{v}-\mathbf{u}\|^2 = \|\mathbf{v}\|^2 + \|\mathbf{u}\|^2$ (Teorema di Pitagora);

4) $\|\mathbf{v}-\mathbf{u}\|^2 + \|\mathbf{v}+\mathbf{u}\|^2 = 2\left(\|\mathbf{v}\|^2 + \|\mathbf{u}\|^2\right)$ (Proprieta' del Parallelogramma).

Svolgimento. La prima proprieta' segue osservando che

$$\|\mathbf{v}-\mathbf{u}\|^2 = \langle\,\mathbf{v}-\mathbf{u},\mathbf{v}-\mathbf{u}\rangle = \langle\,\mathbf{v},\mathbf{v}\rangle - 2\langle\,\mathbf{v},\mathbf{u}\rangle + \langle\,\mathbf{u},\mathbf{u}\rangle$$

$$= \langle \mathbf{v}, \mathbf{v} \rangle - 2\|\mathbf{v}\|\|\mathbf{u}\| \cos \widehat{\mathbf{u}\mathbf{v}} + \langle \mathbf{u}, \mathbf{u} \rangle = \|\mathbf{v}\|^2 - 2\|\mathbf{v}\|\|\mathbf{u}\| \cos \widehat{\mathbf{u}\mathbf{v}} + \|\mathbf{u}\|^2.$$

La seconda proprieta' si prova allo stesso modo. Il Teorema di Pitagora segue da 1), mentre la Proprieta' del Parallelogramma segue sommando la 1) con la 2). ■

Siano $\mathbf{u}$, $\mathbf{v}$ vettori in uno spazio euclideo V. Si definisce *distanza tra* $\mathbf{u}$ *e* $\mathbf{v}$ il numero

$$d(\mathbf{u}, \mathbf{v}) := \|\mathbf{v} - \mathbf{u}\|.$$

Si osservi che $d(\mathbf{u}, \mathbf{v}) = d(\mathbf{v}, \mathbf{u})$, e che se si introducono coordinate ortonormali in V allora si ha

$$d(\mathbf{u}, \mathbf{v}) = \sqrt{\sum_{i=1}^{n} (y_i - x_i)^2},$$

dove n e' la dimensione di V ed $\mathbf{x}$ ed $\mathbf{y}$ denotano le coordinate di $\mathbf{u}$ e $\mathbf{v}$.

Ora fissiamo un vettore $\mathbf{v}$ ed un sottospazio U di V, e consideriamo le distanze $d(\mathbf{u}, \mathbf{v})$ al variare di $\mathbf{u}$ in U, cioe' andiamo a considerare l'insieme numerico

$$\{d(\mathbf{u}, \mathbf{v}) \,:\, \mathbf{u} \in U\}.$$

Si osservi che poiche' $p_U(\mathbf{v}) \in U$ allora $d(p_U(\mathbf{v}), \mathbf{v}) \in \{d(\mathbf{u}, \mathbf{v}) \,:\, \mathbf{u} \in U\}$. Andiamo a provare che

$$(*) \qquad d(p_U(\mathbf{v}), \mathbf{v}) = \min \{d(\mathbf{u}, \mathbf{v}) \,:\, \mathbf{u} \in U\}.$$

Cioe' al variare di $\mathbf{u} \in U$ il valore della distanza $d(\mathbf{u}, \mathbf{v})$ diventa minimo in corrispondenza del vettore $\mathbf{u} = p_U(\mathbf{v})$. Infatti innanzitutto osserviamo che

$$d(\mathbf{u}, \mathbf{v})^2 = \|\mathbf{v} - \mathbf{u}\|^2 = \|(\mathbf{v} - p_U(\mathbf{v})) + (p_U(\mathbf{v}) - \mathbf{u})\|^2.$$

Ora noi sappiamo che $\mathbf{v} - p_U(\mathbf{v})$ e' un vettore ortogonale ad U, quindi per il Teorema di Pitagora abbiamo ancora

$$d(\mathbf{u}, \mathbf{v})^2 = \|(\mathbf{v} - p_U(\mathbf{v})) + (p_U(\mathbf{v}) - \mathbf{u})\|^2 = \|\mathbf{v} - p_U(\mathbf{v})\|^2 + \|p_U(\mathbf{v}) - \mathbf{u}\|^2$$

$$= d(\mathbf{v}, p_U(\mathbf{v}))^2 + d(p_U(\mathbf{v}), \mathbf{u})^2 \geq d(\mathbf{v}, p_U(\mathbf{v}))^2.$$

Cio' prova la (*). Si osservi che il calcolo precedente prova anche che *se* $\mathbf{u}$ *e' un vettore di* U *tale che* $d(\mathbf{u}, \mathbf{v}) = d(\mathbf{v}, p_U(\mathbf{v}))$ *allora* $\mathbf{u} = p_U(\mathbf{v})$. Infine osserviamo che

$$d(\mathbf{v}, p_U(\mathbf{v})) = \|\mathbf{v} - p_U(\mathbf{v})\| = \|p_{U^\perp}(\mathbf{v})\|.$$

Il numero $d(\mathbf{v}, p_U(\mathbf{v}))$ si chiama anche *la distanza del vettore* $\mathbf{v}$ *dal sottospazio* U, e si denota

$$d(\mathbf{v}, U).$$

Riepilogando, *la distanza di un vettore* $\mathbf{v}$ *da un sottospazio* U *e' la minima distanza del vettore* $\mathbf{v}$ *da* $\mathbf{u}$*, al variare di* $\mathbf{u}$ *in* U*, e tale distanza minima coincide con la lunghezza della proiezione ortogonale di* $\mathbf{v}$ *sul complemento ortogonale di* U.

Esempio 37. *Con le notazioni dell'Esempio 35 calcolare la distanza del vettore* $(1,2,3,1)$ *dal sottospazio* U.

Svolgimento. Sappiamo che i vettori $(1,0,0,0),(0,1,1,1)$ *formano una base ortogonale di* $U^\perp$. *Quindi*

$$d((1,2,3,1),U) = \|p_{U^\perp}(1,2,3,1)\| = \|p_{(1,0,0,0)}(1,2,3,1) + p_{(0,1,1,1)}(1,2,3,1)\|$$

$$= \|(1,2,2,2)\| = \sqrt{13}. \quad \blacksquare$$

10. La proiezione ortogonale come operatore lineare.

Sia V uno spazio euclideo, ed U un sottospazio di V. Abbiamo definito, per ogni $\mathbf{v}\in V$, la proiezione ortogonale $p_U(\mathbf{v})$ di $\mathbf{v}$ su U. Abbiamo quindi un'applicazione

$$p_U : \mathbf{v}\in V \to p_U(\mathbf{v}) \in V.$$

Andiamo a provare che p_U e' un operatore lineare. A tale proposito fissiamo una base ortogonale $\mathcal{B} = \{\mathbf{u}_1,\dots,\mathbf{u}_h\}$ di U, e siano $\mathbf{v},\mathbf{w}$ vettori di V. Abbiamo:

$$p_U(\mathbf{v}+\mathbf{w}) = \sum_{j=1}^h p_{\mathbf{u}_j}(\mathbf{v}+\mathbf{w}) = \sum_{j=1}^h \frac{\langle \mathbf{v}+\mathbf{w},\mathbf{u}_j\rangle}{\langle \mathbf{u}_j,\mathbf{u}_j\rangle}\mathbf{u}_j = \sum_{j=1}^h \frac{\langle \mathbf{v},\mathbf{u}_j\rangle + \langle \mathbf{w},\mathbf{u}_j\rangle}{\langle \mathbf{u}_j,\mathbf{u}_j\rangle}\mathbf{u}_j$$

$$= \sum_{j=1}^h \frac{\langle \mathbf{v},\mathbf{u}_j\rangle}{\langle \mathbf{u}_j,\mathbf{u}_j\rangle}\mathbf{u}_j + \sum_{j=1}^h \frac{\langle \mathbf{w},\mathbf{u}_j\rangle}{\langle \mathbf{u}_j,\mathbf{u}_j\rangle}\mathbf{u}_j = p_U(\mathbf{v}) + p_U(\mathbf{w}).$$

Cio' prova che p_U rispetta l'addizione. Poi se c e' uno scalare abbiamo

$$p_U(c\mathbf{v}) = \sum_{j=1}^h p_{\mathbf{u}_j}(c\mathbf{v}) = \sum_{j=1}^h \frac{\langle c\mathbf{v},\mathbf{u}_j\rangle}{\langle \mathbf{u}_j,\mathbf{u}_j\rangle}\mathbf{u}_j = \sum_{j=1}^h \frac{c\langle \mathbf{v},\mathbf{u}_j\rangle}{\langle \mathbf{u}_j,\mathbf{u}_j\rangle}\mathbf{u}_j = c\sum_{j=1}^h \frac{\langle \mathbf{v},\mathbf{u}_j\rangle}{\langle \mathbf{u}_j,\mathbf{u}_j\rangle}\mathbf{u}_j = c p_U(\mathbf{v}).$$

Quindi p_U rispetta anche la moltiplicazione esterna, e dunque e' un operatore lineare.

