Spazi vettoriali metrici e teorema spettrale.

1. Prodotti scalari e hermitiani. Spazi vettoriali metrici

Definizione. Sia V uno spazio vettoriale su $\mathbb{R}$ e sia $g: V \times V \to \mathbb{R}$ una funzione. Se per ogni $v, v_1, v_2, w, w_1, w_2 \in V$ e $\lambda \in \mathbb{R}$ si ha

$$g(v_1 + v_2, w) = g(v_1, w) + g(v_2, w), \quad g(v, w) = g(v, w_1) + g(v, w_2),$$

$$g(\lambda v, w) = \lambda g(v, w) = g(v, \lambda w)$$

g si dice *forma bilineare*. Se per ogni $v, w \in V$ si ha

$$g(v, w) = g(w, v)$$

g si dice *simmetrica*. Una forma bilineare simmetrica si dice *prodotto scalare*.

Definizione. Sia V uno spazio vettoriale su $\mathbb{C}$ e sia $h: V \times V \to \mathbb{C}$ una funzione. Se per ogni $v, v_1, v_2, w, w_1, w_2 \in V$ e $\lambda \in \mathbb{R}$ si ha

$$h(v_1 + v_2, w) = h(v_1, w) + h(v_2, w), \quad h(v, w) = h(v, w_1) + h(v, w_2),$$

$$h(\lambda v, w) = \lambda h(v, w), \quad h(v, \lambda w) = \overline{\lambda} h(v, w)$$

g si dice *forma sesquilineare*. Se per ogni $v, w \in V$ si ha

$$g(v, w) = \overline{g(w, v)}$$

g si dice *hermitiana*. Una forma sesquilineare hermitiana si dice *prodotto hermitiano*.

Esempio 1. Il *prodotto scalare canonico* $< \cdot, \cdot >: \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ di $\mathbb{R}^n$ è la forma bilineare simmetrica definita per ogni scelta di $v, w \in \mathbb{R}^n$ da

$$< v, w >= \sum_{i=1}^{n} v_i w_i = w^T v.$$

Sia $A \in M_n(\mathbb{R})$ una matrice simmetrica reale ossia tale che $A = A^T$. Un prodotto scalare $< \cdot, \cdot >_A: \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ è definito da

$$< v, w >_A = < Av, w > .$$

La bilinearità di $< \cdot, \cdot >_A$ segue immediatamente dalla bilinearità del prodotto scalare canonico e dalla distributività del prodotto di matrici. La simmetria segue da:

$$< v, w >_A = < Av, w > = < w, Av > = (Av)^T w = v^T A^T w = v^T A w = < Aw, v > = < w, v >_A .$$

Esempio 2. Il *prodotto hermitiano canonico* su $\mathbb{C}^n$ è la forma sesquilineare hermitiana $< \cdot, \cdot >: \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$ definita per ogni scelta di $v, w \in \mathbb{C}^n$ da

$$< v, w >= \sum_{i=1}^{n} v_i \overline{w}_i = w^H v$$

dove come al solito per una matrice (o per un vettore inteso come matrice con una sola colonna) l'apice H indica la trasposta coniugata. Sia $A \in M_n(\mathbb{R})$ una matrice hermitiana ossia tale che $A = A^H$. Un prodotto hermitiano $< \cdot, \cdot >_A: \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$ è definito da

$$< v, w >_A = < Av, w > .$$

La sesquilinearità di $< \cdot, \cdot >_A$ segue immediatamente dalla sesquilinearità del prodotto hermitiano canonico e dalla distributività del prodotto di matrici. La hermitianità segue da:

$$< v, w >_A = < Av, w > = \overline{< w, Av >} = \overline{(Av)^H w} = \overline{v^H A^H w} = \overline{v^H A w} = \overline{< Aw, v >} = \overline{< w, v >_A}.$$

In realtà si dimostra facilmente che i prodotti scalari e hermitiani definiti negli Esempi 1 e 2 sono tutti quelli possibili:

Proposizione 1.1:

(i) *Sia* $(\cdot, \cdot): \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ *un prodotto scalare e siano* $e_1, \ldots, e_n$ *i vettori della base canonica di* $\mathbb{R}^n$*. Se* $A = (a_{kj})$ *è la matrice simmetrica definita da* $a_{kj} = (e_j, e_k)$*, allora* $(\cdot, \cdot) = < \cdot, \cdot >_A$.

(ii) *Sia* $(\cdot, \cdot): \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$ *un prodotto hermitiano e siano* $e_1, \ldots, e_n$ *i vettori della base canonica di* $\mathbb{C}^n$*. Se* $A = (a_{kj})$ *è la matrice Hermitiana definita da* $a_{kj} = (e_j, e_k)$*, allora* $(\cdot, \cdot) = < \cdot, \cdot >_A$.

Dimostrazione: (i): Siano $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{R}^n$. Allora

$$\begin{aligned}(x, y) &= \left(\sum_{j=1}^{n} x_j e_j, \sum_{k=1}^{n} y_k e_k \right) = \sum_{j,k=1}^{n} x_j y_k (e_j, e_k) \\ &= \sum_{k=1}^{n} y_k \left(\sum_{j=1}^{n} a_{kj} x_j \right) = \sum_{k=1}^{n} y_k (Ax)_k = y^T A x = < x, y >_A .\end{aligned}$$

(ii): Siano $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{C}^n$. Allora

$$\begin{aligned}(x, y) &= \left(\sum_{j=1}^{n} x_j e_j, \sum_{k=1}^{n} y_k e_k \right) = \sum_{j,k=1}^{n} x_j \overline{y}_k (e_j, e_k) \\ &= \sum_{k=1}^{n} \overline{y}_k \left(\sum_{j=1}^{n} a_{kj} x_j \right) = \sum_{k=1}^{n} \overline{y}_k (Ax)_k = y^H A x = < x, y >_A .\end{aligned}$$

$\square$

Esempio 3. Sia $V = C^0([a, b])$ lo spazio vettoriale delle funzioni continue a valori reali definite sull'intervallo $[a, b] \subset \mathbb{R}$. Definiamo $g: V \times V \to \mathbb{R}$ ponendo

$$g(f, g) = \int_a^b f(t)g(t)dt$$

per ogni $f, g \in V$. È immediato verificare che g è un prodotto scalare.

Definizione. Un prodotto scalare (hermitiano) $< \cdot, \cdot >$ su uno spazio vettoriale V si dice *definito positivo* se per ogni $v \in V \setminus \{O\}$ si ha $< v, v >> 0$ e *definito negativo* se per ogni $v \in V \setminus \{O\}$ si ha $< v, v >< 0$. Si dice che $< \cdot, \cdot >$ è *semidefinito positivo* se per ogni $v \in V$ si ha $< v, v >\geq 0$ e esiste $v_0 \in V \setminus \{O\}$ tale che $< v, v >= 0$. Si dice che $< \cdot, \cdot >$ è *semidefinito negativo* se per ogni $v \in V$ si ha $< v, v >\leq 0$ e esiste $v_0 \in V \setminus \{O\}$ tale che $< v, v >= 0$. Infine $< \cdot, \cdot >$ si dice *indefinito* se esistono $v, w \in V$ tali che $< v, v >> 0$ e $< w, w >< 0$.

Esempi. Il prodotto scalare canonico di $\mathbb{R}^n$ e il prodotto hermitiano canonico di $\mathbb{C}^n$ sono definiti positivi. In teoria della Relatività è importante considerare il prodotto di Minkowski, il prodotto scalare su $\mathbb{R}^4$ definito da

$$< \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} >= x_1y_1 + x_2y_2 + x_3y_3 - x_4y_4.$$

Il prodotto di Minkowski è indefinito. Il prodotto scalare definito nell'Esempio 3 sullo spazio delle funzioni continue sull'intervallo $[a, b]$ è definito positivo. Infatti se $f \in C^0([a, b])$ allora $f^2(t) \geq 0$ per $t \in [a, b]$. Dunque $\int_a^b f^2(t)dt \geq 0$ e si dimostra in teoria dell'integrazione che $\int_a^b f^2(t)dt = 0$ se e solo se f^2, e quindi f, è identicamente nulla.

Ci occuperemo prevalentemente di prodotti scalari definiti positivi. Cominciamo introducendo un'ulteriore nozione:

Definizione. Uno *spazio vettoriale metrico reale (complesso)* è uno spazio vettoriale V su $\mathbb{R}$ ($\mathbb{C}$) su cui è definito un prodotto scalare (hermitiano) $< \cdot, \cdot >$ definito positivo. Su uno spazio vettoriale metrico la funzione $\| \cdot \|: V \to \mathbb{R}_+$, definita da $\|v\| = \sqrt{< v, v >}$, si dice *norma*.

Proposizione 1.2: *Sia V uno spazio vettoriale metrico con prodotto $< \cdot, \cdot >$ e norma $\| \cdot \|$. Allora:*

(i) Per ogni $v \in V$ si ha $\|v\| \geq 0$ e $\|v\| = 0$ se e solo se $v = O$.

