

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 $g: V \times V \rightarrow \mathbb{R}$ 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 $h: V \times V \rightarrow \mathbb{C}$ una funzione. Se per ogni $v, v_1, v_2, w, w_1, w_2 \in V$ e $\lambda \in \mathbb{R}$ si ha

$$h(v_1 + v_2, w) = h(v_1, w) + h(v_2, w), \quad h(v, w) = h(v, w_1) + h(v, w_2), \\ h(\lambda v, w) = \lambda h(v, w), \quad h(v, \lambda w) = \bar{\lambda} h(v, w)$$

h 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* $\langle \cdot, \cdot \rangle: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ di \mathbb{R}^n è la forma bilineare simmetrica definita per ogni scelta di $v, w \in \mathbb{R}^n$ da

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

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

$$\langle v, w \rangle_A = \langle Av, w \rangle.$$

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

$$\langle v, w \rangle_A = \langle Av, w \rangle = \langle w, Av \rangle = (Av)^T w = v^T A^T w = v^T A w = \langle Aw, v \rangle = \langle w, v \rangle_A.$$

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

$$\langle v, w \rangle = \sum_{i=1}^n v_i \bar{w}_i = w^H v$$

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 $\langle \cdot, \cdot \rangle_A: \mathbb{C}^n \times \mathbb{C}^n \rightarrow \mathbb{C}$ è definito da

$$\langle v, w \rangle_A = \langle Av, w \rangle.$$

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

$$\langle v, w \rangle_A = \langle Av, w \rangle = \overline{\langle w, Av \rangle} = \overline{(Av)^H w} = \overline{v^H A^H w} = \overline{v^H A w} = \overline{\langle Aw, v \rangle} = \overline{\langle w, v \rangle_A}.$$

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 $(\cdot, \cdot): \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ un prodotto scalare e siano e_1, \dots, 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 $(\cdot, \cdot) = \langle \cdot, \cdot \rangle_A$.
- (ii) Sia $(\cdot, \cdot): \mathbb{C}^n \times \mathbb{C}^n \rightarrow \mathbb{C}$ un prodotto hermitiano e siano e_1, \dots, 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 $(\cdot, \cdot) = \langle \cdot, \cdot \rangle_A$.

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

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

(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

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

□

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 $g: V \times V \rightarrow \mathbb{R}$ 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) $\langle \cdot, \cdot \rangle$ su uno spazio vettoriale V si dice *definito positivo* se per ogni $v \in V \setminus \{O\}$ si ha $\langle v, v \rangle > 0$ e *definito negativo* se per ogni $v \in V \setminus \{O\}$ si ha $\langle v, v \rangle < 0$. Si dice che $\langle \cdot, \cdot \rangle$ è *semidefinito positivo* se per ogni $v \in V$ si ha $\langle v, v \rangle \geq 0$ e esiste $v_0 \in V \setminus \{O\}$ tale che $\langle v_0, v_0 \rangle = 0$. Si dice che $\langle \cdot, \cdot \rangle$ è *semidefinito negativo* se per ogni $v \in V$ si ha $\langle v, v \rangle \leq 0$ e esiste $v_0 \in V \setminus \{O\}$ tale che $\langle v_0, v_0 \rangle = 0$. Infine $\langle \cdot, \cdot \rangle$ si dice *indefinito* se esistono $v, w \in V$ tali che $\langle v, v \rangle > 0$ e $\langle w, w \rangle < 0$.

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

$$\left\langle \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} \right\rangle = x_1 y_1 + x_2 y_2 + x_3 y_3 - x_4 y_4.$$

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) $\langle \cdot, \cdot \rangle$ definito positivo. Su uno spazio vettoriale metrico la funzione $\| \cdot \|: V \rightarrow \mathbb{R}_+$, definita da $\|v\| = \sqrt{\langle v, v \rangle}$, si dice *norma*.