Ora ricordiamo che $V = U \perp U^\perp$, e che se $\mathbf{v}\in U$ allora $p_U(\mathbf{v}) = \mathbf{v}$, e che se $\mathbf{v}\in U^\perp$ allora $p_U(\mathbf{v}) = \mathbf{0}$. Quindi se $\mathcal{U}$ e' una base di U ed $\mathcal{U}'$ e' una base di $U^\perp$, allora $\mathcal{B} = \mathcal{U}\cup\mathcal{U}'$ e' una base di autovettori per p_U, e la matrice rappresentativa $M_{\mathcal{B}}^{\mathcal{B}}(p_U)$ e' la matrice diagonale la cui diagonale principale e' formata da 1 (tanti quanti la dimensione di U) e da 0 (tanti quanti la dimensione di $U^\perp$). In particolare p_U e' un operatore diagonalizzabile, U e' l'autospazio di p_U relativo all'autovalore $\lambda = 1$, mentre $U^\perp$ e' l'autospazio di p_U relativo all'autovalore $\lambda = 0$, e il polinomio caratteristico di p_U e' $p_{p_U}(t) = (t-1)^{\dim U} t^{\dim U^\perp}$. Possiamo riassumere quanto detto nella seguente

Proposizione 4. *Sia* U *un sottospazio di dimensione* h *di uno spazio euclideo* V *di dimensione* n. *Allora l'applicazione proiezione ortogonale su* U $p_U : \mathbf{v}\in V \to$

$p_U(\mathbf{v}) \in V$ *e' un operatore lineare diagonalizzabile, con spettro dato da* 1 *e* 0*. Lo spazio* U *coincide con l'autospazio* V_1*, mentre* $U^\perp$ *coincide con l'autospazio* V_0*. Una base di autovettori per* p_U *si ottiene riunendo una base di* U *con una base di* $U^\perp$*. L'espressione esplicita di* p_U *rispetto ad una base siffatta e'*

$$p_U(x_1, x_2, \dots, x_h, x_{h+1}, \dots, x_n) = (x_1, x_2, \dots, x_h, 0, \dots, 0).$$

Per la precisione, se $U = V$ allora $p_U = id_V$ e c'e' solo l'autovalore 1, mentre se $U = \{\mathbf{0}\}$, allora $p_U \equiv \mathbf{0}$ e c'e' solo l'autovalore 0.

Un'altra interessante proprieta' delle proiezioni ortogonali e' la seguente, cioe' che per ogni $\mathbf{v}$, $\mathbf{w}$ in V si ha

$$(*) \qquad \langle\, p_U(\mathbf{v}), \mathbf{w} \rangle = \langle\, \mathbf{v}, p_U(\mathbf{w}) \rangle.$$

Questa proprieta' si esprime dicendo che *l'operatore di proiezione e' simmetrico (o autoaggiunto)*. Per provare questa proprieta' possiamo ragionare nel seguente modo. Sia $\mathcal{U} = \{\mathbf{u}_1, \dots, \mathbf{u}_h\}$ una base ortonormale di U, e $\mathcal{U}' = \{\mathbf{u}_{h+1}, \dots, \mathbf{u}_n\}$ una base ortonormale di $U^\perp$. Allora $\mathcal{B} = \mathcal{U} \cup \mathcal{U}'$ e' una base ortonormale di V. Utilizzando le coordinate rispetto a tale base possiamo calcolare:

$$\langle\, p_U(\mathbf{v}), \mathbf{w} \rangle = [p_U(\mathbf{v})]_{\mathcal{B}}^T \mathbf{y} = (x_1, x_2, \dots, x_h, 0, \dots, 0) \cdot \mathbf{y}$$

$$= x_1 y_1 + \dots + x_h y_h = (x_1, x_2, \dots, x_h, x_{h+1}, \dots, x_n)(y_1, y_2, \dots, y_h, 0, \dots, 0)^T$$

$$= \mathbf{x}^T [p_U(\mathbf{w})]_{\mathcal{B}} = \langle\, \mathbf{v}, p_U(\mathbf{w}) \rangle.$$

Dalla proprieta' appena dimostrata segue che *la matrice rappresentativa* M *di* p_U *rispetto ad una qualsiasi base ortonormale e' simmetrica.* Infatti utilizzando le coordinate rispetto ad una base ortonormale la (*) ci dice che

$$(M\mathbf{x})^T \mathbf{y} = \mathbf{x}^T (M\mathbf{y})$$

cioe'

$$\mathbf{x}^T M^T \mathbf{y} = \mathbf{x}^T M \mathbf{y}$$

per ogni $\mathbf{x}, \mathbf{y} \in \mathbf{R}^n$, da cui $M = M^T$. Inoltre poiche' evidentemente $p_U \circ p_U = p_U$, si ha anche $M^2 = M$. Pertanto un operatore di proiezione ortogonale e' rappresentato rispetto ad una base ortonormale da una matrice M soddisfacente le condizioni

$$M^T = M, \quad M^2 = M.$$

Vedremo nel successivo paragrafo 12 che tali condizioni non solo sono necessarie, ma sono anche sufficienti affinche' un operatore $f : V \to V$ sia un operatore di proiezione ortogonale (ed in tal caso f e' la proiezione ortogonale sull'autospazio V_1, cioe' sul complemento ortogonale del suo nucleo). In particolare un operatore del tipo

$$\mathbf{x} \in \mathbf{R}^n \to M\mathbf{x} \in \mathbf{R}^n$$

rappresenta una proiezione ortogonale in $(\mathbf{R}^n, \textit{prodotto punto})$ se e solo se $M = M^T$ ed $M = M^2$.

Esempio 38. *Nello spazio euclideo* $(\mathbf{R}^n, \textit{prodotto punto})$ *calcolare la matrice rappresentativa rispetto alla base canonica* $\mathcal{E}$ *della proiezione ortogonale* $p_{\mathbf{u}}$ *sul vettore* $\mathbf{u}$.

Svolgimento. Poniamo $\mathbf{u} = (u_1, \dots, u_n)$. *Poiche' la colonna di posto* j *di* $M_{\mathcal{E}}^{\mathcal{E}}(p_{\mathbf{u}})$ *e'*

$$p_{\mathbf{u}}(\mathbf{e}_j) = \frac{\langle \mathbf{u}, \mathbf{e}_j \rangle}{\langle \mathbf{u}, \mathbf{u} \rangle} \mathbf{u} = \frac{u_j}{\mathbf{u}^T \mathbf{u}} \mathbf{u}$$

allora

$$M_{\mathcal{E}}^{\mathcal{E}}(p_{\mathbf{u}}) = \frac{\mathbf{u}\mathbf{u}^T}{\mathbf{u}^T \mathbf{u}}. \quad \blacksquare$$

11. Le matrici ortogonali.

Molto particolari ed interessanti sono le matrici del cambiamento delle coordinate tra basi ortonormali in uno spazio euclideo: tali matrici si dicono *ortogonali*. Queste matrici si possono caratterizzare in vari modi, infatti si ha la seguente

Proposizione 5. *Sia* P *una matrice reale* $n \times n$. *Sono equivalenti le seguenti proprieta':*

(i) P *e' una matrice ortogonale;*

(ii) $PP^T = P^T P = I$;

(iii) P *e' invertibile e* $P^{-1} = P^T$;

(iv) le righe di P *formano una base ortonormale per* $(\mathbf{R}^n, \textit{prodotto punto})$;

(v) P^T *e' ortogonale;*

(vi) le colonne di P *formano una base ortonormale per* $(\mathbf{R}^n, \textit{prodotto punto})$.

Dimostrazione. Sia P una matrice ortogonale. Per definizione esiste uno spazio euclideo V e basi ortonormali $\mathcal{B}$ e $\mathcal{B}'$ di V tali che $P = M_{\mathcal{B}}^{\mathcal{B}'}(id_V)$. Siano G e G' le matrici di Gram del prodotto scalare di V rispetto a $\mathcal{B}$ e $\mathcal{B}'$. Sappiamo che $G' = P^T G P$. D'altra parte essendo le basi ortonormali deve essere $G = G' = I$. Quindi $I = P^T P$. Cio' prova che (i) implica (ii), che evidentemente equivale a (iii).