(ii) Per ogni scalare λ e $v, w \in V$ si ha

$$\|\lambda v\| = |\lambda|\|v\|, \qquad \|v + w\|^2 = \|v\|^2 + \|w\|^2 + 2Re < v, w > .$$

(iii) Per ogni $v, w \in V$ vale la diseguaglianza di Cauchy-Schwarz*:*

$$| < v, w > | \leq \|v\|\|w\| \tag{1.1}$$

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

(iv) *Per ogni* $v, w \in V$ *vale la* diseguaglianza triangolare*:*

$$\|v + w\| \leq \|v\| + \|w\|. \tag{1.2}$$

(v) *Per ogni* $v, w \in V$ *valgono le seguenti* formule di polarizzazione*:*

$$< v, w >= \frac{1}{4}\left(\|v + w\|^2 - \|v - w\|^2\right) \quad \text{per}\, V \,\text{su}\, \mathbb{R} \tag{1.3}$$

$$Re < v, w >= \frac{1}{4}\left(\|v + w\|^2 - \|v - w\|^2\right) \quad \text{e} \quad Im < v, w >= \frac{1}{4}\left(\|v + iw\|^2 - \|v - iw\|^2\right) \quad \text{per}\, V \,\text{su}\, \mathbb{C}. \tag{1.4}$$

Dimostrazione: La (i) segue immediatamente dal fatto che $< \cdot, \cdot >$ è definito positivo. Per (ii) calcoliamo nel caso complesso (quello reale è ancora più semplice):

$$\|\lambda v\|^2 =< \lambda v, \lambda v >= \lambda\overline{\lambda} < v, v >= |\lambda|^2\|v\|^2,$$

$$\|v+w\|^2 =< v+w, v+w >=< v, v > + < w, w > + < v, w > + < w, v >= \|v\|^2+\|w\|^2+2Re < v, w >$$

Dimostriamo ora la (iii). Se $v = 0$ oppure $w = 0$ l'enunciato vale. Supponiamo allora che $v \neq 0$ e $w \neq 0$. Allora, usando (ii), per ogni coppia di scalari a, b si ha

$$0 \leq \|av + bw\|^2 = |a|^2\|v\|^2 + |b|^2\|w\|^2 + 2Re(a\overline{b} < v, w >).$$

Se si pone $a = \|w\|^2 > 0$ e $b = - < v, w >$, la diseguaglianza diventa

$$\begin{aligned} 0 &\leq \|w\|^4\|v\|^2 + | < v, w > |^2\|w\|^2 + 2Re[\|w\|^2(-\overline{< v, w >} < v, w >)] \\ &= \|w\|^4\|v\|^2 - | < v, w > |^2\|w\|^2 . \end{aligned}$$

Dividendo per $\|w\|^2 > 0$, la diseguaglianza di Cauchy-Schwarz segue. Si osservi che l'eguaglianza vale se e solo se $\|av + bw\|^2 = 0$ e quindi se e solo se $av + bw = O$, ossia se v e w sono linearmente dipendenti. Per quanto riguarda (iv), si osservi prima di tutto che che per ogni v, w

$$|Re < v, w > | \leq | < v, w > | \leq \|v\|\|w\|.$$

Dunque

$$\begin{aligned} \|v + w\|^2 &= \|v\|^2 + \|w\|^2 + 2Re < v, w >\leq \|v\|^2 + \|w\|^2 + 2|Re < v, w > | \\ &\leq \|v\|^2 + \|w\|^2 + 2\|v\|\|w\| = (\|v\| + \|w\|)^2 \end{aligned}$$

e quindi (iv) vale. Infine dimostriamo le formule di polarizzazione nel caso complesso. Si ha

$$\|v + w\|^2 = \|v\|^2 + \|w\|^2 + 2Re < v, w > \quad \text{e} \quad \|v - w\|^2 = \|v\|^2 + \|w\|^2 - 2Re < v, w >$$

e quindi

$$4Re < v, w >= \|v + w\|^2 - \|v - w\|^2.$$

In modo analogo calcoliamo

$$\|v + iw\|^2 = \|v\|^2 + \|w\|^2 + 2Im < v, w > \quad \text{e} \quad \|v - iw\|^2 = \|v\|^2 + \|w\|^2 - 2Im < v, w >$$

e quindi

$$4Im < v, w >= \|v + iw\|^2 - \|v - iw\|^2.$$

Nel caso reale si procede allo stesso modo. □

Osservazione. Sia V uno spazio vettoriale metrico con norma $\|\cdot\|$. La funzione $d\colon V\times V \to \mathbb{R}^+$ definita da $d(v_1, v_2) = \|v_1 - v_2\|$ si dice *distanza*. Allora si vede immediatamente che per ogni $v_1, v_2 \in V$ si ha $d(v_1, v_2) \geq 0$, con l'uguaglianza se e solo se $v_1 = v_2$, e che $d(v_1, v_2) = d(v_2, v_1)$. Dalla diseguaglianza triangolare (1.2) segue la diseguaglianza triangolare per la distanza ossia $d(v_1, v_2) \leq d(v_1, v_3) + d(v_3, v_2)$ per ogni $v_1, v_2, v_3 \in V$. Infatti

$$d(v_1, v_2) = \|v_1 - v_2\| = \|v_1 - v_3 + v_3 - v_2\| \leq \|v_1 - v_3\| + \|v_3 - v_2\| = d(v_1, v_3) + d(v_3, v_2).$$

Grazie alla diseguaglianza di Cauchy-Schwarz per uno spazio vettoriale metrico reale ha senso definire una nozione di angolo.:

Definizione. Sia V uno spazio vettoriale metrico su $\mathbb{R}$ e siano $v, w \in V$ due vettori non nulli. L'*angolo* (convesso) fra v e w è il numero reale $\theta = \angle(vw) \in [0, \pi]$ tale che

$$\cos\theta = \frac{< v, w >}{\|v\|\|w\|}. \tag{1.5}$$

La definizione è ben posta dato che la diseguaglianza di Cauchy-Schwarz in questo caso equivale a

$$\frac{|< v, w >|}{\|v\|\|w\|} \leq 1$$

e quindi esiste un unico $\theta \in [0, \pi]$ tale che valga la (1.5).

Con argomenti di geometria elementare si vede facilmente che questa nozione di angolo nel caso di $\mathbb{R}^2$ o $\mathbb{R}^3$, pensati come spazi vettoriali metrici con i rispettivi prodotti scalari canonici, coincide con quella usuale. Nel caso di spazi vettoriali metrici complessi per dare una nozione di angolo occorre ricorrere a considerazioni che esulano dai contenuti del corso. Invece è semplice definire la nozione di ortogonalità per qualunque spazio vettoriale metrico.

Definizione. Sia V uno spazio vettoriale metrico con prodotto $< \cdot, \cdot >$. Due vettori $v, w \in V$ si dicono *ortogonali* (o *perpendicolari*) se $< v, w >= 0$ e si scrive $v \perp w$. Se $S_1, S_2 \subset V$ scriveremo $S_1 \perp S_2$ se ogni vettore di S_1 è ortogonale a ogni vettore in S_2. Se S_1 consiste di un solo elemento v si scrive semplicemente $v \perp S_2$.

Esercizio. Sia V uno spazio vettoriale metrico e $U \subset V$ un suo sottospazio. Sia $\{u_1, \ldots, u_r\}$ una base di V e sia $v \in V$ arbitrario. Dimostrare che $v \perp U$ se e solo se $< v, u_i >= 0$ per ogni $i = 1, \ldots, r$.

Proposizione 1.3: *Sia V uno spazio vettoriale metrico e $v_1, \ldots, v_n$ vettori non nulli a due a due ortogonali. Allora $v_1, \ldots, v_n$ sono linearmente indipendenti. In particolare se $n = dim(V)$, allora $\{v_1, \ldots, v_n\}$ è una base di V.*

Dimostrazione: Per ipotesi $< v_i, v_j >= 0$ se $i \neq j$ e $< v_i, v_i >\neq 0$ per ogni i dato che $v_i \neq O$. Dunque siano $\alpha_1, \ldots, \alpha_n$ scalari tali che

$$\alpha_1 v_1 + \ldots \alpha_n v_n = O.$$

Allora per ogni $i = 1, \ldots, n$ si ha

$$0 =< v_i, O >=< v_i, \alpha_1 v_1 + \ldots \alpha_n v_n >= \alpha_i < v_i, v_i >$$

da cui segue che necessariamente $\alpha_i = 0$ per ogni $i = 1, \ldots, n$. Dunque $v_1, \ldots, v_n$ sono linearmente indipendenti. Il resto dell'enunciato è immediato dato che se $dim(V) = n$ allora $\{v_1, \ldots, v_n\}$ è un insieme massimale di vettori linearmente indipendenti. $\square$

Definizione. Una base *ortogonale* di uno spazio vettoriale metrico V è una base costituita da vettori a due a due ortogonali. Una base ortogonale $\{v_1, \dots, v_n\}$ di V si dice *ortonormale* se è composta di vettori di norma 1, ossia se $\|v_i\| = 1$ per ogni $i = 1, \dots, n$.

Esempio 1. Si consideri $\mathbb{R}^n$ ($\mathbb{C}^n$) con il prodotto scalare (hermitiano) canonico. Allora la base canonica è una base ortonormale.