Proposizione 1.2: Sia V uno spazio vettoriale metrico con prodotto $\langle \cdot, \cdot \rangle$ 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

$$\|\lambda v\| = |\lambda| \|v\|, \quad \|v + w\|^2 = \|v\|^2 + \|w\|^2 + 2\operatorname{Re} \langle v, w \rangle.$$

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

$$|\langle v, w \rangle| \leq \|v\| \|w\| \quad (1.1)$$

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:

$$\|v + w\| \leq \|v\| + \|w\|. \quad (1.2)$$

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

$$\langle v, w \rangle = \frac{1}{4} (\|v + w\|^2 - \|v - w\|^2) \quad \text{per } V \text{ su } \mathbb{R} \quad (1.3)$$

$$\operatorname{Re} \langle v, w \rangle = \frac{1}{4} (\|v + w\|^2 - \|v - w\|^2) \quad \text{e} \quad \operatorname{Im} \langle v, w \rangle = \frac{1}{4} (\|v + iw\|^2 - \|v - iw\|^2) \quad \text{per } V \text{ su } \mathbb{C}. \quad (1.4)$$

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

$$\|\lambda v\|^2 = \langle \lambda v, \lambda v \rangle = \lambda \bar{\lambda} \langle v, v \rangle = |\lambda|^2 \|v\|^2,$$

$$\|v + w\|^2 = \langle v + w, v + w \rangle = \langle v, v \rangle + \langle w, w \rangle + \langle v, w \rangle + \langle w, v \rangle = \|v\|^2 + \|w\|^2 + 2\operatorname{Re} \langle v, w \rangle.$$

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

$$0 \leq \|av + bw\|^2 = |a|^2 \|v\|^2 + |b|^2 \|w\|^2 + 2\operatorname{Re}(a\bar{b} \langle v, w \rangle).$$

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

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

Dividendo per $\|w\|^2 > 0$, la diseguaglianza di Cauchy-Schwarz segue. Si osservi che l'uguaglianza 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 per ogni v, w

$$|\operatorname{Re} \langle v, w \rangle| \leq |\langle v, w \rangle| \leq \|v\| \|w\|.$$

Dunque

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

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

$$\|v + w\|^2 = \|v\|^2 + \|w\|^2 + 2\operatorname{Re} \langle v, w \rangle \quad \text{e} \quad \|v - w\|^2 = \|v\|^2 + \|w\|^2 - 2\operatorname{Re} \langle v, w \rangle$$

e quindi

$$4\operatorname{Re} \langle v, w \rangle = \|v + w\|^2 - \|v - w\|^2.$$

In modo analogo calcoliamo

$$\|v + iw\|^2 = \|v\|^2 + \|w\|^2 + 2\operatorname{Im} \langle v, w \rangle \quad \text{e} \quad \|v - iw\|^2 = \|v\|^2 + \|w\|^2 - 2\operatorname{Im} \langle v, w \rangle$$

e quindi

$$4\operatorname{Im} \langle v, w \rangle = \|v + iw\|^2 - \|v - iw\|^2.$$

Nel caso reale si procede allo stesso modo. □

Osservazione. Sia V uno spazio vettoriale metrico con norma $\|\cdot\|$. La funzione $d: V \times V \rightarrow \mathbb{R}^+$ 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 disuguaglianza triangolare (1.2) segue la disuguaglianza 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 disuguaglianza 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

$$\cos \theta = \frac{\langle v, w \rangle}{\|v\| \|w\|}. \quad (1.5)$$

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

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

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 $\langle \cdot, \cdot \rangle$. Due vettori $v, w \in V$ si dicono *ortogonali* (o *perpendicolari*) se $\langle v, w \rangle = 0$ 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 $\{u_1, \dots, u_r\}$ una base di U e sia $v \in V$ arbitrario. Dimostrare che $v \perp U$ se e solo se $\langle v, u_i \rangle = 0$ per ogni $i = 1, \dots, r$.