Supponiamo ora che $PP^T = I$, e denotiamo con P_i la riga di posto i di P e con $(P^T)^j$ la colonna di posto j di P^T, che coincide con P_j. Poiche' $PP^T = I$ allora per definizione di prodotto righe per colonna il prodotto punto di P_i con $(P^T)^j = P_j$ e' uguale a 1 se $i = j$, ed e' uguale a 0 se $i \neq j$. Questo significa che le righe di P formano una base ortonormale per lo spazio euclideo $(\mathbf{R}^n, \textit{prodotto punto})$. Cio' prova che (iii) implica (iv).

Supponiamo che le righe di P formino una base ortonormale $\mathcal{B}$ per lo spazio euclideo $(\mathbf{R}^n, \textit{prodotto punto})$. Allora la matrice P^T, che e' uguale a $M_{\mathcal{E}}^{\mathcal{B}}(id_{\mathbf{R}^n})$, e' una matrice ortogonale per definizione. Cio' prova che (iv) implica (v).

Se P^T e' una matrice ortogonale, allora abbiamo appena visto che vale la proprieta' (ii), quindi $P^T(P^T)^T = I$, cioe' $P^TP = I$, e possiamo provare (vi) ripetendo lo stesso argomento usato per provare che (iii) implica (iv). Cio' prova che (v) implica (vi).

Infine, per provare che (vi) implica (i), osserviamo che se le colonne di P formano una base ortonormale $\mathcal{B}$ per lo spazio euclideo $(\mathbf{R}^n, \textit{prodotto punto})$, allora la matrice P, che e' uguale a $M_{\mathcal{E}}^{\mathcal{B}}(id_{\mathbf{R}^n})$, e' una matrice ortogonale per definizione. Cio' prova che (vi) implica (i). ∎

Denotiamo con $O(n)$ l'insieme delle matrici reali $n \times n$ ortogonali. Osserviamo che $I \in O(n)$, che se $P, Q \in O(n)$ allora $PQ \in O(n)$ e che se $P \in O(n)$ allora $P^{-1} \in O(n)$. Per via di queste proprieta' si dice che $O(n)$ forma un *gruppo* rispetto alla moltiplicazione, detto il *gruppo ortogonale*. Andiamo a descrivere l'insieme $O(n)$ nel caso $n = 1$ ed $n = 2$.

Se $n = 1$ evidentemente si ha $O(1) = \{-1, 1\}$.

Esaminiamo il caso $n = 2$. Sia P una matrice ortogonale 2×2. Sappiamo allora che le colonne di P formano una base ortonormale per $(\mathbf{R}^2, \textit{prodotto punto})$. Sappiamo anche che, fissata una base ortonormale $\{\overrightarrow{OE_1}, \overrightarrow{OE_2}\}$ nello spazio dei vettori geometrici del piano applicati nel punto O, possiamo identificare lo spazio euclideo dei vettori geometrici del piano munito del prodotto scalare, con $(\mathbf{R}^2, \textit{prodotto punto})$. In tale identificazione le due colonne di P corrispondono a due vettori geometrici $\overrightarrow{OA}$ ed $\overrightarrow{OB}$ di lunghezza 1. Quindi gli estremi A e B si trovano sulla circonferenza di raggio 1 centrata in O. In particolare, se α e' l'angolo che il vettore $\overrightarrow{OA}$ forma con il vettore $\overrightarrow{OE_1}$, le coordinate del vettore $\overrightarrow{OA}$ (che formano la prima colonna di P) sono $(\cos\alpha, \sin\alpha)^T$. Poiche' $\overrightarrow{OB}$ e' ortogonale ad $\overrightarrow{OA}$, allora le sue coordinate (che formano la seconda colonna di P) sono $(-\sin\alpha, \cos\alpha)^T$ oppure $(\sin\alpha, -\cos\alpha)^T$. In altre parole la matrice P o e' del tipo

$$R_\alpha = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}$$

oppure e' del tipo

$$S_\alpha = \begin{bmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{bmatrix}.$$

Nel primo caso riconosciamo la matrice R_α come la matrice che rappresenta una rotazione antioraria di angolo α, nel secondo la matrice S_α come la matrice che rappresenta la simmetria ortogonale rispetto alla retta che forma l'angolo $\frac{1}{2}\alpha$ con $\overrightarrow{OE_1}$. In particolare possiamo dire che

$$O(2) = \left\{ \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix} : \alpha \in \mathbf{R} \right\} \cup \left\{ \begin{bmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{bmatrix} : \alpha \in \mathbf{R} \right\}.$$

Si osservi che l'unione e' disgiunta in quanto $\det R_\alpha = 1$, mentre $\det S_\alpha = -1$.

Esiste una descrizione analoga per $O(3)$, nella quale la generica matrice ortogonale appare in funzione di tre angoli, detti gli *angoli di Eulero*. Piu' in generale la generica matrice ortogonale $n \times n$ si puo' rappresentare in funzione di $\binom{n}{2}$ angoli.

Esempio 39. *Completare la seguente matrice a matrice ortogonale:*

$$\begin{bmatrix} -\frac{1}{\sqrt{3}} & * & * \\ \frac{1}{\sqrt{3}} & * & * \\ \frac{1}{\sqrt{3}} & * & * \end{bmatrix}.$$

Svolgimento. Possiamo procedere cosi'. Completiamo il vettore $(-1,1,1)$ *a base di* $\mathbf{R}^3$ *aggiungendo i vettori canonici* $(1,0,0)$ *e* $(0,1,0)$*, e poi applichiamo il procedimento di ortogonalizzazione di Gram-Schmidt alla sequenza* $(-1,1,1)$*,* $(1,0,0)$*,* $(0,1,0)$*. Quindi otteniamo* $\mathbf{c}_1 = (-1,1,1)$*,* $\mathbf{c}_2 = (\frac{2}{3}, \frac{1}{3}, \frac{1}{3})$*,* $\mathbf{c}_3 = (0, \frac{1}{2}, -\frac{1}{2})$*. Normalizzando otteniamo la seguente base ortonormale*

$$\frac{1}{\sqrt{3}}(-1,1,1),\ \frac{1}{\sqrt{6}}(2,1,1),\ \frac{1}{\sqrt{2}}(0,1,-1).$$

In conclusione una matrice che risponde ai requisiti richiesti e'

$$\begin{bmatrix} -\frac{1}{\sqrt{3}} & \frac{2}{\sqrt{6}} & 0 \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} \end{bmatrix}. \quad \blacksquare$$

12. Operatori lineari in uno spazio euclideo.

In questo paragrafo andremo a studiare gli endomorfismi di uno spazio vettoriale in presenza di un prodotto scalare. Cominciamo con l'osservare la seguente

Proposizione 6. *Sia* $(V, \langle\ ,\ \rangle)$ *uno spazio euclideo, ed* $f : V \to V$ *un operatore lineare. Allora esiste un unico operatore* $g : V \to V$ *tale che per ogni* $\mathbf{u}$*,* $\mathbf{v}$ *in* V *si ha*

$$\langle\, f(\mathbf{u}), \mathbf{v}\,\rangle = \langle\, \mathbf{u}, g(\mathbf{v})\,\rangle.$$

Dimostrazione. Fissiamo una base ortonormale $\mathcal{B}$ di V, e poniamo $A := M_{\mathcal{B}}^{\mathcal{B}}(f)$. Poi consideriamo l'operatore $g : V \to V$ definito ponendo

$$[g(\mathbf{u})]_{\mathcal{B}} := A^T [\mathbf{u}]_{\mathcal{B}}.$$

In altre parole g e' l'operatore che rispetto alla base ortonormale $\mathcal{B}$ e' rappresentato dalla matrice A^T. Andiamo a provare che g e' un operatore che verifica la proprieta' richiesta. Innanzitutto osserviamo che essendo $\mathcal{B}$ ortonormale si ha $\langle\, f(\mathbf{u}), \mathbf{v}\,\rangle = [f(\mathbf{u})]_{\mathcal{B}}^T [\mathbf{v}]_{\mathcal{B}}$. Quindi, denotando come di consueto con $\mathbf{x}$ le coordinate di $\mathbf{u}$ e con $\mathbf{y}$ le coordinate di $\mathbf{v}$, abbiamo

$$\langle\, f(\mathbf{u}), \mathbf{v}\,\rangle = [f(\mathbf{u})]_{\mathcal{B}}^T [\mathbf{v}]_{\mathcal{B}} = (A\mathbf{x})^T \mathbf{y} = (\mathbf{x}^T A^T)\mathbf{y} = \mathbf{x}^T (A^T \mathbf{y}) = \langle\, \mathbf{u}, g(\mathbf{v})\,\rangle.$$

Cio' prova che l'operatore g esiste.