Esempio 2. Non è difficile scrivere tutte le basi ortogonali di $\mathbb{R}^2$ con il prodotto scalare canonico. Infatti $\mathcal{B} = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \right\}$ è una base ortonormale di $\mathbb{R}^2$ se e solo se

$$x_1^2 + x_2^2 = 1 = y_1^2 + y_2^2 \qquad \text{e} \qquad x_1 y_1 + x_2 y_2 = 0.$$

Dalla prima relazione segue che, per opportuni $\theta, \varphi \in \mathbb{R}$, si ha $x_1 = \cos\theta, x_2 = \sin\theta$, $y_1 = \cos\varphi, y_2 = \sin\varphi$. Dalla seconda si ottiene

$$0 = \cos\theta\cos\varphi + \sin\theta\sin\varphi = \cos(\theta - \varphi).$$

Pertanto si dovrà avere $\varphi = \theta - \frac{\pi}{2}$ oppure $\varphi = \theta - \frac{3}{2}\pi$. Dato che

$$\cos(\theta - \frac{3}{2}\pi) = -\sin\theta, \quad \sin(\theta - \frac{3}{2}\pi) = \cos\theta, \quad \cos(\theta - \frac{\pi}{2}) = \sin\theta, \quad \sin(\theta - \frac{\pi}{2}) = -\cos\theta,$$

allora $\mathcal{B}$ è una base ortonormale di $\mathbb{R}^2$ se e solo se esiste $\theta \in \mathbb{R}^2$ tale che

$$\mathcal{B} = \left\{ \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}, \begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix} \right\} \text{ oppure } \mathcal{B} = \left\{ \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}, \begin{pmatrix} \sin\theta \\ -\cos\theta \end{pmatrix} \right\}.$$

Le basi ortonormali sono strumenti fondamentali per studiare gli spazi vettoriali metrici perché possono essere utilizzate per esprimere in modo particolarmente semplice le proprietà metriche. Questo fatto è illustrato ad esempio nella seguente

Proposizione 1.4: *Sia $\{v_1, \dots, v_n\}$ una base ortonormale di V. Allora valgono le seguenti formule:*

(i) Per ogni $v \in V$ si ha

$$v = \sum_{i=1}^{n} < v, v_i > v_i \qquad \text{(somma di Fourier)} \tag{1.6}$$

(ii) Per ogni $v, w \in V$ si ha

$$< v, w >= \sum_{i=1}^{n} < v, v_i >< v_i, w > \qquad \text{(formula di Parseval)} \tag{1.7}$$

(iii) Per ogni $v \in V$ si ha:

$$\|v\|^2 = \sum_{i=1}^{n} | < v, v_i > |^2 \qquad \text{(teorema di Pitagora)} \tag{1.8}$$

Dimostrazione: Sia $v \in V$ arbitrario. Allora $v = \sum_{i=1}^{n} \alpha_i v_i$ per opportuni scalari $\alpha_1, \ldots, \alpha_n$. Si fissi un arbitrario $j \in \{1, \ldots, n\}$. Allora utilizzando il fatto che $\{v_1, \ldots, v_n\}$ è una base ortonormale:

$$< v, v_j >= \left\langle \sum_{i=1}^{n} \alpha_i v_i, v_j \right\rangle = \sum_{i=1}^{n} \alpha_i < v_i, v_j >= \alpha_j$$

e quindi (i) segue.

Siano ora $v, w \in V$ arbitrari. Allora utilizzando (i), abbiamo

$$< v, w > = \left\langle \sum_{i=1}^{n} < v, v_i > v_i, \sum_{j=1}^{n} < w, v_j > v_j \right\rangle = \sum_{i,j=1}^{n} < v, v_i > \overline{< w, v_j >} < v_i, v_j >$$
$$= \sum_{i=1}^{n} < v, v_i > \overline{< w, v_i >} = \sum_{i=1}^{n} < v, v_i >< v_i, w >$$

e quindi (ii) è dimostrata. Infine (iii) è (ii) nel caso in cui $w = v$. □

Naturalmente occorre porsi il problema dell'esistenza di basi ortonormali. A questione risponde il seguente risultato che non solo garantisce l'esistenza di basi ortonormali, ma fornisce un efficace metodo per costruirle.

Teorema 1.5: (Procedimento di ortonormalizzazione di Gram-Schmidt) Sia V uno spazio vettoriale metrico e siano $v_1, \ldots, v_r \in V$ vettori linearmente indipendenti. Allora i vettori $w_1, \ldots, w_n$ definiti per ricorrenza da

$$w_1 = v_1 \text{ e } w_j = v_j - \sum_{h=1}^{j-1} \frac{< v_j, w_h >}{< w_h, w_h >} w_h \quad \forall\ j = 2, \ldots r \tag{1.9}$$

soddisfano le seguenti proprietà:

(i) $Span(w_1, \ldots, w_j) = Span(v_1, \ldots, v_j)$ per ogni $j = 1, \ldots r$;

(ii) w_j è ortogonale a $Span(w_1, \ldots, w_{j-1})$ per ogni $j = 2, \ldots r$ e quindi,in particolare $w_1, \ldots, w_r$ sono a due a due ortogonali;

(iii) Se $\{v_1, \ldots, v_r\}$ è una base di V e $w_1, \ldots, w_n$ sono definiti da (1.9), allora $\{\frac{w_1}{\|w_1\|}, \ldots, \frac{w_r}{\|w_r\|}\}$ è una base ortonormale di V.

In particolare quindi uno spazio vettoriale metrico finitamente generato ammette una base ortonormale.

Dimostrazione: Si procede per induzione sull'indice j. Se $j = 1$ gli enunciati sono banali. Supponiamo che il risultato sia vero per $j-1$. Allora in particolare $Span(w_1, \ldots, w_{j-1}, v_j) = Span(v_1, \ldots, v_{j-1}, v_j)$ e, per costruzione $w_j \in Span(w_1, \ldots, w_{j-1}, v_j)$. Dato che per oni $k < j$ si ha

$$\begin{aligned} < w_j, w_k > &= \left\langle v_j - \sum_{h=1}^{j-1} \frac{< v_j, w_h >}{< w_h, w_h >} w_h, w_k \right\rangle =< v_j, w_k > - \sum_{h=1}^{j-1} \frac{< v_j, w_h >}{< w_h, w_h >} < w_h, w_k >, \\ &=< v_j, w_k > - \frac{< v_j, w_k >}{< w_k, w_k >} < w_k, w_k >= 0 \end{aligned} \tag{1.10}$$

allora, i vettori $w_1, \ldots, w_{j-1}, w_j$ sono a due a due a due ortogonali in $Span(v_1, \ldots, v_{j-1}, v_j)$ e quindi $Span(w_1, \ldots, w_j) = Span(v_1, \ldots, v_j)$. Da (1.10) segue anche che w_j è ortogonale a $Span(w_1, \ldots, w_{j-1})$, e quindi la dimostrazione è completa. □

Come immediata conseguenza abbiamo i seguenti

Corollario 1.6: *Sia V uno spazio vettoriale metrico di dimensione n e siano $v_1, \dots, v_r \in V$ vettori a due a due ortogonali ciascuno di norma 1. Allora esistono $n-r$ vettori $v_{r+1}, \dots, v_n$ tali che $\{v_1, \dots, v_r, v_{r+1}, \dots, v_n\}$ sia una base ortonormale di V.*

Dimostrazione: Basta completare $v_1, \dots, v_r \in V$ a una base $\{v_1, \dots, v_r, u_{r+1}, \dots, u_n\}$ di V e poi applicare il Teorema 1.5. □

Corollario 1.7: *Sia V uno spazio vettoriale metrico di dimensione n e sia $U \subset V$ un sottospazio di dimensione r di V. Allora esiste un unico sottospazio $U^\perp \subset V$, chiamato il complemento ortogonale di U, tale che $U \perp U^\perp$ e $V = U \oplus U^\perp$.*

Dimostrazione: Se $\{v_1, \dots, v_r\}$ è una base ortonormale di U, sia $\{v_1, \dots, v_r, v_{r+1}, \dots, v_n\}$ un suo completamento a una base ortonormale di V. Allora $U^\perp = Span(v_{r+1}, \dots, v_n)$. □

Molte applicazioni geometriche delle nozioni presentate in questo paragrafo sono dovute alla seguente

Proposizione 1.8: *Sia V uno spazio vettoriale metrico di dimensione finita e U un suo sottospazio. Esiste una applicazione lineare $P_U: V \to V$ chiamata* proiezione ortogonale *sul sottospazio U, che ha le seguenti proprietà:*

i) $ImP_U = U$ e per ogni $v \in V$ il vettore $P_U(v)$ è l'unico vettore tale che $v - P_U(v)$ è ortogonale a tutti i vettori di U;

ii) Se $\{v_1, \dots, v_r\}$ è una base ortonormale di U e $\{v_1, \dots, v_r, v_{r+1}, \dots, v_n\}$ è un suo completamento a una base ortonormale di V, allora per ogni vettore $v = \sum_{i=1}^n < v, v_i > v_i \in V$ si ha

$$P_U(v) = \sum_{i=1}^r < v, v_i > v_i; \tag{1.11}$$

iii) il vettore $P_U(v)$ è il vettore di U che minimizza la distanza da v, ossia per ogni vettore $u \in U$ vale la diseguaglianza

$$\|v - u\| \geq \|v - P_U(v)\| \tag{1.12}$$

e vale l'eguaglianza se e solo se $u = P_U(v)$.