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

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

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

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

$$0 = \langle v_i, O \rangle = \langle v_i, \alpha_1 v_1 + \dots + \alpha_n v_n \rangle = \alpha_i \langle v_i, v_i \rangle$$

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

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

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

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

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

$$v = \sum_{i=1}^n \langle v, v_i \rangle v_i \quad (\text{somma di Fourier}) \quad (1.6)$$

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

$$\langle v, w \rangle = \sum_{i=1}^n \langle v, v_i \rangle \langle v_i, w \rangle \quad (\text{formula di Parseval}) \quad (1.7)$$

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

$$\|v\|^2 = \sum_{i=1}^n |\langle v, v_i \rangle|^2 \quad (\text{teorema di Pitagora}) \quad (1.8)$$

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

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

e quindi (i) segue.

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

$$\begin{aligned} \langle v, w \rangle &= \left\langle \sum_{i=1}^n \langle v, v_i \rangle v_i, \sum_{j=1}^n \langle w, v_j \rangle v_j \right\rangle = \sum_{i,j=1}^n \langle v, v_i \rangle \overline{\langle w, v_j \rangle} \langle v_i, v_j \rangle \\ &= \sum_{i=1}^n \langle v, v_i \rangle \overline{\langle w, v_i \rangle} = \sum_{i=1}^n \langle v, v_i \rangle \langle v_i, w \rangle \end{aligned}$$

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 $v_1, \dots, v_r \in V$ vettori linearmente indipendenti. Allora i vettori w_1, \dots, w_n definiti per ricorrenza da

$$w_1 = v_1 \text{ e } w_j = v_j - \sum_{h=1}^{j-1} \frac{\langle v_j, w_h \rangle}{\langle w_h, w_h \rangle} w_h \quad \forall j = 2, \dots, r \quad (1.9)$$

soddisfano le seguenti proprietà:

- (i) $\text{Span}(w_1, \dots, w_j) = \text{Span}(v_1, \dots, v_j)$ per ogni $j = 1, \dots, r$;
- (ii) w_j è ortogonale a $\text{Span}(w_1, \dots, w_{j-1})$ per ogni $j = 2, \dots, r$ e quindi, in particolare w_1, \dots, w_r sono a due a due ortogonali;
- (iii) Se $\{v_1, \dots, v_r\}$ è una base di V e w_1, \dots, w_n sono definiti da (1.9), allora $\{\frac{w_1}{\|w_1\|}, \dots, \frac{w_r}{\|w_r\|}\}$ 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 $\text{Span}(w_1, \dots, w_{j-1}, v_j) = \text{Span}(v_1, \dots, v_{j-1}, v_j)$ e, per costruzione $w_j \in \text{Span}(w_1, \dots, w_{j-1}, v_j)$. Dato che per ogni $k < j$ si ha

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

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

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. \square

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 $U^\perp = \text{Span}(v_{r+1}, \dots, v_n)$. \square

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 $P_U: V \rightarrow V$ chiamata proiezione ortogonale sul sottospazio U , che ha le seguenti proprietà:

- i) $\text{Im}P_U = U$ 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 $v = \sum_{i=1}^n \langle v, v_i \rangle v_i \in V$ si ha

$$P_U(v) = \sum_{i=1}^r \langle v, v_i \rangle v_i; \quad (1.11)$$

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 disuguaglianza

$$\|v - u\| \geq \|v - P_U(v)\| \quad (1.12)$$

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 $P_U: V \rightarrow V$ tale che $\text{Im}P_U = U$. 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, $v - P_U(v) = \sum_{i=r+1}^n \langle v, v_i \rangle v_i$ e quindi $\langle v - P_U(v), v_i \rangle = 0$ 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

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

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

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

con l'eguaglianza che vale se e solo se $\|P_U(v) - u\|^2 = 0$ ossia se e solo se $u = P_U(v)$. \square

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

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

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):

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

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

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:

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

Una 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 $\langle v_1, z \rangle = \langle v_2, z \rangle = \langle v_3, z \rangle = 0$. Aggiungendo il vettore $z/\|z\|$ si ottiene una base ortonormale di \mathbb{R}^4 .