Rimane da provare l'unicita' di g. A tale proposito sia $h : V \to V$ un operatore tale che per ogni $\mathbf{u}$ e $\mathbf{v}$ in V si ha $\langle\, f(\mathbf{u}), \mathbf{v}\,\rangle = \langle\, \mathbf{u}, h(\mathbf{v})\,\rangle$. Allora per ogni $\mathbf{u}$ e $\mathbf{v}$ si ha anche

$$\langle\, \mathbf{u}, g(\mathbf{v})\,\rangle = \langle\, \mathbf{u}, h(\mathbf{v})\,\rangle$$

cioe'

$$\langle\, \mathbf{u}, g(\mathbf{v}) - h(\mathbf{v})\,\rangle = 0$$

per ogni $\mathbf{u}$ e $\mathbf{v}$. In particolare per ogni $\mathbf{v} \in V$ si ha

$$\langle\, g(\mathbf{v}) - h(\mathbf{v}), g(\mathbf{v}) - h(\mathbf{v})\,\rangle = 0.$$

Poiche' il prodotto scalare e' definito positivo cio' implica che $g(\mathbf{v}) - h(\mathbf{v}) = \mathbf{0}$, cioe' che $g(\mathbf{v}) = h(\mathbf{v})$ per ogni $\mathbf{v} \in V$. Questo significa che $h = g$, e prova l'unicita' di g. ■

Denotiamo con f^* l'operatore g definito nella proposizione precedente. L'operatore f^* si dice *l'operatore aggiunto di* f, o anche *l'operatore trasposto di* f. Quest'ultima locuzione e' motivata dal fatto, che consegue dalla stessa dimostrazione precedente, che se A e' la matrice rappresentativa di f rispetto ad una qualunque base *ortonormale* $\mathcal{B}$ allora la matrice rappresentativa di f^* rispetto alla base $\mathcal{B}$ e' A^T. Diremo anche che un operatore $f : V \to V$ e' *autoaggiunto (o simmetrico)* se $f = f^*$. Quindi un operatore f e' autoaggiunto se e solo se per ogni $\mathbf{u}$, $\mathbf{v}$ in V si ha

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

Cio' equivale a dire che la matrice rappresentativa di f rispetto ad una qualunque base ortonormale e' una matrice simmetrica.

Esempio 40. *(i) Nello spazio euclideo* $\mathbf{R}^2$ *munito del prodotto punto, consideriamo l'operatore*

$$f(x, y) := (x + 2y, 3x - 5y).$$

Allora abbiamo

$$f^*(x, y) = (x + 3y, 2x - 5y).$$

Infatti, posto $\mathbf{u} = (x, y)$ *e* $\mathbf{v} = (x', y')$, *si ha*

$$\langle\, f(\mathbf{u}), \mathbf{v}\,\rangle = \langle\, (x + 2y, 3x - 5y), (x', y')\,\rangle = (x + 2y)x' + (3x - 5y)y'$$

$$= x(x' + 3y') + y(2x' - 5y') = \langle\, (x, y), (x' + 3y', 2x' - 5y')\,\rangle = \langle\, \mathbf{u}, f^*(\mathbf{v})\,\rangle.$$

(ii) Nello spazio euclideo $\mathbf{R}^2$ *munito del prodotto punto, l'operatore*

$$f(x, y) := (x + 2y, 2x + 24y)$$

e' autoaggiunto.

(iii) Sia G una matrice reale e simmetrica $n \times n$. Nello spazio euclideo $\mathbf{R}^n$ munito del prodotto punto, l'operatore

$$\mathbf{x} \in \mathbf{R}^n \to G\mathbf{x} \in \mathbf{R}^n$$

e' autoaggiunto.

(iv) Ogni operatore di proiezione ortogonale e' un operatore autoaggiunto. ■

Per gli operatori autoaggiunti vale il seguente importante

Teorema degli assi principali. *Sia $(V, \langle\ ,\ \rangle)$ uno spazio euclideo, ed $f : V \to V$ un operatore autoaggiunto. Allora esiste una base ortonormale $\mathcal{B}$ di V tale che $M_{\mathcal{B}}^{\mathcal{B}}(f)$ sia una matrice diagonale. Cioe' esiste una base ortonormale di V formata da autovettori per f. In particolare f e' diagonalizzabile.*

Per dimostrare il Teorema degli assi principali abbiamo bisogno del seguente

Lemma 2. *Sia G una matrice reale simmetrica. Allora tutti gli autovalori di G sono reali.*

Dimostrazione del Lemma 2. Sia $\lambda \in \mathbf{C}$ un autovalore per G. Dobbiamo dimostrare che $\lambda \in \mathbf{R}$. Per fare cio' proveremo che $\lambda = \lambda^c$ (λ^c significa *coniugato* di λ). A tale proposito consideriamo l'operatore

$$\mathbf{x} \in \mathbf{C}^n \to G\mathbf{x} \in \mathbf{C}^n,$$

dove n denota l'ordine di G. In corrispondenza dell'autovalore λ esiste un autovettore $\mathbf{x} \in \mathbf{C}^n$, cioe' un vettore non nullo $\mathbf{x} \in \mathbf{C}^n$ tale che $G\mathbf{x} = \lambda\mathbf{x}$. Ora andiamo a calcolare il numero complesso

$$(\mathbf{x}^c)^T G\mathbf{x}$$

in due modi diversi. Innanzitutto possiamo scrivere

$$(*) \qquad (\mathbf{x}^c)^T G\mathbf{x} = (\mathbf{x}^c)^T(\lambda\mathbf{x}) = \lambda\left[(\mathbf{x}^c)^T\mathbf{x}\right].$$

D'altra parte poiche' $G\mathbf{x} = \lambda\mathbf{x}$ allora

$$\lambda\mathbf{x}^T = (\lambda\mathbf{x})^T = (G\mathbf{x})^T = \mathbf{x}^T G^T = \mathbf{x}^T G$$

(nell'ultimo passaggio abbiamo usato l'ipotesi che G e' simmetrica) da cui, essendo G reale,

$$(\mathbf{x}^c)^T G = (\mathbf{x}^T G)^c = (\lambda\mathbf{x}^T)^c = \lambda^c(\mathbf{x}^c)^T.$$

Quindi abbiamo anche

$$(\mathbf{x}^c)^T G\mathbf{x} = \lambda^c\left[(\mathbf{x}^c)^T\mathbf{x}\right].$$

E confrontando con (*) deduciamo

$$\lambda \left[(\mathbf{x}^c)^T \mathbf{x} \right] = \lambda^c \left[(\mathbf{x}^c)^T \mathbf{x} \right].$$

Poiche' $\mathbf{x} \neq \mathbf{0}$ allora $\left[(\mathbf{x}^c)^T \mathbf{x} \right] \neq 0$, e quindi $\lambda = \lambda^c$. Cio' prova che $\lambda \in \mathbf{R}$ e conclude la dimostrazione del Lemma 2.

Siamo in condizione di dimostrare il Teorema degli assi principali. Procediamo per induzione sulla dimensione n di V. Il caso $n = 1$ e' ovvio, per cui possiamo assumere $n \geq 2$, e supporre il Teorema vero per ogni spazio euclideo di dimensione $\leq n - 1$.

Sia $\mathcal{E}$ una base ortonormale di V, ed $M_{\mathcal{E}}^{\mathcal{E}}(f)$ la matrice rappresentativa. Tale matrice e' simmetrica reale, percio' per il lemma precedente esiste un autovalore reale λ per f. In corrispondenza di λ esiste un vettore $\mathbf{b}_1 \in V$ non nullo tale che $f(\mathbf{b}_1) = \lambda \mathbf{b}_1$. Possiamo assumere $\|\mathbf{b}_1\| = 1$. Denotiamo con U il sottospazio generato da $\mathbf{b}_1$. Sappiamo che $V = U \perp U^{\perp}$, e che dunque $U^{\perp}$ ha dimensione $n - 1$.

Andiamo a provare che $U^{\perp}$ e' $f-$invariante. Tenuto conto che un vettore $\mathbf{b}$ sta in $U^{\perp}$ se e solo se e' ortogonale a $\mathbf{b}_1$, e' sufficiente provare che se $\langle \mathbf{b}, \mathbf{b}_1 \rangle = 0$ allora anche $\langle f(\mathbf{b}), \mathbf{b}_1 \rangle = 0$. Per provare cio' sia $\mathbf{b}$ un vettore in $U^{\perp}$. Allora poiche' f e' simmetrico si ha $\langle f(\mathbf{b}), \mathbf{b}_1 \rangle = \langle \mathbf{b}, f(\mathbf{b}_1) \rangle$ e quindi

$$\langle f(\mathbf{b}), \mathbf{b}_1 \rangle = \langle \mathbf{b}, f(\mathbf{b}_1) \rangle = \langle \mathbf{b}, \lambda \mathbf{b}_1 \rangle = \lambda \langle \mathbf{b}, \mathbf{b}_1 \rangle = 0.$$

Cio' prova che $U^{\perp}$ e' $f-$invariante.