Dimostrazione: È immediato convincersi che la formula (1.11) definisce una applicazione lineare $P_U: V \to V$ tale che $ImP_U = U$. Avremo dunque dimostrato completamente (i) e (ii) se proveremo che per ogni $v \in V$ il vettore $P_U(v)$ è l'unico vettore tale che $v - P_U(v)$ è ortogonale a tutti i vettori di U. Anche questo fatto è immediato. Infatti, per costruzione, $v - P_U(v) = \sum_{i=r+1}^n < v, v_i > v_i$ e quindi $< v - P_U(v), v_i >= 0$ per ogni $i = 1, \dots, r$. D'altra parte se $u \in U$ allora $u = \sum_{i=1}^r \alpha_i v_i$ per opportuni scalari $\alpha_1, \dots, \alpha_r$. Inoltre $v - u \perp U$ se e solo se per ogni $j = 1, \dots, r$ si ha

$$0 =< v - u, v_j >=< v, v_j > - \left\langle \sum_{i=1}^r \alpha_i v_i, v_j \right\rangle =< v, v_j > -\alpha_j$$

e quindi necessariamente se e solo se $u = P_U(v)$. Rimane da dimostrare (1.12). Sia $u \in U$ arbitrario, allora $v - u = (P_U(v) - u) + (v - P_U(v))$ e $(P_U(v) - u) \perp (v - P_U(v))$ dato che $v - P_U(v) \perp U$ e $P_U(v) - u \in U$. Pertanto

$$\begin{aligned}\|v - u\|^2 &= \|P_U(v) - u\|^2 + \|v - P_U(v)\|^2 + 2Re\,\langle P_U(v) - u, v - P_U(v)\rangle \\ &= \|P_U(v) - u\|^2 + \|v - P_U(v)\|^2 \geq \|v - P_U(v)\|^2\end{aligned}$$

con l'eguaglianza che vale se e solo se $\|P_U(v) - u\|^2 = O$ ossia se e solo se $u = P_U(v)$. □

Esempio. Per fissare le idee sulle nozioni presentate risolviamo il seguente esercizio. Sia U il sottospazio di $\mathbb{R}^4$ definito da

$$U = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} | x_1 + x_2 = x_3 + x_4 \right\}.$$

Troveremo una base ortonormale per U (pensato come spazio metrico con il prodotto scalare canonico) e l'estenderemo a una base ortonormale di $\mathbb{R}^4$. Infine troveremo la matrice della proiezione ortogonale P_U su U relativa alla base canonica di $\mathbb{R}^4$. Si vede facilmente che

$$\left\{ v_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, v_2 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, v_3 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} \right\}$$

è una base di U. Utilizzando il procedimento di Gram-Schmidt troviamo ora una base $\{w_1, w_2, w_3\}$ di U costituita di vettori a due a due ortogonali. Poniamo allora $w_1 = v_1$; utilizzando (1.9):

$$w_2 = v_2 - \frac{< v_2, w_1 >}{< w_1, w_1 >} w_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1/2 \\ 0 \\ -1/2 \\ 1 \end{pmatrix}$$

$$\begin{aligned} w_3 &= v_3 - \frac{< v_3, w_1 >}{< w_1, w_1 >} w_1 - \frac{< v_3, w_2 >}{< w_2, w_2 >} w_2 \\ &= \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} - \frac{< v_3, w_1 >}{< w_1, w_1 >} \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} - \frac{< v_3, w_2 >}{< w_2, w_2 >} \begin{pmatrix} 1/2 \\ 0 \\ -1/2 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} - \frac{-1/2}{3/2} \begin{pmatrix} 1/2 \\ 0 \\ -1/2 \\ 1 \end{pmatrix} = \begin{pmatrix} -1/3 \\ 1 \\ 1/3 \\ 1/3 \end{pmatrix}. \end{aligned}$$

Dunque una base ortonormale per U è data da

$$\mathcal{B}' = \left\{ \frac{w_1}{\|w_1\|}, \frac{w_2}{\|w_2\|}, \frac{w_3}{\|w_3\|} \right\} = \left\{ \frac{1}{\sqrt{2}} w_1, \frac{\sqrt{2}}{\sqrt{3}} w_2, \frac{\sqrt{3}}{2} w_3 \right\}.$$

Per completare $\mathcal{B}'$ a una base ortonormale di $\mathbb{R}^4$, procediamo nel modo seguente. Completiamo $\{w_1, w_2, w_3\}$ a una base di $\mathbb{R}^4$: basterà aggiungere un vettore della base canonica che non appartiene a U. Ad esempio un completamento è dato da $\{w_1, w_2, w_3, e_4\}$. Poi applichiamo il procedimento di Gram-Schmidt per sostituire e_4 con un vettore w_4 in modo che $\{w_1, w_2, w_3, w_4\}$ sia una base di vettori a due a due ortogonali. Dunque:

$$\begin{aligned} w_4 &= e_4 - \frac{< e_4, w_1 >}{< w_1, w_1 >} w_1 - \frac{< e_4, w_2 >}{< w_2, w_2 >} w_2 - \frac{< e_4, w_3 >}{< w_3, w_3 >} w_3 \\ &= \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} - 0 w_1 - \frac{2}{3} w_2 - \frac{1}{4} w_3 = \begin{pmatrix} -1/4 \\ -1/4 \\ 1/4 \\ 1/4 \end{pmatrix}. \end{aligned}$$

Una una base ortonormale $\mathcal{B}$ di $\mathbb{R}^4$ che completa $\mathcal{B}'$ è allora data da

$$\begin{aligned}\mathcal{B} &= \left\{ \frac{w_1}{\|w_1\|}, \frac{w_2}{\|w_2\|}, \frac{w_3}{\|w_3\|}, \frac{w_4}{\|w_4\|} \right\} = \left\{ \frac{1}{\sqrt{2}} w_1, \frac{\sqrt{2}}{\sqrt{3}} w_2, \frac{\sqrt{3}}{2} w_3, 2w_4 \right\} \\ &= \left\{ \begin{pmatrix} 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \\ 0 \end{pmatrix}, \begin{pmatrix} 1/\sqrt{6} \\ 0 \\ -1/\sqrt{6} \\ \sqrt{2}/\sqrt{3} \end{pmatrix}, \begin{pmatrix} -1/2\sqrt{3} \\ \sqrt{3}/2 \\ 1/2\sqrt{3} \\ 1/2\sqrt{3} \end{pmatrix}, \begin{pmatrix} -1/2 \\ -1/2 \\ 1/2 \\ 1/2 \end{pmatrix} \right\} \end{aligned} .$$

Un metodo alternativo per trovare un vettore che completi $\mathcal{B}'$ a una base ortonormale di $\mathbb{R}^4$ è il seguente. Dato che $\{v_1, v_2, v_3\}$ è una base di U si trova un vettore $z \neq O$ ortogonale a U risolvendo il sistema omogeneo $< v_1, z >=< v_2, z >=< v_3, z >= 0$. Aggiungendo il vettore $z/\|z\|$ si ottiene una base ortonormale di $\mathbb{R}^4$.

Cerchiamo ora di studiare la proiezione ortogonale $P_U \colon \mathbb{R}^4 \to \mathbb{R}^4$ sul sottospazio U. Dato che, per definizione si ha

$$P_U\left(\frac{w_1}{\|w_1\|}\right) = \frac{w_1}{\|w_1\|}, \ P_U\left(\frac{w_2}{\|w_2\|}\right) = \frac{w_2}{\|w_2\|}, \ P_U\left(\frac{w_3}{\|w_3\|}\right) = \frac{w_3}{\|w_3\|}, \ P_U\left(\frac{w_4}{\|w_4\|}\right) = O,$$

la matrice associata a P_U relativa alla base $\mathcal{B}$ è data da

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

Per trovare la matrice A' di P_U relativa alla base canonica $\mathcal{C}$ di $\mathbb{R}^4$, si ricordi che $A' = B^{-1}AB$ dove B è la matrice del cambio di base da $\mathcal{B}$ a $\mathcal{C}$ e B^{-1} è la matrice del cambio di base da $\mathcal{C}$ a $\mathcal{B}$. Dunque B^{-1} è la matrice che ha per colonne le coordinate dei vettori della base $\mathcal{B}$ relative alla base canonica $\mathcal{C}$ e B è l'inversa di B^{-1}. Si ha allora

$$B^{-1} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & -\frac{1}{2\sqrt{3}} & -\frac{1}{2} \\ 0 & 0 & \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{2\sqrt{3}} & \frac{1}{2} \\ 0 & \frac{\sqrt{2}}{\sqrt{3}} & \frac{1}{2\sqrt{3}} & \frac{1}{2} \end{pmatrix} \ \text{e} \ B = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{6}} & 0 & -\frac{1}{\sqrt{6}} & \frac{\sqrt{3}}{2} \\ -\frac{1}{2\sqrt{3}} & \frac{\sqrt{3}}{2} & \frac{1}{2\sqrt{3}} & \frac{1}{2\sqrt{3}} \\ -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

Si osservi che B è la trasposta di B^{-1}: questo e un fatto non casuale che si chiarirà fra qualche pagina. Possiamo allora calcolare

$$
\begin{aligned}
A' &= B^{-1}AB \\
&= \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & -\frac{1}{2\sqrt{3}} & -\frac{1}{2} \\ 0 & 0 & \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{2\sqrt{3}} & \frac{1}{2} \\ 0 & \frac{\sqrt{2}}{\sqrt{3}} & \frac{1}{2\sqrt{3}} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{6}} & 0 & -\frac{1}{\sqrt{6}} & \frac{\sqrt{3}}{2} \\ -\frac{1}{2\sqrt{3}} & \frac{\sqrt{3}}{2} & \frac{1}{2\sqrt{3}} & \frac{1}{2\sqrt{3}} \\ -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{pmatrix} \\
&= \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & -\frac{1}{2\sqrt{3}} & 0 \\ 0 & 0 & \frac{\sqrt{3}}{2} & 0 \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{2\sqrt{3}} & 0 \\ 0 & \frac{\sqrt{2}}{\sqrt{3}} & \frac{1}{2\sqrt{3}} & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{6}} & 0 & -\frac{1}{\sqrt{6}} & \frac{\sqrt{3}}{2} \\ -\frac{1}{2\sqrt{3}} & \frac{\sqrt{3}}{2} & \frac{1}{2\sqrt{3}} & \frac{1}{2\sqrt{3}} \\ -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{pmatrix} \\
&= \begin{pmatrix} 3/4 & -1/4 & 1/4 & 1/4 \\ -1/4 & 3/4 & 1/4 & 1/4 \\ 1/4 & 1/4 & 3/4 & -1/4 \\ 1/4 & 1/4 & -1/4 & 3/4 \end{pmatrix}
\end{aligned}
$$