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

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

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

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

Si osservi che B è la trasposta di B^{-1} : questo è 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 esclusivamente 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 $\langle v, w \rangle = 0$ per ogni $v \in V \iff w = O$.

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

$$w = \sum_{j=1}^n \langle w, v_j \rangle v_j = O.$$

□

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

$$\langle T(v), w \rangle_W = \langle v, T^*(w) \rangle_V. \quad (1.13)$$

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

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

Dimostrazione: Esistenza di T^* : Siano $\mathcal{B} = \{v_1, \dots, v_n\}$ e $\mathcal{C} = \{w_1, \dots, w_m\}$ rispettivamente di V e W . Se A è la matrice associata a T relativa a \mathcal{B} e \mathcal{C} , allora l'applicazione $T^*: W \rightarrow V$ 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

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

Pertanto

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

e analogamente

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

e quindi la (1.13) è verificata.

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

$$\langle v, T^*(w) \rangle_V = \langle T(v), w \rangle_W = \langle v, S(w) \rangle_V$$

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

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

Osservazione 1. Sia $T: V \rightarrow W$ un'applicazione lineare fra spazi metrici e sia $T^*: W \rightarrow V$ 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 *qualsunque* 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 $L_A: \mathbb{R}^n \rightarrow \mathbb{R}^m$ è $L_{A^T}: \mathbb{R}^m \rightarrow \mathbb{R}^n$. Basta considerare su \mathbb{R}^n e \mathbb{R}^m le basi canoniche e 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 $w \in \mathbb{R}^m$

$$\langle L_A(v), w \rangle_{\mathbb{R}^m} = w^T Av = (w^T Av)^T = v^T A^T w = \langle L_{A^T}(w), v \rangle_{\mathbb{R}^n} = \langle v, L_{A^T}(w) \rangle_{\mathbb{R}^n}$$

Esempio. Si consideri su \mathbb{C}^n il prodotto scalare canonico. Se $A \in M_{m,n}(\mathbb{C})$, l'aggiunta di $L_A: \mathbb{C}^n \rightarrow \mathbb{C}^m$ è $L_{A^H}: \mathbb{C}^m \rightarrow \mathbb{C}^n$. Basta considerare su \mathbb{C}^n e \mathbb{C}^m le basi canoniche e 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 $w \in \mathbb{C}^m$

$$\langle L_A(v), w \rangle_{\mathbb{C}^m} = w^H Av = \overline{(w^H Av)^H} = \overline{v^H A^H w} = \langle L_{A^H}(w), v \rangle_{\mathbb{C}^n} = \langle v, L_{A^H}(w) \rangle_{\mathbb{C}^n}$$

Definizione. Un endomorfismo $T: V \rightarrow W$ di uno spazio metrico si dice *autoaggiunto* se $T = T^*$ ossia se $\langle T(v), w \rangle = \langle v, T(w) \rangle$ per ogni $v, w \in V$.

Proposizione 1.11: Sia $T: V \rightarrow V$ 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. □

Definizione. Un endomorfismo $T: V \rightarrow V$ di uno spazio metrico si dice *isometria* se $\langle T(v), T(w) \rangle = \langle v, w \rangle$ 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:

$$(i) \implies (ii) \implies (iii) \implies (iv) \implies (i) \text{ e } (ii) \iff (v).$$

(i) \implies (ii) Per ogni $v, w \in V$ si ha $\langle v, w \rangle = \langle T(v), T(w) \rangle = \langle v, T^*(T(w)) \rangle$ e quindi $\langle v, w - T^*(T(w)) \rangle = 0$.

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 biiettivo e quindi invertibile. Per l'unicità dell'inverso, ancora da (1.15) segue $T^{-1} = T^*$.