Poiche' $U^{\perp}$ e' $f-$invariante possiamo considerare la restrizione di f su $U^{\perp}$, cioe' l'operatore $\varphi : \mathbf{b} \in U^{\perp} \to f(\mathbf{b}) \in U^{\perp}$. Ora $U^{\perp}$ e' anch'esso uno spazio euclideo, di dimensione $n - 1$, e φ e' un operatore autoaggiunto in quanto agisce come f. Per cui per ipotesi induttiva esiste una base ortonormale $\mathcal{B}'$ di $U^{\perp}$ formata da autovettori per φ. Ma allora

$$\{\mathbf{b}_1\} \cup \mathcal{B}'$$

e' una base ortonormale di V, formata da autovettori per f. Cio' conclude la dimostrazione del Teorema degli assi principali.

Ora andiamo a provare alcune conseguenze del Teorema degli assi principali.

Corollario 7. *Sia G una matrice simmetrica reale $(n \times n)$. Allora esiste una matrice ortogonale P (cioe' una matrice per cui $P^T P = P P^T = I$) tale che la matrice $D := P^{-1} G P$ (che e' uguale a $P^T G P$) sia diagonale. Sulla diagonale principale di D compaiono tutti gli autovalori (reali) di G, ed il rango di G coincide con il numero di autovalori non nulli (contati con la relativa molteplicita'), l'indice di G coincide con il numero di autovalori positivi (contati con la relativa molteplicita'), e la segnatura di G coincide con la differenza tra il numero di autovalori positivi ed il numero di autovalori negativi (contati con la relativa molteplicita').*

Dimostrazione. Si consideri l'operatore $f : \mathbf{x} \in \mathbf{R}^n \to G\mathbf{x} \in \mathbf{R}^n$. Poiche' G e' simmetrica allora f e' un operatore autoaggiunto per lo spazio euclideo $\mathbf{R}^n$ munito del prodotto punto. Per il Teorema degli assi principali esiste una base ortonormale $\mathcal{B}$ di

$\mathbf{R}^n$ tale che $D := M^{\mathcal{B}}_{\mathcal{B}}(f)$ e' una matrice diagonale. Posto $P := M^{\mathcal{B}}_{\mathcal{E}}(id_{\mathbf{R}^n})$ ($\mathcal{E}$ denota la base canonica), allora P e' una matrice ortogonale e

$$D = M^{\mathcal{E}}_{\mathcal{B}}(id_{\mathbf{R}^n})M^{\mathcal{E}}_{\mathcal{E}}(f)M^{\mathcal{B}}_{\mathcal{E}}(id_{\mathbf{R}^n}) = P^{-1}GP.$$

Poiche' $D = P^{-1}GP$ allora G e' simile a D e sulla diagonale di D appaiono gli autovalori di G. D'altra parte si ha anche $D = P^TGP$ e quindi G e' congruente a D, e le entrate sulla diagonale di D consentono di calcolare il rango e l'indice di G. ■

Si osservi che la matrice P cui si perviene applicando l'algoritmo di Gauss-Lagrange alla matrice $[G \,|\, I]$ non e' la matrice che interviene nell'enunciato del Corollario 7.

Osserviamo anche che il Corollario 7 appare come un nuovo algoritmo per calcolare l'indice di una matrice simmetrica G: esso coincide con il numero degli autovalori positivi di G, contati con la debita molteplicita'. Per calcolare tale numero si puo' calcolare il polinomio caratteristico $p_G(t)$ di G ed applicare il Criterio di Cartesio che dice quanto segue: *sia $p(t) = a_0 + a_1t + \cdots + a_nt^n$ un polinomio a coefficienti reali, con radici tutte reali; scritta la successione ordinata dei coefficienti non nulli di $p(t)$, si conti il numero ν delle variazioni di segno nel passaggio da ciascun coefficiente al successivo; allora ν e' uguale al numero delle radici positive di $p(t)$, contate con la debita molteplicita'.*

Corollario 8 (Diagonalizzazione simultanea di due matrici simmetriche). *Siano A e B due matrici reali e simmetriche $n \times n$. Si assuma $A > 0$. Allora esiste una matrice invertibile P tale che P^TBP sia una matrice diagonale e $P^TAP = I$.*

Dimostrazione. Poiche' A e' definita positiva allora la sua forma canonica rispetto alla congruenza e' I, cioe' esiste una matrice invertibile Q tale che $Q^TAQ = I$. Consideriamo la matrice simmetrica $G := Q^TBQ$. Per il Corollario 7 sappiamo che esiste una matrice ortogonale R tale che la matrice $D := R^TGR$ e' diagonale. Poniamo allora $P := QR$. Abbiamo

$$P^TAP = (QR)^TA(QR) = R^TQ^TAQR = R^TR = I$$

e

$$P^TBP = R^TQ^TBQR = R^TGR = D. \quad ■$$

Corollario 9 (Diagonalizzazione simultanea di due forme quadratiche). *Sia (V, q) uno spazio euclideo, e $p : V \to \mathbf{R}$ una forma quadratica. Allora esiste una base ortonormale $\mathcal{B}$ per (V, q) tale che la matrice di Gram $G^{\mathcal{B}}_{\mathcal{B}}(p)$ sia una matrice diagonale.*

Dimostrazione. Sia $\mathcal{E}$ una base di V. Poniamo $A := G^{\mathcal{E}}_{\mathcal{E}}(q)$ (che e' una matrice definita positiva) e $B := G^{\mathcal{E}}_{\mathcal{E}}(p)$. Per il Corollario 8 esiste una matrice invertibile P tale che $P^TAP = I$ e tale che $D := P^TBP$ sia una matrice diagonale. Sia $\mathcal{B}$ la base di V tale che $P = M^{\mathcal{B}}_{\mathcal{E}}(id_V)$. Allora

$$G^{\mathcal{B}}_{\mathcal{B}}(q) = M^{\mathcal{B}}_{\mathcal{E}}(id_V)^TG^{\mathcal{E}}_{\mathcal{E}}(q)M^{\mathcal{B}}_{\mathcal{E}}(id_V) = P^TAP = I$$

e quindi $\mathcal{B}$ e' una base ortonormale per (V, q), e si ha anche

$$G^{\mathcal{B}}_{\mathcal{B}}(p) = M^{\mathcal{B}}_{\mathcal{E}}(id_V)^TG^{\mathcal{E}}_{\mathcal{E}}(p)M^{\mathcal{B}}_{\mathcal{E}}(id_V) = P^TBP = D. \quad ■$$

Il seguente ulteriore corollario, noto come il Teorema spettrale, ci dice che un operatore simmetrico ammette una semplice descrizione geometrica in termini di proiezioni ortogonali.

Corollario 10 (Il Teorema spettrale). *Sia $(V, \langle\ ,\ \rangle)$ uno spazio euclideo, ed $f : V \to V$ un operatore autoaggiunto. Siano $\lambda_1, \ldots, \lambda_h$ gli autovalori reali distinti di f, e $V_{\lambda_1}, \ldots, V_{\lambda_h}$ gli autospazi corrispondenti. Allora*

$$V = V_{\lambda_1} \perp \cdots \perp V_{\lambda_h},$$

cioe' V e' la somma diretta ortogonale degli autospazi di f. Quindi l'operatore f agisce al seguente modo. Per ogni vettore $\mathbf{v} \in V$ innanzitutto decompone $\mathbf{v}$ nelle sue proiezioni ortogonali sugli autospazi $\mathbf{v} = p_{V_{\lambda_1}}(\mathbf{v}) + \cdots + p_{V_{\lambda_h}}(\mathbf{v})$, e poi ricompone tali proiezioni dopo averle pesate con i corrispondenti autovalori. Cioe'

$$\mathbf{v} = p_{V_{\lambda_1}}(\mathbf{v}) + \cdots + p_{V_{\lambda_h}}(\mathbf{v}) \xrightarrow{\ f\ } f(\mathbf{v}) = \lambda_1 p_{V_{\lambda_1}}(\mathbf{v}) + \cdots + \lambda_h p_{V_{\lambda_h}}(\mathbf{v}).$$

L'espressione

$$f(\mathbf{v}) = \lambda_1 p_{V_{\lambda_1}}(\mathbf{v}) + \cdots + \lambda_h p_{V_{\lambda_h}}(\mathbf{v})$$

si chiama la decomposizione spettrale di f.