Da questo punto in poi consideriamo escusivamente spazi vettoriali metrici di dimensione finita. Ci sarà utile nel seguito il seguente

Lemma 1.9: *Sia V uno spazio vettoriale metrico e $w \in V$. Allora $< v, w >= 0$ per ogni $v \in V \iff w = O$.*

Dimostrazione: Se $w = O$ allora ovviamente $< v, w >= 0$ per ogni $v \in V$. Viceversa sia $< v, w >= 0$ per ogni $v \in V$. Se $\{v_1, \ldots, v_n\}$ è una base ortonormale di V allora

$$w = \sum_{j=1}^{n} < w, v_j > v_j = O.$$

□

Proposizione 1.10: *Sia $T\colon V \to W$ un'applicazione lineare fra spazi metrici su $\mathbb{K} = \mathbb{R}$ o $\mathbb{C}$. Esiste una unica applicazione lineare $T^*\colon V \to W$ tale che per ogni $v \in V$ e $w \in W$ si abbia:*

$$< T(v), w >_W = < v, T^*(w) >_V . \tag{1.13}$$

Se A è la matrice associata a T relativa a basi ortonormali $\mathcal{B}$ di V e $\mathcal{C}$ di W, allora la matrice associata a T^* relativa a $\mathcal{C}$ e $\mathcal{B}$ è

$$A^* = \begin{cases} A^T & \text{se } \mathbb{K} = \mathbb{R} \\ A^H = \overline{A}^T & \text{se } \mathbb{K} = \mathbb{C} \end{cases} . \tag{1.14}$$

Dimostrazione: Esistenza di T^*: Siano $\mathcal{B} = \{v_1, \ldots, v_n\}$ e $\mathcal{C} = \{w_1, \ldots, w_m\}$ rispettivamente di V e W. Se A è la matrice associata a T relativa a $\mathcal{B}$ e $\mathcal{C}$, allora l'applicazione $T^*: W \to V$ che relativamente a $\mathcal{C}$ e $\mathcal{B}$ ha matrice $A^* = (a^*_{ij})$ definita da (1.14) effettivamente verifica (1.13). Siano

$$v = \sum_{i=1}^{n} x_i v_i \in V \quad w = \sum_{j=1}^{m} y_j w_j \in W.$$

Allora per $T(v)$ è l'unico vettore di W che ha coordinate Ax rispetto alla base $\mathcal{C}$ e quindi

$$T(v) = \sum_{l=1}^{m} \left(\sum_{i=1}^{n} a_{li} x_i \right) w_l$$

e

$$T^*(w) = \sum_{k=1}^{n} \left(\sum_{j=1}^{m} a^*_{kj} y_j \right) v_k = \sum_{k=1}^{n} \left(\sum_{j=1}^{n} \overline{a}_{jk} y_j \right) v_k.$$

Pertanto

$$\begin{aligned} < T(v), w >_W &= \left\langle \sum_{l=1}^{m} \left(\sum_{i=1}^{n} a_{li} x_i \right) w_l, \sum_{j=1}^{m} y_j w_j \right\rangle_W \\ &= \sum_{l=1}^{m} \sum_{j=1}^{m} \left(\sum_{i=1}^{n} a_{li} x_i \right) \overline{y}_j < w_l, w_j >_W \\ &= \sum_{j=1}^{m} \sum_{i=1}^{n} a_{ji} x_i y_j \end{aligned}$$

e analogamente

$$\begin{aligned} < v, T^*(w) >_V &= \left\langle \sum_{i=1}^{n} x_i v_i, \sum_{k=1}^{n} \left(\sum_{j=1}^{m} \overline{a}_{jk} y_j \right) v_k \right\rangle_V \\ &= \sum_{i=1}^{n} \sum_{k=1}^{n} x_i \left(\sum_{j=1}^{m} a_{jk} \overline{y}_j \right) < v_i, v_k >_V \\ &= \sum_{i=1}^{n} \sum_{j=1}^{m} a_{ji} x_i y_j \end{aligned}$$

e quindi la (1.13) è verificata.

Unicità di T^*: Se $S: V \to W$ è un'altra applicazione lineare che soddisfa (1.13), allora per ogni $v \in V$ e $w \in W$

$$< v, T^*(w) >_V = < T(v), w >_W = < v, S(w) >_V$$

e quindi $< v, T^*(w) - S(w) >_V = 0$. Per il Lemma 1.9 allora $T^*(w) - S(w) = O$ per ogni $w \in W$ e dunque $T^* = S$. $\square$

Definizione. Sia $T\colon V \to W$ un'applicazione lineare fra spazi metrici. L'applicazione $T^*\colon W \to V$ definita nella Proposizione 1.10 si dice *aggiunta* di T.

Osservazione 1. Sia $T\colon V \to W$ un'applicazione lineare fra spazi metrici e sia $T^*\colon W \to V$ la sua aggiunta. Per l'unicità dell'aggiunta e per l'arbitrarietà della scelta delle basi ortonormali usate nella Proposizione 1.10 per definire T^*, segue che per *qualunque* scelta di basi ortonormali $\mathcal{B}$ per V e $\mathcal{C}$ per W, se A è la matrice di T relativa a $\mathcal{B}$ e $\mathcal{C}$, allora la matrice A^* di T^* relativa a $\mathcal{C}$ e $\mathcal{B}$ è data dalla (1.14).

Esempio. Si consideri su $\mathbb{R}^n$ il prodotto scalare canonico. Se $A \in M_{m,n}(\mathbb{R})$, L'aggiunta di $L_A\colon \mathbb{R}^n \to \mathbb{R}^m$ è $L_{A^T}\colon \mathbb{R}^m \to \mathbb{R}^n$. Basta considerare su $\mathbb{R}^n$ e $\mathbb{R}^m$ le basi canoniche e applicare applicare la Proposizione 1.10. Ovviamente in questo caso si può anche osservare direttamente che $L_A^* = L_{A^T}$ dato che per ogni $v \in \mathbb{R}^n$ e $w \in \mathbb{R}^m$

$$< L_A(v), w >_{R^m} = w^T A v = (w^T A v)^T = v^T A^T w = < L_{A^T}(w), v >_{R^n} = < v, L_{A^T}(w) >_{R^n}$$

Esempio. Si consideri su $\mathbb{C}^n$ il prodotto scalare canonico. Se $A \in M_{m,n}(\mathbb{C})$, L'aggiunta di $L_A\colon \mathbb{C}^n \to \mathbb{C}^m$ è $L_{A^H}\colon \mathbb{C}^m \to \mathbb{C}^n$. Basta considerare su $\mathbb{C}^n$ e $\mathbb{C}^m$ le basi canoniche e applicare applicare la Proposizione 1.10. Anche in questo caso è facile dimostrare direttamente che $L_A^* = L_{A^H}$ dato che per ogni $v \in \mathbb{C}^n$ e $w \in \mathbb{C}^m$

$$< L_A(v), w >_{C^m} = w^H A v = \overline{(w^H A v)^H} = \overline{v^H A^H w} = \overline{< L_{A^H}(w), v >_{C^n}} = < v, L_{A^H}(w) >_{C^n}$$

Definizione. Un endomorfismo $T\colon V \to W$ di uno spazio metrico si dice *autoaggiunto* se $T = T^*$ ossia se $< T(v), w > = < v, T(w) >$ per ogni $v, w \in V$.

Proposizione 1.11: *Sia $T\colon V \to V$ un endomorfismo di uno spazio metrico su $\mathbb{K} = \mathbb{R}$ o $\mathbb{C}$. Siano $\mathcal{B}$ una base ortonormale V e A sia la matrice di T relativa a $\mathcal{B}$. Allora T è autoaggiunto $\iff$*

$$A = A^* = \begin{cases} A^T & \text{se } \mathbb{K} = \mathbb{R} \\ \\ A^H = \overline{A}^H & \text{se } \mathbb{K} = \mathbb{C} \end{cases}.$$

Dimostrazione: Immediata dall'Osservazione 1. $\square$

Definizione. Un endomorfismo $T\colon V \to V$ di uno spazio metrico si dice *isometria* se $< T(v), T(w) > = < v, w >$ per ogni $v, w \in V$. A volte le isometrie sono chiamate *rotazioni*.