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

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

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

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

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

Allora

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

(iv) \implies (i) Utilizzando la formula di polarizzazione, nel caso reale abbiamo:

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

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

(ii) \iff (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

$$\begin{aligned} (L_A)^{-1} &= L_{A^{-1}} = L_{A^T} = (L_A)^* \quad \text{se } \mathbb{K} = \mathbb{R}, \\ (L_A)^{-1} &= L_{A^{-1}} = L_{A^H} = (L_A)^* \quad \text{se } \mathbb{K} = \mathbb{C}. \end{aligned}$$

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

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

ossia se e solo se $\{A^1, \dots, A^n\}$ 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:

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

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

Utilizzeremo le seguenti notazioni: $O(n) = \{A \in M_{n,n}(\mathbb{R}) \mid A^T A = I_n\}$ per il gruppo delle matrici ortogonali di ordine n e $U(n) = \{A \in M_{n,n}(\mathbb{C}) \mid A^H A = I_n\}$ 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 $T: V \rightarrow V$ un endomorfismo e siano A la matrice di T relativa a \mathcal{B} e A' la matrici di T relativa \mathcal{C} . Allora

$$A' = B^{-1} A B = \begin{cases} B^T A B & \text{se } \mathbb{K} = \mathbb{R} \\ B^H A B & \text{se } \mathbb{K} = \mathbb{C}. \end{cases}$$

Due matrici A, A' reali quadrate di ordine n si dicono *congruenti* se esiste una matrice invertibile B tale che $A' = B^T A B$.

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 $T: V \rightarrow V$ 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 $L = L_A: \mathbb{C}^n \rightarrow \mathbb{C}^n$. 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

$$\lambda \|u\|^2 = \lambda \langle u, u \rangle = \langle \lambda u, u \rangle = \langle Lu, u \rangle = \langle u, Lu \rangle = \bar{\lambda} \langle u, u \rangle = \bar{\lambda} \|u\|^2$$

e quindi $(\lambda - \bar{\lambda})\|u\|^2 = 0$. Dato che $\|u\|^2 \neq 0$, allora deve essere $\lambda - \bar{\lambda} = 0$ 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 $T: V \rightarrow V$. Le seguenti affermazioni sono equivalenti:

(i) Esiste una base ortonormale $\{v_1, \dots, v_n\}$ di V e numeri reali $\lambda_1, \dots, \lambda_n$ 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) \iff (ii) Se vale (i) allora si ha $T(v_i) = \lambda_i v_i$ e quindi vale (ii). D'altro canto se $\{v_1, \dots, v_n\}$ 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) \iff (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 $W = (\text{Span}(v_1))^\perp$. Allora W è un sottospazio di V di dimensione $n - 1$ e, per costruzione, si ha che $w \in W$ se e solo se $\langle w, v_1 \rangle = 0$ e quindi $W = \{w \in V \mid \langle w, v_1 \rangle = 0\}$.

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

$$\langle u, v_1 \rangle = \langle T(w), v_1 \rangle = \langle w, T(v_1) \rangle = \langle w, \lambda v_1 \rangle = \bar{\lambda} \langle w, v_1 \rangle = 0$$

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 $T: W \rightarrow W$ è un endomorfismo autoaggiunto di uno spazio vettoriale metrico di dimensione $n - 1$. Esiste allora, per l'ipotesi induttiva, una base ortonormale $\{v_2, \dots, v_n\}$ formata da autovettori di T ciascuno con autovalore reale. Dato che $\langle v_j, v_1 \rangle = 0$ per $j = 2, \dots, n$, allora $\{v_1, v_2, \dots, v_n\}$ è una base ortonormale di V formata da autovettori di T ciascuno con autovalore reale e quindi (ii) vale. \square

Conseguenza immediata del Teorema 2.2 è il seguente

Corollario 2.3: (Teorema spettrale) Sia $T: V \rightarrow V$ 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 $D = B^{-1}AB = B^T AB$ 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 $D = B^{-1}AB = B^H AB$ è una matrice diagonale reale.