Dimostrazione. Per il Teorema degli assi principali sappiamo che esiste una base ortonormale $\mathcal{B}$ di V che diagonalizza f. Siano $\mathbf{b}^i_1, \ldots, \mathbf{b}^i_{m_i}$ $(m_i := m_a(\lambda_i))$ gli autovettori di f relativi all'autovalore λ_i che appaiono nella base ortonormale $\mathcal{B}$. Poiche' f e' diagonalizzabile allora $m_i = m_a(\lambda_i) = m_g(\lambda_i)$. Per cui i vettori $\mathbf{b}^i_1, \ldots, \mathbf{b}^i_{m_i}$ formano una base di V_{λ_i}. Poiche' $\mathcal{B}$ e' la riunione di tutti i vettori $\mathbf{b}^i_1, \ldots, \mathbf{b}^i_{m_i}$ (per $i = 1, \ldots, h$) allora $V = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_h}$, e tale somma e' ortogonale perche' $\mathcal{B}$ e' una base ortonormale. Infine la decomposizione spettrale segue dal fatto appena provato che $V = V_{\lambda_1} \perp \cdots \perp V_{\lambda_h}$, da quanto visto nell'Osservazione 4, e dal fatto che i vettori in V_{λ_i} sono tutti autovettori. ■

Esempio 41. *Utilizzando il Criterio di Cartesio calcolare l'indice della matrice*

$$G := \begin{bmatrix} -2 & -3 & -1 \\ -3 & -5 & -1 \\ -1 & -1 & 1 \end{bmatrix}.$$

Svolgimento. Il polinomio caratteristico di G e' $p_G(t) = 2 + 8t - 6t^2 - t^3$. La sequenza ordinata dei coefficienti non nulli presenta una sola variazione, dunque l'indice di G e' 1. ■

Esempio 42. *Utilizzando il Criterio di Cartesio calcolare l'indice della matrice*

$$G := \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & -1 \\ 2 & -1 & 5 \end{bmatrix}.$$

Svolgimento. Il polinomio caratteristico di G e' $p_G(t) = -6t + 7t^2 - t^3$. La sequenza ordinata dei coefficienti non nulli presenta due variazioni, dunque l'indice di G e' 2. ■

Esempio 43. *Si consideri la matrice*

$$G := \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & -1 \\ 2 & -1 & 5 \end{bmatrix}.$$

Determinare una matrice ortogonale P tale che $P^{-1}GP$ sia diagonale.

Svolgimento. Il polinomio caratteristico di G e' $p_G(t) = -6t+7t^2-t^3 = -t(t-1)(t-6)$. L'autospazio V_0 e' generato da $(-2,1,1)$, l'autospazio V_1 e' generato da $(1,2,0)$, e l'autospazio V_6 e' generato da $(2,-1,5)$. Si osservi che, come annunciato dal Teorema spettrale, gli autospazi sono mutualmente ortogonali. Normalizzando i vettori precedenti otteniamo la matrice cercata

$$P := \begin{bmatrix} -\frac{2}{\sqrt{6}} & \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{30}} \\ \frac{1}{\sqrt{6}} & \frac{2}{\sqrt{5}} & -\frac{1}{\sqrt{30}} \\ \frac{1}{\sqrt{6}} & 0 & \frac{5}{\sqrt{30}} \end{bmatrix}.$$

In particolare si ha

$$P^{-1}GP = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 6 \end{bmatrix}. \quad ■$$

Esempio 44. *Determinare la decomposizione spettrale dell'operatore*

$$f(x,y,z) := (x+2z, y-z, 2x-y+5z).$$

Svolgimento. La matrice rappresentativa di f rispetto alla base canonica e' proprio la matrice G dell'esempio precedente. Quindi sappiamo che

$$\mathbf{R}^3 = V_6 \perp V_1 \perp V_0 = Span((2,-1,5)) \perp Span((1,2,0)) \perp Span((-2,1,1)).$$

Ne consegue che la decomposizione cercata e'

$$f(x,y,z) = 6p_{(2,-1,5)}(x,y,z) + p_{(1,2,0)}(x,y,z) =$$

$$\frac{1}{5}(2x-y+5z)(2,-1,5) + \frac{1}{5}(x+2y)(1,2,0). \quad ■$$

Un'altra interessante conseguenza del Teorema degli assi principali consiste nel fornire un algoritmo per il calcolo dei massimi e dei minimi di una forma quadratica. Piu' precisamente, consideriamo lo spazio euclideo $\mathbf{R}^n$ munito del prodotto punto. Denotiamo con $\| \ \|$ la norma in tale spazio. Poi consideriamo una qualunque forma quadratica $q : \mathbf{R}^n \to \mathbf{R}$. Risulta di un certo interesse il seguente problema: determinare, tra tutti i vettori $\mathbf{u}$ tali che $\|\mathbf{u}\| \leq 1$, il valore massimo M assunto da $q(\mathbf{u})$, cioe' calcolare

$$M := \max_{\|\mathbf{u}\| \leq 1} q(\mathbf{u}).$$

Illustriamo con un esempio come si puo' risolvere questo problema utilizzando il Teorema degli assi principali.

Esempio 45 (massimo vincolato di una forma quadratica). *Si consideri la forma quadratica $q : \mathbf{R}^3 \to \mathbf{R}$ definita ponendo*

$$q(\mathbf{u}) := x_1^2 + 4x_1x_3 + x_2^2 - 2x_2x_3 + 5x_3^2.$$

(a) Calcolare $M := \max \{q(\mathbf{u}) : \|\mathbf{u}\| \leq 1\}$;

(b) determinare tutti i vettori $\mathbf{u} \in \mathbf{R}^3$ per i quali $\|\mathbf{u}\| \leq 1$ e $q(\mathbf{u}) = M$.

Svolgimento. La matrice di Gram di q rispetto alla base canonica e'

$$G := G_{\mathcal{E}}^{\mathcal{E}}(q) = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & -1 \\ 2 & -1 & 5 \end{bmatrix}.$$

In base agli esempi precedenti sappiamo che la matrice ortogonale

$$P := \begin{bmatrix} -\frac{2}{\sqrt{6}} & \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{30}} \\ \frac{1}{\sqrt{6}} & \frac{2}{\sqrt{5}} & -\frac{1}{\sqrt{30}} \\ \frac{1}{\sqrt{6}} & 0 & \frac{5}{\sqrt{30}} \end{bmatrix}$$

soddisfa la condizione

$$P^T G P = D := \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 6 \end{bmatrix}.$$

Questa formula ci dice che il cambiamento di coordinate $\mathbf{x} = P\mathbf{x}'$ nell'espressione esplicita di q diagonalizza q, cioe'

$$q(\mathbf{u}) := x_1^2 + 4x_1x_3 + x_2^2 - 2x_2x_3 + 5x_3^2 = {x_2'}^2 + 6{x_3'}^2.$$

E' evidente che

$${x_2'}^2 + 6{x_3'}^2 \leq 6({x_1'}^2 + {x_2'}^2 + {x_3'}^2).$$

D'altra parte P e' una matrice ortogonale, quindi le sue colonne formano una base ortonormale per $(\mathbf{R}^n, \text{prodotto punto})$, e percio'

$${x_1'}^2 + {x_2'}^2 + {x_3'}^2 = \|\mathbf{u}\|^2.$$

Se si vincola $\mathbf{u}$ *ad avere lunghezza* ≤ 1 *allora* $\|\mathbf{u}\|^2 \leq 1$ *e in conclusione otteniamo*

$$q(\mathbf{u}) \leq 6.$$

Cio' prova che $M \leq 6$*. Poi osserviamo che*

$$q(\frac{2}{\sqrt{30}}, -\frac{1}{\sqrt{30}}, \frac{5}{\sqrt{30}}) = (x_2'^2 + 6x_3'^2)_{|\mathbf{x}'=(0,0,1)^T} = 6$$

il che prova che $M = 6$ *(in generale* M *coincide con l'autovalore massimo di* G*). Cio' risponde alla prima domanda.*

Quanto alla seconda domanda, sia $\mathbf{u}$ *un vettore di lunghezza* ≤ 1 *tale che* $q(\mathbf{u}) = M = 6$*. Allora abbiamo*

$$6 = q(\mathbf{u}) = x_2'^2 + 6x_3'^2 \leq 6(x_1'^2 + x_2'^2 + x_3'^2) = 6\|\mathbf{u}\|^2 \leq 6.$$

Percio' deve essere

$$q(\mathbf{u}) = x_2'^2 + 6x_3'^2 = 6(x_1'^2 + x_2'^2 + x_3'^2) = 6\|\mathbf{u}\|^2 = 6.$$

Da cui deduciamo che $\|\mathbf{u}\| = 1$ *e che* $6x_1'^2 + 5x_2'^2 = 0$*, cioe'* $\mathbf{x}' = (0, 0, \pm 1)^T$*. Quindi ci sono solo due vettori di lunghezza* ≤ 1 *che raggiungono il massimo* $M = 6$ *e sono i versori di* V_6*, cioe'*

$$\mathbf{u} = \pm(\frac{2}{\sqrt{30}}, -\frac{1}{\sqrt{30}}, \frac{5}{\sqrt{30}}). \quad \blacksquare$$

13. La decomposizione ai valori singolari.

I risultati sulle matrici simmetriche stabiliti nei paragrafi precedenti ammettono delle generalizzazioni per matrici A non necessariamente simmetriche ne' quadrate. Per esempio e' possibile dimostrare una sorta di Teorema degli assi principali per matrici non quadrate, detto *Teorema di decomposizione ai valori singolari.* L'idea consiste nel fatto che, nonostante A sia una matrice qualunque, tuttavia A^TA e AA^T sono sempre matrici simmetriche. A tali matrici si possono quindi applicare i teoremi sulle matrici simmetriche, che restituiscono informazioni sulla matrice originaria A.