Proposizione 1.12: *Le seguenti affermazioni sono equivalenti:*

(i) T è un'isometria;

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

(iii) T trasforma basi ortonormali in basi ortonormali;

(iv) $\|T(v)\| = \|v\|$ per ogni $v \in V$;

(v) Se A è la matrice di T rispetto a una base ortonormale, allora A è invertibile e

$$A^{-1} = A^* = \begin{cases} A^T & \text{se } \mathbb{K} = \mathbb{R} \\ \\ A^H = \overline{A}^H & \text{se } \mathbb{K} = \mathbb{C} \end{cases}.$$

Dimostrazione: La dimostrazione seguirà il seguente schema:

$$(i) \Longrightarrow (ii) \Longrightarrow (iii) \Longrightarrow (iv) \Longrightarrow (i) \text{ e } (ii) \Longleftrightarrow (v).$$

$(i) \Longrightarrow (ii)$ Per ogni $v, w \in V$ si ha $< v, w >=< T(v), T(w) >=< v, T^*(T(w)) >$ e quindi

$$< v, w - T^*(T(w)) >= 0.$$

Per il lemma 1.9 allora

$$T^*(T(w)) = w \tag{1.15}$$

per ogni $w \in V$. Da (1.15) segue immediatamente che T è iniettiva. Infatti se per $w_1, w_2 \in V$ si ha $T(w_1) = T(w_2)$ allora $w_1 = T^*(T(w_1)) = T^*(T(w_2)) = w_2$. Dato che è un endomorfismo, T è biettivo e quindi invertibile. Per l'unicità dell'inverso, ancora da (1.15) segue $T^{-1} = T^*$.

$(ii) \Longrightarrow (iii)$ Sia $\{v_1, \ldots, v_n\}$ una base ortonormale. Si ha

$$< T(v_i), T(v_j) >=< v_i, T^*(T(v_j)) >=< v_i, T^{-1}(T(v_j)) >=< v_i, v_j >= \begin{cases} 0 \text{ se } i \neq j \\ 1 \text{ se } i = j \end{cases}$$

e quindi $\{T(v_1), \ldots, T(v_n)\}$ è una base ortonormale.

$(iii) \Longrightarrow (iv)$ Per ipotesi, se $\{v_1, \ldots, v_n\}$ è una base ortonormale, anche $\{T(v_1), \ldots, T(v_n)\}$ lo è. Sia $v \in V$. Allora

$$v = \sum_{i=1}^{n} < v, v_i > v_i \quad \text{e} \quad T(v) = T\left(\sum_{i=1}^{n} < v, v_i > v_i\right) = \sum_{i=1}^{n} < v, v_i > T(v_i).$$

Allora

$$\begin{aligned} \|T(v)\|^2 &= \left\langle \sum_{i=1}^{n} < v, v_i > T(v_i), \sum_{k=1}^{n} < v, v_k > T(v_k) \right\rangle \\ &= \sum_{i=1}^{n} \sum_{k=1}^{n} < v, v_i > \overline{< v, v_k >} < T(v_i), T(v_k) >= \sum_{i=1}^{n} | < v, v_i > |^2 = \|v\|^2. \end{aligned}$$

$(iv) \Longrightarrow (i)$ Utilizzando la formuala di polarizzazione, nel caso reale abbiamo:

$$\begin{aligned} < T(v), T(w) > &= \frac{1}{4}(\|T(v) + T(w)\|^2 - \|T(v) - T(w)\|^2 = \\ &\frac{1}{4}(\|T(v + w)\|^2 - \|T(v - w)\|^2 = \frac{1}{4}(\|v + w\|^2 - \|v - w\|^2 =< v, w > . \end{aligned}$$

Il caso complesso si dimostra allo stesso modo utilizzando le corrispondenti formule per il prodotto hermitiano.

$(ii) \Longleftrightarrow (v)$ Fissata una base ortonormale $\mathcal{B}$ di V, se A è la matrice di T relativa a $\mathcal{B}$, allora T è invertibile se e solo se A è invertibile e $T^{-1} = T^*$ se e solo se $A^{-1} = A^*$. □

Definizione. Una matrice $A \in M_{n,n}(\mathbb{R})$ si dice *ortogonale* se $A^T A = I_n$. Una matrice $A \in M_{n,n}(\mathbb{C})$ si dice *unitaria* se $A^H A = I_n$.

Corollario 1.13: *Sia A una matrice quadrata di ordine n reale (complessa). Le seguenti affermazioni sono equivalenti:*

(i) A è ortogonale (unitaria);

(ii) L_A è una isometria di $\mathbb{R}^n$ $(\mathbb{C}^n)$ con il prodotto scalare (hermitiano) canonico;

(iii) Le colonne di A formano una base ortonormale di $\mathbb{R}^n$ $(\mathbb{C}^n)$ con il prodotto scalare (hermitiano) canonico.

Dimostrazione: $(i) \iff (ii)$ La matrice A è ortogonale (unitaria) se e solo se $A^T A = I_n$ $(A^H A = I_n)$ e quindi se e solo se

$$(L_A)^{-1} = L_{A^{-1}} = L_{A^T} = (L_A)^* \text{ se } \mathbb{K} = \mathbb{R},$$

$$(L_A)^{-1} = L_{A^{-1}} = L_{A^H} = (L_A)^* \text{ se } \mathbb{K} = \mathbb{C}.$$

$(i) \iff (iii)$ Se $A = (A^1, \ldots, A^n)$, allora A è ortogonale (unitaria) se e solo se per ogni $i, j = 1, \ldots, n$ si ha

$$(A^i)^T A^j = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases} \text{ se } \mathbb{K} = \mathbb{R} \qquad (A^i)^H A^j = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases} \text{ se } \mathbb{K} = \mathbb{C}$$

ossia se e solo se $\{A^1, \ldots, A^n\}$ è una base ortonormale di $\mathbb{R}^n$ $(\mathbb{C}^n)$ con il prodotto scalare (hermitiano) canonico. $\square$

Esercizio. La composizione di isometrie è una isometria. L'inversa di una isometria è un'isometria.

Esercizio. Il prodotto di due matrici ortogonali (unitarie) è una matrice ortogonale (unitaria). L'inversa di una matrice ortogonale (unitaria) è una matrice ortogonale (unitaria).

Esercizio. Dimostrare che una matrice A di ordine 2 è ortogonale se e solo se esiste $\theta \in \mathbb{R}$ tale che si può scrivere in uno dei modi seguenti:

$$A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \text{ oppure } A = \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}.$$

Consiglio: ricordare la descrizione data delle basi ortonormali di $\mathbb{R}^2$.

Utilizzeremo le seguenti notazioni: $O(n) = \{A \in M_{n,n}(\mathbb{R}) | A^T A = I_n\}$ per il gruppo delle matrici ortogonali di ordine n e $U(n) = \{A \in M_{n,n}(\mathbb{R}) | A^H A = I_n\}$ per il gruppo delle matrici unitarie di ordine n.

Osservazione. Siano $\mathcal{B}$, $\mathcal{C}$ due basi ortonormali di uno spazio vettoriale metrico V su $\mathbb{K} = \mathbb{R}$ o $\mathbb{C}$. La matrice B del cambio di base da $\mathcal{B}$ a $\mathcal{C}$ è ortogonale se $\mathbb{K} = \mathbb{R}$, unitaria se $\mathbb{K} = \mathbb{C}$. Questo è immediato se si osserva che B è la matrice dell'applicazione identica di V relativa alle basi $\mathcal{B}$, $\mathcal{C}$. Dato che l'identità è sempre un'isometria (esercizio banale!) allora necessariamente B è ortogonale se $\mathbb{K} = \mathbb{R}$, unitaria se $\mathbb{K} = \mathbb{C}$.

Osservazione. Siano $\mathcal{B}$, $\mathcal{C}$ due basi ortonormali di uno spazio vettoriale metrico V su $\mathbb{K} = \mathbb{R}$ o $\mathbb{C}$. Sia $T: V \to V$ un endomorfismo e siano A la matrice di T relativa a $\mathcal{B}$ e A' la matrici di T relativa $\mathcal{C}$. Allora

$$A' = B^{-1}AB = \begin{cases} B^TAB & \text{se } \mathbb{K} = \mathbb{R} \\ B^HAB & \text{se } \mathbb{K} = \mathbb{C}. \end{cases}$$

Due matrici A, A' reali quadrate di ordine n si dicono *congruenti* se esiste una matrice invertibile B tale che $A' = B^TAB$.

2. Diagonalizzazione ortogonale di endomorfismi autoaggiunti

In questo paragrafo dimostreremo che ogni endomorfismo autoaggiunto di uno spazio vettoriale metrico è diagonalizzabile e che la base rispetto alla quale ha matrice diagonale si può scegliere ortonormale. Cominciamo con la seguente cruciale osservazione:

Teorema 2.1: *Sia $T: V \to V$ un endomorfismo autoaggiunto di uno spazio vettoriale metrico V di dimensione n su $\mathbb{K} = \mathbb{R}$ o $\mathbb{C}$. Allora T ha n autovalori (contati con la molteplicità) e sono tutti in $\mathbb{R}$.*

Dimostrazione: Rispetto a una base ortonormale T ha una matrice A simmetrica nel caso reale, hermitiana nel caso complesso. Gli autovalori di T sono esattamente le radici del suo polinomio caratteristico di A che sono nel campo $\mathbb{K}$. Per il teorema fondamentale dell'algebra, il polinomio caratteristico di A ha esattamente n radici (contate con la molteplicità) in $\mathbb{C}$. Dobbiamo allora solo dimostrare che le radici del polinomio caratteristico di A sono reali. Si consideri $L = L_A: \mathbb{C}^n \to \mathbb{C}^n$. Dato che $A = A^H$ (sia nel caso reale, sia nel caso complesso), allora L è un endomorfismo autoaggiunto di $\mathbb{C}^n$ con il prodotto hermitiano canonico. Sia λ una radice del polinomio caratteristico di A. Allora λ è un autovalore di L. Sia $u \in \mathbb{C}^n \setminus \{0\}$ tale che $L(u) = \lambda u$. Si ha

$$\lambda \|u\|^2 = \lambda < u, u >=< \lambda u, u >=< Lu, u >=< u, Lu >= \overline{\lambda} < u, u >= \overline{\lambda}\|u\|^2$$

e quindi $(\lambda - \overline{\lambda})\|u\|^2 = 0$. Dato che $\|u\|^2 \neq 0$, allora deve essere $\lambda - \overline{\lambda} = 0$ ossia $\lambda \in \mathbb{R}$. $\square$

Il risultato principale che vogliamo dimostrare è il seguente

Teorema 2.2: *Sia V uno spazio vettoriale metrico di dimensione n e sia $T: V \to V$. Le seguenti affermazioni sono equivalenti:*

(i) Esiste una base ortonormale $\{v_1, \ldots, v_n\}$ di V e numeri reali $\lambda_1, \ldots, \lambda_n$ tali che se $u = \sum_{i=1}^n x_i v_i$ allora $T(u) = \sum_{i=1}^n \lambda_i x_i v_i$.