Dimostrazione: (i) Si consideri $L_A: \mathbb{R}^n \rightarrow \mathbb{R}^n$. 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. \square

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 $T: V \rightarrow V$ 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

$$\mu \langle v, w \rangle = \langle \mu v, w \rangle = \langle T(v), w \rangle = \langle v, T(w) \rangle = \langle v, \nu w \rangle = \nu \langle v, w \rangle$$

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

La Proposizione 2.5 ci dice che uno spazio vettoriale metrico si decompone in somma diretta ortogonale di autospazi di un endomorfismo autoaggiunto. Suggerisce quindi di procedere per diagonalizzare un endomorfismo autoaggiunto $T: V \rightarrow V$ 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, grazie alla Proposizione 2.5 è una base ortonormale diagonalizzante per T .

Data una matrice simmetrica reale A si può adattare il procedimento per trovare una matrice ortogonale B tale che $D = B^T A B = B^{-1} A B$ 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 ortonormale 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 molteplicità 2. Se V_1, V_{-1} sono i rispettivi autospazi si ha

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

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

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

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

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 algebrica, 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 $L_A: \mathbb{C}^n \rightarrow \mathbb{C}^n$ e sia z_1 un autovettore relativo a λ_1 con $\|z_1\| = 1$. Sia infine $\{z_1, z_2, \dots, z_n\}$ una base ortonormale di \mathbb{C}^n . La matrice B di L_A relativa a questa base ha questa forma:

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

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 $S' = (U'_1)^{-1}A_1U'_1 = U_1^H A_1 U_1$ è triangolare superiore. Se definiamo

$$U_1 = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & U'_1 & \\ 0 & & & \end{pmatrix},$$

allora U_1 è una matrice unitaria di ordine n tale che

$$\begin{aligned} S &= U_1^{-1}BU_1 = U_1^H BU_1 \\ &= \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & (U'_1)^H & & \\ 0 & & & \end{pmatrix} \begin{pmatrix} \lambda_1 & * & \dots & * \\ 0 & & & \\ \vdots & & A_1 & \\ 0 & & & \end{pmatrix} \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & U'_1 & \\ 0 & & & \end{pmatrix} \\ &= \begin{pmatrix} \lambda_1 & * & \dots & * \\ 0 & & & \\ \vdots & & (U'_1)^H A_1 U'_1 & \\ 0 & & & \end{pmatrix} = \begin{pmatrix} \lambda_1 & * & \dots & * \\ 0 & & & \\ \vdots & & S' & \\ 0 & & & \end{pmatrix} \end{aligned}$$

è triangolare superiore. Ma allora la matrice unitaria $U = U_0 U_1$ è la matrice cercata dato che, rimettendo insieme i calcoli fatti, si ha $S = U^{-1}AU = U^H AU$. \square

3. Forme quadratiche reali

Definizione 3.1 Una forma quadratica su \mathbb{R}^n è una funzione $\phi: \mathbb{R}^n \rightarrow \mathbb{R}$ definita da $\phi(x) = \langle x, x \rangle_\phi$ dove $\langle \bullet, \bullet \rangle_\phi$ un prodotto scalare (non necessariamente definito positivo) su \mathbb{R}^n .