Piu' precisamente possiamo procedere nel seguente modo. Sia A una matrice reale $m \times n$. Allora la matrice A^TA e' una matrice quadrata $n \times n$, ed e' anche simmetrica in quanto $(A^TA)^T = A^T(A^T)^T = A^TA$. Per il Teorema degli assi principali sappiamo allora che esiste una matrice ortogonale Q (quindi $Q^{-1} = Q^T$) di ordine n tale che $Q^T(A^TA)Q$ sia una matrice diagonale. Denotiamo con $\mu_1, \dots, \mu_n$ la diagonale principale di $Q^T(A^TA)Q$. Quindi abbiamo

$$Q^T(A^TA)Q = \begin{bmatrix} \mu_1 & & & & \\ & \mu_2 & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & \mu_n \end{bmatrix}.$$

Scambiando eventualmente le colonne $\mathbf{q}^1, \mathbf{q}^2, \dots, \mathbf{q}^n$ di Q possiamo assumere che

$$\mu_1 \geq \mu_2 \geq \cdots \geq \mu_n.$$

Andiamo a provare che $\mu_n \geq 0$, cioe' che tutti gli autovalori di $A^T A$ sono non negativi. A tale proposito, tenuto conto che $(A^T A)\mathbf{q}^j = \mu_j \mathbf{q}^j$ in quanto le colonne di Q sono autovettori per $A^T A$, andiamo a calcolare il prodotto scalare $\langle A\mathbf{q}^i, A\mathbf{q}^j \rangle$. Si ha

$$\langle A\mathbf{q}^i, A\mathbf{q}^j \rangle = (A\mathbf{q}^i)^T (A\mathbf{q}^j) = {\mathbf{q}^i}^T (A^T A)\mathbf{q}^j = {\mathbf{q}^i}^T (\mu_j \mathbf{q}^j) = \mu_j \langle \mathbf{q}^i, \mathbf{q}^j \rangle.$$

Poiche' le colonne di Q formano una base ortonormale per $\mathbf{R}^n$ deduciamo che

$$\langle A\mathbf{q}^i, A\mathbf{q}^j \rangle = \begin{cases} 0 & \text{se } i \neq j \\ \mu_j & \text{se } i = j. \end{cases}$$

In particolare

$$\mu_j = \langle A\mathbf{q}^j, A\mathbf{q}^j \rangle = \|A\mathbf{q}^j\|^2$$

da cui $\mu_j \geq 0$ per ogni $j = 1, \dots, n$.

Poiche' $\mu_j \geq 0$ possiamo definire i numeri reali

$$\sigma_j := \sqrt{\mu_j} = \|A\mathbf{q}^j\|.$$

Tali numeri sono detti anche *i valori singolari di A*. Se denotiamo con $r + 1$ il primo indice j per cui $\sigma_j = 0$, possiamo scrivere la sequenza dei valori singolari di A nel seguente modo

$$\sigma_1 \geq \sigma_2 \geq \dots \sigma_r > \sigma_{r+1} = \sigma_{r+2} = \cdots = \sigma_n = 0.$$

La seguente proposizione ci consente di identificare il numero r come il rango di A.

Proposizione 7. *I vettori numerici $A\mathbf{q}^1, \dots, A\mathbf{q}^r$ formano una base ortogonale per lo spazio delle colonne di A. In particolare r e' il rango di A.*

Per dimostrare cio' abbiamo bisogno della seguente osservazione:

Lemma 3. *Siano $\mathbf{u}_1, \dots, \mathbf{u}_r$ vettori non nulli e mutualmente ortogonali di uno sapzio euclideo V. Allora $\mathbf{u}_1, \dots, \mathbf{u}_r$ sono linearmente indipendenti.*

Dimostrazione del Lemma 3. Sia $a_1\mathbf{u}_1 + \cdots + a_r\mathbf{u}_r = \mathbf{0}$ una relazione tra i vettori assegnati. Per ogni $j = 1, \dots, r$ abbiamo

$$0 = \langle a_1\mathbf{u}_1 + \cdots + a_r\mathbf{u}_r, \mathbf{u}_j \rangle = \sum_{i=1}^{n} a_i \langle \mathbf{u}_i, \mathbf{u}_j \rangle = a_j \langle \mathbf{u}_j, \mathbf{u}_j \rangle.$$

Cio' implica $a_j = 0$ poiche' $\mathbf{u}_j \neq \mathbf{0}$, e conclude la dimostrazione del Lemma 3.

Possiamo ora dimostrare la Proposizione 7. Abbiamo visto che se $i \neq j$ allora $\langle A\mathbf{q}^i, A\mathbf{q}^j \rangle = 0$, per cui i vettori $A\mathbf{q}^1, \ldots, A\mathbf{q}^r$ sono mutualmente ortogonali. Poi, per la stessa definizione di r, sono vettori non nulli. Per il lemma appena dimostrato possiamo dire allora che il sistema $A\mathbf{q}^1, \ldots, A\mathbf{q}^r$ e' linearmente indipendente. Rimane da provare che tale sistema genera lo spazio delle colonne di A.

A tale proposito consideriamo l'operatore $\mathbf{x} \in \mathbf{R}^n \to A\mathbf{x} \in \mathbf{R}^m$. Sappiamo che lo spazio delle colonne di A e' lo spazio immagine di tale operatore. Quindi se $\mathbf{y}$ e' il generico vettore dello spazio delle colonne di A, esiste un vettore $\mathbf{x} \in \mathbf{R}^n$ tale che $\mathbf{y} = A\mathbf{x}$. Poiche' le colonne di Q formano una base di $\mathbf{R}^n$ potremo scrivere $\mathbf{x} = x'_1\mathbf{q}^1 + \cdots + x'_n\mathbf{q}^n$ secondo opportune coordinate x'_i. Quindi

$$\mathbf{y} = A\mathbf{x} = A(x'_1\mathbf{q}^1 + \cdots + x'_n\mathbf{q}^n) = x'_1A\mathbf{q}^1 + \cdots + x'_nA\mathbf{q}^n = x'_1A\mathbf{q}^1 + \cdots + x'_rA\mathbf{q}^r$$

in quanto sappiamo che per $j > r$ $A\mathbf{q}^j = \mathbf{0}$. Cio' prova che $\mathbf{y}$ e' combinazione lineare dei vettori $A\mathbf{q}^1, \ldots, A\mathbf{q}^r$, e conclude la dimostrazione della Proposizione 7.

In base alla proposizione appena dimostrata possiamo dire che i vettori colonna

$$\frac{1}{\sigma_1}A\mathbf{q}^1, \ldots, \frac{1}{\sigma_r}A\mathbf{q}^r$$

formano una base ortonormale per lo spazio delle colonne di A. Utilizzando il procedimento di ortogonalizzazione di Gram-Schmidt possiamo aggiungere alle precedenti colonne opportune $m - r$ colonne ortonormali

$$\frac{1}{\sigma_1}A\mathbf{q}^1, \ldots, \frac{1}{\sigma_r}A\mathbf{q}^r, \mathbf{p}^{r+1}, \ldots, \mathbf{p}^m$$

in modo da formare una matrice ortonormale P di ordine m.

Ora definiamo la seguente matrice $m \times n$

$$\Sigma = \begin{bmatrix} \sigma_1 & & & & & & & & \\ & \sigma_2 & & & & & & & \\ & & \cdot & & & & & & \\ & & & \cdot & & & & & \\ & & & & \sigma_r & & & & \\ & & & & & 0 & & & \\ & & & & & & \cdot & & \\ & & & & & & & \cdot & \\ & & & & & & & & 0 \end{bmatrix}.$$

Tenuto conto delle definizioni adottate, un confronto diretto mostra che $AQ = P\Sigma$. Resta pertanto dimostrato il seguente importante

Teorema di decomposizione ai valori singolari. *Sia A una matrice qualunque, e siano P, Σ e Q le matrici definite in precedenza a partire da A. Allora si ha*

$$A = P\Sigma Q^T. \quad \blacksquare$$

Come immediata conseguenza possiamo dimostrare il seguente

Teorema di decomposizione polare. *Sia A una matrice quadrata nonsingolare. Allora esistono una matrice simmetrica S definita positiva ed una matrice ortogonale O tali che*

$$A = S \cdot O. \quad \blacksquare$$

Infatti se $A = P\Sigma Q^T$ e' la decomposizione ai valori singolari di A allora possiamo scrivere

$$A = P\Sigma Q^T = (P\Sigma P^T)(PQ^T),$$

e porre $S := P\Sigma P^T$ ed $O := PQ^T$.