(ii) V ha una base ortonormale formata da autovettori di T ciascuno con autovalore reale (rispetto alla quale T ha matrice diagonale).

(iii) T è autoaggiunta.

Dimostrazione: $(i) \Longleftrightarrow (ii)$ Se vale (i) allora si ha $T(v_i) = \lambda_i v_i$ e quindi vale (ii). D'altro canto se $\{v_1, \ldots, v_n\}$ è una base ortonormale formata da autovettori di T ciascuno con autovalore reale λ_i corrispondente a v_i allora la (i) segue immediatamente per la linearità di T.

$(ii) \Longleftrightarrow (iii)$ Se vale (ii) allora T ha matrice diagonale rispetto a una base ortonormale. Dato che una matrice diagonale reale coincide con la sua trasposta e con la sua aggiunta, necessariamente $T = T^*$. Supponiamo viceversa che valga (iii). Procederemo per induzione sulla dimensione n di V. Se $n = 1$ allora (ii) è banale. Supponiamo che se T è autoaggiunta valga (ii) per spazi di dimensione $n - 1$. Per il Teorema 2.1, esiste un autovalore $\lambda \in \mathbb{R}$ di T. Sia v_1 un autovettore con autovalore λ e con $\|v_1\| = 1$. Sia $W = (Span(v_1))^\perp$. Allora W è un sottospazio di V di dimensione $n-1$ e, per costruzione, si ha che $w \in W$ se e solo se $< w, v_1 >= 0$ e quindi $W = \{w \in V| < w, v_1 >= 0\}$.

Si ha inoltre $T(W) \subset W$. Infatti $u \in T(W) \Leftrightarrow u = T(w)$ per qualche $w \in W$. Allora

$$< u, v_1 >=< T(w), v_1 >=< w, T(v_1) >=< w, \lambda v_1 >= \overline{\lambda} < w, v_1 >= 0$$

e quindi $u \in W$. Il prodotto definito su V, ristretto a W, definisce una struttura di spazio vettoriale metrico per W. Dato che $T(W) \subset W$ allora $T: W \to W$ è un endomorfismo autaggiunto di uno spazio vettoriale metrico di dimensione $n-1$. Esiste allora, per l'ipotesi induttiva, una base ortonormale $\{v_2, \ldots, v_n\}$ formata da autovettori di di T ciascuno con autovalore reale. Dato che $< v_j, v_1 >= 0$ per $j = 2, \ldots, n$, allora $\{v_1, v_2, \ldots, v_n\}$ è una base ortonormale di V formata da autovettori di di T ciascuno con autovalore reale e quindi (ii) vale. □

Conseguenza immediata del Teorema 2.2 è il seguente

Corollario 2.3: (Teorema spettrale) *Sia $T: V \to V$ un endomorfismo di uno spazio vettoriale metrico reale. Esiste una base ortonormale formata da autovettori di T se e solo se T è autoaggiunto.*

In termini di matrici il Teorema 2.2 si traduce in questo modo:

Teorema 2.4: *(i) Sia $A \in M_{n,n}(\mathbb{R})$ una matrice simmetrica reale. Allora esiste una matrice ortogonale $B \in \mathcal{O}(n)$ tale che $D = B^{-1}AB = B^TAB$ è una matrice diagonale.*

(ii) Sia $A \in M_{n,n}(\mathbb{C})$ una matrice hermitiana. Allora esiste una matrice unitaria $B \in \mathcal{U}(n)$ tale che $D = B^{-1}AB = B^HAB$ è una matrice diagonale reale.

Dimostrazione: (i) Si consideri $L_A: \mathbb{R}^n \to \mathbb{R}^n$. Allora L_A è un endomorfismo autoaggiunto di $\mathbb{R}^n$ con il prodotto scalare canonico. Per il Teorema 2.2 esiste una base ortonormale $\mathcal{B}$ rispetto alla quale L_A ha matrice diagonale. La matrice B del cambio di base dalla base canonica alla base $\mathcal{B}$, che ha per colonne le coordinate dei vettori di $\mathcal{B}$, è la matrice cercata.

(ii) La dimostrazione, analoga al caso reale, è lasciata per esercizio. □

Vogliamo illustrare ora come in pratica si possa diagonalizzare un endomorfismo autoaggiunto. A tal fine cominciamo con la seguente utile

Proposizione 2.5: *Siano V uno spazio vettoriale metrico e $T: V \to V$ un endomorfismo autoaggiunto. Autovettori di T relativi a autovalori distinti sono ortogonali.*

Dimostrazione: Siano $\mu \neq \nu$ autovalori di T e v, w autovettori relativi rispettivamente a μ e ν. Allora

$$\mu < v, w >=< \mu v, w >=< T(v), w >=< v, T(w) >=< v, \nu w >= \nu < v, w >$$

e quindi, dato che $\mu \neq \nu$ deve essere $< v, w >= 0$. □

La Proposizione 2.5 ci dice che uno spazio vettoriale metrico si decompone in autospazi di un endomorfismo autoaggiunto. Suggerisce quindi di procedere per diagonalizzare un endomorfismo autaggiunto $T\colon V \to V$ di uno spazio metrico V seguendo il seguente schema:

I Passo: Fissata una base ortonormale, si scrive la matrice associata a T, si calcola il polinomio caratteristico sfruttando la matrice e si trovano gli autovalori di T

II Passo: Per ciascun autovalore si trova una base ortonormale per il corrispondente autospazio. L'unione delle basi degli autospazi determinate in questo modo è una base diagonalizzante per T.

Data una matrice simmetrica reale A si può adattare il procedimento per trovare una matrice ortogonale B tale che $D = B^T AB = B^{-1}AB$ sia diagonale:

I Passo: Si scrive il polinomio caratteristico e si calcolano gli autovalori di A.

II Passo: Per ciascun autovalore si trova una base ortonormale per il corrispondente autospazio. L'unione delle basi degli autospazi determinate in questo modo è una base diagonalizzante per L_A. La matrice B che ha per colonne i vettori della base ottenuta in questo modo è la matrice cercata.

In tutti e due i casi nel I Passo occorre trovare le radici di un polinomio. Mentre sotto le ipotesi imposte si è sicuri che esistono il numero "giusto" di radici reali, trovarle effettivamente è tutta un'altra faccenda! Per diagonalizzare mediante matrici unitarie matrici hermitiane si procede in modo analogo.

Esempio Si consideri la matrice $A = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$. Il suo polinomio caratteristico è $p_A(\lambda) = (\lambda^2 - 1)^2$. Dunque gli autovalori di A sono $1, -1$ ciascuno con moteplicità 2. Se V_1, V_{-1} sono i rispettivi autospazi si ha

$$V_1 = Ker(A - I_4) = Ker \begin{pmatrix} -1 & 0 & 0 & 1 \\ 0 & -1 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 1 & 0 & 0 & -1 \end{pmatrix} = Span \left(\begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} \right)$$

$$V_{-1} = Ker(A + I_4) = Ker \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix} = Span \left(\begin{pmatrix} 1 \\ 0 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix} \right)$$

Dunque basi ortonormali per V_1 e V_{-1} sono date ad esempio rispettivamente da:

$$\left\{ \begin{pmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ 0 \\ \frac{1}{\sqrt{2}} \end{pmatrix}, \begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix} \right\} \quad \left\{ \begin{pmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ 0 \\ -\frac{1}{\sqrt{2}} \end{pmatrix}, \begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \end{pmatrix} \right\}$$

e la matrice ortonormale diagonalizzante B è data da

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

Concludiamo il paragrafo con un'osservazione che è alle volte utile. Abbiamo ripetutamente osservato che un endomorfismo arbitrario di uno spazio vettoriale non è in genere diagonalizzabile. D'altra parte si dimostra facilmente che dato un endomorfismo di uno spazio vettoriale metrico esiste una base ortonormale rispetto alla quale l'endomorfismo ha matrice triangolare superiore: dunque se non si può sempre diagonalizzare un endomorfismo, almeno lo si può triangolarizzare! Diamo di seguito la dimostrazione di questo risultato nella versione per matrici:

Teorema 2.6: *Per ogni $A \in M_n(\mathbb{C})$ esiste una matrice unitaria U tale che $S = U^{-1}AU = U^H AU$ è una matrice triangolare superiore. Analogamente, per ogni $A \in M_n(\mathbb{R})$ che ha n autovalori, contati con la loro molteplicità lagebrica, reali. Allora esiste una matrice ortogonale O tale che $S = O^{-1}AO = O^T AU$ è una matrice triangolare superiore.*

Dimostrazione:

Dimostriamo il caso complesso procedendo per induzione. Il caso reale è del tutto analogo. Per $n = 1$ il risultato è ovvio. Supponiamo sia vero per matrici quadrate di ordine $n-1$. Sia λ_1 un autovalore dell'endomorfismo $L_A\colon \mathbb{C}^n \to \mathbb{C}^n$ e sia z_1 un autovettore relativo a λ_1 con $\|z_1\| = 1$. Sia infine $\{z_1, z_2, \ldots, z_n\}$ una base ortonormale di $\mathbb{C}^n$. La matrice B di L_A relativa a questa base ha questa forma:

$$B = \begin{pmatrix} \lambda_1 & * & \ldots & * \\ 0 & & & \\ \vdots & & A_1 & \\ 0 & & & \end{pmatrix}$$

dove A_1 è una matrice quadrata di ordine $n-1$. Dato che A e B sono matrici dello stesso endomorfismo L_A relative a basi ortonormali diverse, esiste una matrice unitaria U_0 di ordine n tale che $B = U_0^{-1}AU_0 = U_0^H AU_0$.