Sappiamo che dato un prodotto scalare $\langle \bullet, \bullet \rangle_\phi$, se $A = (a_{ij}) \in M_{n,n}$ è la matrice simmetrica definita da

$$a_{ij} = \langle e_j, e_i \rangle_\phi \quad (3.1)$$

si ha $\langle x, y \rangle_\phi = y^T A x$ per ogni $x, y \in \mathbb{R}^n$. Dunque, in particolare si ha

$$\phi(x) = \langle x, x \rangle_\phi = x^T A x = \sum_{i,j=1}^n a_{ij} x_i x_j. \quad (3.2)$$

Chiameremo la matrice A definita in (3.1) la *matrice associata a ϕ rispetto alla base canonica*

Osservazione. Dalla (3.2) si vede subito che una forma quadratica su \mathbb{R}^n non è altro che un polinomio reale omogeneo di grado 2 in n variabili ossia un polinomio reale $p(x_1, \dots, x_n)$ in n variabili tale che per ogni $t \in \mathbb{R}^n$ risulta:

$$p(tx_1, \dots, tx_n) = t^2 p(x_1, \dots, x_n). \quad (3.3)$$

Viceversa dato un polinomio reale omogeneo di grado 2 in n variabili $p(x_1, \dots, x_n)$, si vede subito che

$$p(x_1, \dots, x_n) = \sum_{i,j=1}^n a_{ij} x_i x_j$$

dove $A = (a_{ij}) \in M_{n,n}$ è una matrice simmetrica. Dunque il polinomio $p(x_1, \dots, x_n)$ definisce una forma quadratica con matrice associata A .

Esempio. Sia $f: \mathbb{R}^n \rightarrow \mathbb{R}$ una funzione che ammette derivate parziali seconde continue. 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

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

si dice *forma hessiana* di f in x_0 . La forma hessiana è importante nello studio dei punti di massimo e di minimo. Nei corsi di analisi, utilizzando lo sviluppo di Taylor in x_0 , si dimostra facilmente che se in x_0 si annullano le derivate prime di f e risulta $H_f(x) > 0$ se $x \neq 0$, allora x_0 è un punto di minimo locale per f . Sempre senza grandi difficoltà si dimostra che se $f: \mathbb{R}^n \rightarrow \mathbb{R}$ è una funzione che ammette derivate parziali seconde continue e che è omogenea di grado 2, ossia è tale che per ogni $x \in \mathbb{R}^n$ e $t \in \mathbb{R}^n$:

$$f(tx) = t^2 f(x),$$

allora per ogni $x \in \mathbb{R}^n$ si ha

$$f(x) = \frac{1}{4} H_f(x).$$

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

(i) *definite positive* se

$$\phi(x) = x^T A x = \sum_{i,j=1}^n a_{ij} x_i x_j > 0 \quad \forall x \neq 0; \quad (3.4)$$

(i) *definite negative* se

$$\phi(x) = x^T A x = \sum_{i,j=1}^n a_{ij} x_i x_j < 0 \quad \forall x \neq 0; \quad (3.5)$$

(i) *semidefinite positive* se

$$\phi(x) = x^T A x = \sum_{i,j=1}^n a_{ij} x_i x_j \geq 0 \quad \forall x \text{ e } \phi(x_0) = x_0^T A x_0 = 0 \text{ per qualche } x_0 \neq 0; \quad (3.6)$$

(i) *semidefinite negative* se

$$\phi(x) = x^T A x = \sum_{i,j=1}^n a_{ij} x_i x_j \leq 0 \quad \forall x \text{ e } \phi(x_0) = x_0^T A x_0 = 0 \text{ per qualche } x_0 \neq 0. \quad (3.7)$$

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

La domanda cui vogliamo rispondere è la seguente: esiste un cambio di base per \mathbb{R}^n tale che nelle nuove coordinate una forma quadratica prende una espressione così semplice che le sue proprietà (ad esempio quelle di positività viste sopra) diventano semplici da “leggere”?