Osservazione 4. *(i) La dimostrazione del Teorema di decomposizione ai valori singolari ci consente di calcolare esplicitamente le matrici P, Σ e Q. Quindi e' un nuovo algoritmo. Lo stesso si puo' dire per la decomposizione polare.*

(ii) Si puo' dimostrare che la decomposizione polare $A = S \cdot O$ e' unica.

(iii) La dimostrazione della decomposizione polare si applica anche se A e' singolare, producendo una decomposizione del tipo $A = S \cdot O$ con S semidefinita positiva (ed O ortogonale). In questo caso viene meno l'unicita'.

(iv) La decomposizione polare consente di interpretare geometricamente un qualsiasi endomorfismo $f : V \to V$. Infatti la decomposizione polare ci dice che f si puo' decomporre nel prodotto $f = g \circ h$ di due operatori g ed h, con g operatore autoaggiunto (che si descrive tramite la decomposizione spettrale in termini di proiezioni ortogonali), ed h operatore con matrice rappresentativa data da una matrice ortogonale. Tali operatori si chiamano isometrie e si possono descrivere geometricamente. Nel caso $\dim V = 2$ abbiamo gia' visto che h e' una rotazione oppure una riflessione (simmetria ortogonale). Quindi ogni operatore $f : V \to V$ di uno spazio euclideo di dimensione 2 si puo' decomporre nella composizione di una rotazione o una riflessione, con un operatore combinazione lineare di proiezioni ortogonali. Esiste una descrizione analoga in ogni dimensione.

(v) Se A e' una matrice simmetrica $n \times n$ con autovalori(eventualmente ripetuti) $\lambda_1, \lambda_2, \dots, \lambda_n$, allora per ogni $i = 1, \dots, n$ si ha $\sigma_i = |\lambda_i|$. Cioe' i valori singolari di una matrice simmetrica sono i valori assoluti dei suoi autovalori. Infatti se R e' una matrice ortogonale tale che la matrice $D = R^{-1}AR$ sia diagonale allora $R^T A^T AR = D^2$, cioe' i numeri μ_i sono i quadrati λ_i^2 delle componenti λ_i della diagonale principale di D. Ma tali componenti λ_i altro non sono che gli autovalori di A.

(vi) Sia $A = P\Sigma Q^T$ una decomposizione ai valori singolari per una matrice A $m \times n$. Allora $AA^T = P\Sigma\Sigma^T P^T$. Quindi gli autovalori di AA^T sono le componenti sulla diagonale principale di $\Sigma\Sigma^T$, che sono $\sigma_1^2, \sigma_2^2, \dots, \sigma_r^2$, insieme ad $m - r$ componenti nulle. Quindi i valori singolari non nulli di A si possono calcolare diagonalizzando indifferentemente A^TA o AA^T. In particolare A, A^TA ed AA^T hanno tutte lo stesso

rango r, e le matrici simmetriche A^TA ed AA^T hanno gli stessi autovalori non nulli, con le stesse molteplicita'. Quindi, se il polinomio caratteristico di A^TA e' del tipo $p_{A^TA}(t) = q(t)t^{n-r}$ con $q(0) \neq 0$, allora $p_{AA^T}(t) = q(t)t^{m-r}$.

Esempio 46. *Determinare una decomposizione ai valori singolari per la matrice*

$$A := \begin{bmatrix} 1 & 2 & 0 \\ 2 & 1 & 0 \end{bmatrix}.$$

Svolgimento. *Innanzitutto calcoliamo*

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

Il polinomio caratteristico di A^TA e' $p(t) = -t(t-1)(t-9)$. Quindi i valori singolari di A sono

$$\sigma_1 = 3,\ \sigma_2 = 1,\ \sigma_3 = 0,$$

e la matrice Σ e'

$$\Sigma = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

Per calcolare Q andiamo a diagonalizzare A^TA con una matrice ortogonale. I suoi autospazi sono

$$V_9 = Span(1,1,0), \quad V_1 = Span(1,-1,0), \quad V_0 = Span(0,0,1).$$

Quindi la matrice Q e':

$$Q = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Poi calcoliamo la matrice P che e' la matrice 2×2 con colonne $\frac{1}{\sigma_1}A\mathbf{q}^1, \frac{1}{\sigma_2}A\mathbf{q}^2$ cioe'

$$P = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}.$$

In conclusione la decomposizione ai valori singolari cercata e':

$$A = P\Sigma Q^T = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \cdot \begin{bmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad \blacksquare$$

Esempio 47. *Determinare la decomposizione polare della matrice*

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

Svolgimento. *Innanzitutto calcoliamo una decomposizione ai valori singolari per A che e':*

$$A = P\Sigma Q^T = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \cdot \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}.$$

Allora

$$S = P\Sigma P^T = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$$

ed

$$O = PQ^T = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

Quindi la decomposizione polare di A e':

$$A = \begin{bmatrix} 1 & 2 \\ -2 & -1 \end{bmatrix} = S \cdot O = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \cdot \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}. \quad \blacksquare$$

Esempio 48. *Descrivere geometricamente l'operatore $f : \mathbf{R}^2 \to \mathbf{R}^2$ definito ponendo*

$$f(x, y) := (x + 2y, -2x - y).$$

Svolgimento. *Sia A la matrice rappresentativa di f rispetto alla base canonica, che coincide con quella dell'esempio precedente, e di cui conosciamo la decomposizione polare. Quindi $f = g \circ h$ dove h e' la rotazione in senso antiorario di $\frac{3}{2}\pi$:*

$$h : (x, y) \in \mathbf{R}^2 \to (y, -x) \in \mathbf{R}^2$$

e g e' quell'operatore autoaggiunto che si ottiene combinando la proiezione ortogonale su $(1, -1)$ e la proiezione ortogonale su $(1, 1)$ con pesi 3 ed 1 (tale combinazione e' cio' che abbiamo chiamato la decomposizione spettrale di g):

$$g : (x, y) \in \mathbf{R}^2 \to 3p_{(1,-1)}(x, y) + p_{(1,1)}(x, y) \in \mathbf{R}^2.$$

In particolare

$$f(x, y) = (x + 2y, -2x - y) = 3p_{(1,-1)}(y, -x) + p_{(1,1)}(y, -x). \quad \blacksquare$$

Esempio 49. *Sia A una matrice $m \times n$. Sia σ_1 il massimo valore singolare di A. Provare che σ_1 e' il massimo valore assunto dalla lunghezza del vettore $A\mathbf{x}$, al variare di $\mathbf{x}$ tra i vettori di $\mathbf{R}^n$ di lunghezza ≤ 1. Cioe'*

$$\sigma_1 = \max_{\|\mathbf{x}\| \leq 1} \|A\mathbf{x}\|.$$

Svolgimento. E' sufficiente osservare che

$$\max_{\|\mathbf{x}\| \leq 1} \|A\mathbf{x}\| = \max_{\|\mathbf{x}\| \leq 1} \sqrt{\langle A\mathbf{x}, A\mathbf{x} \rangle} = \max_{\|\mathbf{x}\| \leq 1} \sqrt{\mathbf{x}^T (A^T A)\mathbf{x}}$$

e ricordare che la forma quadratica $q(\mathbf{x}) := \mathbf{x}^T (A^T A)\mathbf{x}$ assume come valore massimo il massimo autovalore di $A^T A$ (si veda l'Esempio 45), che e' $\mu_1 = \sigma_1^2$ per definizione di valore singolare. ■

Il numero $\max_{\|\mathbf{x}\| \leq 1} \|A\mathbf{x}\|$ si denota anche con il simbolo

$$\|A\|_2$$

e prende il nome di *norma spettrale di A*. Si puo' provare che la norma spettrale di una matrice gode delle stesse proprieta' formali della norma di uno spazio euclideo.

Bibilografia consigliata per questo capitolo:

S. Abeasis, *Complementi di algebra lineare e geometria*, Ed. Zanichelli.

S. Abeasis, *Elementi di algebra lineare e geometria*, Ed. Zanichelli.

N. Bakhvalov, *Methodes Numeriques*, Ed. Mir.

M. Barnabei e F.Bonetti, *Forme quadratiche e forme bilineari simmetriche*, Ed. Pitagora.

V. A. Ilyin and E. G. Poznyak, *Linear Algebra*, Ed. Mir.

S. Lipschutz, *Algebra lineare*, Ed. Schaum.

S. Mac Lane e G. Birkhoff, *Algebra*, Ed. Mursia.

V.V.Voyevodin, *Linear Algebra*, Ed. Mir.