Per l'ipotesi induttiva esiste una matrice unitaria U_1' di ordine $n-1$ tale che $S' = (U_1')^{-1}AU_1' = U_1'^H AU_1'$ è triangolare superiore. Se definiamo

$$U_1 = \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & & & \\ \vdots & & U_1' & \\ 0 & & & \end{pmatrix},$$

allora U_1 è una matrice unitaria di ordine n tale che

$$\begin{aligned} S =& U_1^{-1}BU_1 = U_1^H BU_1 \\ =& \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & & & \\ \vdots & & (U_1')^H & \\ 0 & & & \end{pmatrix} \begin{pmatrix} \lambda_1 & * & \ldots & * \\ 0 & & & \\ \vdots & & A_1 & \\ 0 & & & \end{pmatrix} \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & & & \\ \vdots & & U_1' & \\ 0 & & & \end{pmatrix} \\ =& \begin{pmatrix} \lambda_1 & * & \ldots & * \\ 0 & & & \\ \vdots & & (U_1')^H A_1 U_1' & \\ 0 & & & \end{pmatrix} = \begin{pmatrix} \lambda_1 & * & \ldots & * \\ 0 & & & \\ \vdots & & S' & \\ 0 & & & \end{pmatrix} \end{aligned}$$

è triangolare superiore. Ma allora la matrice unitaria $U = U_0U_1$ è la matrice cercata dato che, rimettendo insieme i calcoli fatti, si ha $S = U^{-1}AU = U^H AU$. □

3. Forme quadratiche reali

Definizione 3.1 Sia $A \in M_{n,n}$ una matrice simmetrica. La forma *forma quadratica* reale associata ad $A = (a_{ij})$ è la funzione $\phi: \mathbb{R}^n \to \mathbb{R}$ definita da

$$\phi(x) = x^T A x = \sum_{i,j=1}^{n} a_{ij} x_i x_j. \tag{3.1}$$

Esempio Sia $f: \mathbb{R}^n \to \mathbb{R}$ una funzione infinitamente derivabile. Per $x_0 \in \mathbb{R}^n$ sia

$$A = \frac{\partial^2 f}{\partial x_i \partial x_j}(x_0)$$

la matrice delle derivate seconde di F in x_0. La forma quadratica associata ad A data da

$$\phi(x) = x^T A x = \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(x_0) x_i x_j$$

si dice *forma hessiana* di f in x_0. La forma hessiana è importante nello studio dei punti di massimo e di mininimo. Ad esempio infatti, utilizzando lo sviluppo di Taylor in x_0, è facile dimostrare che se in x_0 si annullano le derivate prime di f e risulta $\phi(x) > 0$ se $x \neq 0$, allora x_0 è un punto di minimo locale per f.

Definizione 3.2 Sia $\phi: \mathbb{R}^n \to \mathbb{R}$ la forma quadratica associata alla matrice simmetrica $A = (a_{ij})$. La forma ϕ e la matrice A si dicono

(i) *definite positive* se

$$\phi(x) = x^T A x = \sum_{i,j=1}^{n} a_{ij} x_i x_j > 0 \quad \forall x \neq 0; \tag{3.2}$$

(i) *definite negative* se

$$\phi(x) = x^T A x = \sum_{i,j=1}^{n} a_{ij} x_i x_j < 0 \quad \forall x \neq 0; \tag{3.3}$$

(i) *semidefinite positive* se

$$\phi(x) = x^T A x = \sum_{i,j=1}^{n} a_{ij} x_i x_j \geq 0 \quad \forall x \,\mathrm{e}\, \phi(x_0) = x_0^T A x_0 = 0 \text{ per qualche } x_0 \neq 0; \tag{3.4}$$

(i) *semidefinite negative* se

$$\phi(x) = x^T A x = \sum_{i,j=1}^{n} a_{ij} x_i x_j \leq 0 \quad \forall x \,\mathrm{e}\, \phi(x_0) = x_0^T A x_0 = 0 \text{ per qualche } x_0 \neq 0. \tag{3.5}$$

La forma ϕ e la matrice A si dicono *indefinite* se ϕ assume sia valori negativi sia positivi.

Come applicazione dei risultati precedenti, abbiamo la seguente:

Proposizione 3.1: *Sia $\phi\colon \mathbb{R}^n \to \mathbb{R}$ la forma quadratica associata alla matrice simmetrica $A = (a_{ij})$. Siano $\lambda_1, \ldots, \lambda_n$ gli autovalori (eventualmente con ripetizione) di A. Esiste una base ortonormale $\mathcal{B}$ di $\mathbb{R}^n$ tale che se x' sono le coordinate rispetto a $\mathcal{B}$ di un vettore $v \in \mathbb{R}^n$, allora ϕ è data da*

$$\phi(v) = \lambda_1(x'_1)^2 + \ldots \lambda_n(x'_n)^2. \tag{3.6}$$

Inoltre ϕ e A sono definite positive (negative) se e solo se tutti gli autovalori di A sono positivi (negativi); ϕ e A sono semidefinite positive (negative) se e solo se tutti gli autovalori di A sono non negativi (nonpositivi); ϕ e A sono indefinite se e solo se A ha autovalori sia positivi sia negativi.

Dimostrazione: Sia $\mathcal{B}$ una base ortonormale di $\mathbb{R}^n$ formata da autovettori di L_A e sia B la matrice del cambio di base dalla base canonica alla base $\mathcal{B}$. Allora, se x designano le coordinate di $v \in \mathbb{R}^n$ relative alla base canonica e x' quelle relative a $\mathcal{B}$, si ha $x = Bx'$. Dunque, ricordando che $D = B^T AB$ è una matrice diagonale con sulla diagonale gli autovalori di A, otteniamo

$$\phi(v) = x^T Ax = (Bx')^T A(Bx') = (x')^T (B^T AB)(x') = \lambda_1(x'_1)^2 + \ldots \lambda_n(x'_n)^2 \tag{3.7}$$

Dalla (3.7) è semplice studiare il segno della forma quadratica ϕ Ad esempio è immediato concludere che se tutti gli autovalori di A sono positivi allora ϕ è definita positiva. D'altra parte se ϕ è definita positiva, e $\mathcal{B} = \{v_1, \ldots, v_n\}$ allora per ogni j si ha $0 < \phi(v_j) = \lambda_j$. Le altre affermazioni si dimostrano in modo analogo. $\square$

Definizione 4.2 L'espressione (3.6) si dice *forma canonica metrica* della forma quadratica ϕ.

Esercizio. Si consideri la forma quadratica su $\mathbb{R}^3$ definita da $\phi(x) = 25x_1^2 - 7x_2^2 + 7x_3^2 + 48x_2x_3$. Trovare la forma canonica metrica di ϕ, decidere se è defoinita positiva e determinare una base ortonormale rispetto alla quale ϕ è in forma canonica metrica.

Si vede subito che ϕ è la forma quadratica associata alla matrice $A = \begin{pmatrix} 25 & 0 & 0 \\ 0 & -7 & 24 \\ 0 & 24 & 7 \end{pmatrix}$.

Il polinomio caratteristico di A è $p(\lambda) = -(\lambda - 25)^2(\lambda + 25)$ e quindi A ha autovalori 25 con molteplicità 2 e -25 con molteplicità 1. Dunque ϕ ha forma canonica metrica $\phi(x') = 25{x'}_1^2 + 25{x'}_2^2 - 25{x'}_3^2$. Calcolando gli autospazi, si vede facilmente che una base ortonormale dell'autospazio relativo all'autovalore 25 è data da $\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 3/5 \\ 4/5 \end{pmatrix} \right\}$ e una base ortonormale per l'autospazio relativo all'autovalore -25 è data da $\left\{ \begin{pmatrix} 0 \\ -4/5 \\ 3/5 \end{pmatrix} \right\}$. Si può allora concludere che una base ortonormale rispetto alla quale ϕ è in forma canonica metrica è $\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 3/5 \\ 4/5 \end{pmatrix}, \begin{pmatrix} 0 \\ -4/5 \\ 3/5 \end{pmatrix} \right\}$.