Per cominciare vediamo come cambia l’espressione di una forma quadratica cambiando base:

Proposizione 3.1: Sia $\phi: \mathbb{R}^n \rightarrow \mathbb{R}$ una forma quadratica e sia $A = (a_{ij})$ la matrice associata a ϕ rispetto alla base canonica. Sia $\mathcal{B} = \{\epsilon_1, \dots, \epsilon_n\}$ una base di \mathbb{R}^n e sia B la matrice del cambio di base dalla base canonica alla base \mathcal{B} . Se $A' = B^T A B$, allora per ogni $v \in \mathbb{R}^n$, se x' sono le coordinate di v relative a \mathcal{B} , si ha

$$\phi(v) = (x')^T A' x'. \quad (3.8)$$

La matrice $A' = (a'_{ij})$ si dice matrice associata a ϕ rispetto alla base \mathcal{B} e si ha $a'_{ij} = \langle \epsilon_j, \epsilon_i \rangle_\phi$.

Dimostrazione: Se denotiamo con x le coordinate di $v \in \mathbb{R}^n$, allora $x = Bx'$ e quindi

$$\phi(v) = x^T A x = (Bx')^T A Bx' = (x')^T B^T A Bx' = (x')^T A' x'.$$

Per quanto riguarda l’ultima affermazione, si osservi che se $B = (B^1, \dots, B^n)$ è l’espressione per colonne di B si ha $B^k = \epsilon_k$ per ogni k . Inoltre ricordiamo che se $B^T = ((B^T)_1, \dots, (B^T)_n)$ è l’espressione per righe di B^T , si ha $(B^T)_k = (B^k)^T$ per ogni k . Con questo in mente si ha:

$$a'_{ij} = (B^T)_i (A B)^j = (B^i)^T A B^j = \langle \epsilon_j, \epsilon_i \rangle_\phi.$$

□

Come applicazione dei risultati precedenti, abbiamo la seguente:

Proposizione 3.2: Sia $\phi: \mathbb{R}^n \rightarrow \mathbb{R}$ la forma quadratica associata alla matrice simmetrica $A = (a_{ij})$. Siano $\lambda_1, \dots, \lambda_n$ 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

$$\phi(v) = \lambda_1 (x'_1)^2 + \dots + \lambda_n (x'_n)^2. \quad (3.9)$$

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 $D = B^T A B$ è una matrice diagonale con sulla diagonale gli autovalori di A , otteniamo immediatamente dalla Proposizione 3.1:

$$\phi(v) = \lambda_1 (x'_1)^2 + \dots + \lambda_n (x'_n)^2 \quad (3.10)$$

Dalla (3.10) 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 $\mathcal{B} = \{v_1, \dots, v_n\}$ allora per ogni j si ha $0 < \phi(v_j) = \lambda_j$. Le altre affermazioni si dimostrano in modo analogo. □

Definizione 4.2 L'espressione (3.9) si dice *forma canonica metrica* della forma quadratica ϕ .

Si osservi che mentre per calcolare la forma canonica metrica di una forma quadratica occorre riuscire a calcolare tutti gli autovalori della matrice associata A , per studiarne il segno basta conoscere il segno degli autovalori. A tal fine viene in aiuto il Criterio di Cartesio che, non ha in dimostrazione semplice per polinomi di grado arbitrario, ma è molto semplice da enunciare e da utilizzare. Lo ricordiamo rapidamente.

Criterio di Cartesio Sia $P(t) = a_k t^k + \dots + a_d t^d$, dove $0 \leq d \leq k$ e $a_d \neq 0$, un polinomio a coefficienti reali con tutte le radici reali. Allora:

- (i) 0 è radice di $P(t)$ se e solo se $d \geq 1$ e in questo caso è radice di molteplicità d .
- (ii) il numero p delle radici positive del polinomio $P(t)$, contate con la loro molteplicità, è uguale al numero delle variazioni di segno nella successione dei coefficienti non nulli di $P(t)$.
- (iii) il numero delle radici negative del polinomio $P(t)$, contate con la loro molteplicità, è $n = k - d - p$.

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 definita 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 $\phi(x') = 25x_1'^2 + 25x_2'^2 - 25x_3'^2$. 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 $\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 3/5 \\ 4/5 \end{pmatrix}, \begin{pmatrix} 0 \\ -4/5 \\ 3/5 \end{pmatrix} \right\}$.