

Spazi vettoriali metrici.

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

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

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

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* $\langle \cdot, \cdot \rangle: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ 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:

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

In realtà si dimostra facilmente che i prodotti scalari e hermitiani definiti negli Esempi 1 e 2 sono tutti quelli possibili. Otterremo questo fatto come conseguenza della descrizione dei prodotti scalari in termini di coordinate:

Proposizione 1.1:

- (i) Sia V uno spazio vettoriale reale di dimensione n e sia $(\cdot, \cdot): V \times V \rightarrow \mathbb{R}$ un prodotto scalare. Fissata una base $\mathcal{B} = \{v_1, \dots, v_n\}$ di V , sia $A = (a_{kj})$ è la matrice simmetrica definita da $a_{kj} = (v_j, v_k)$. Allora, se per $u, w \in V$ si ha $x = F_{\mathcal{B}}(u)$ e $y = F_{\mathcal{B}}(w)$, ossia $x, y \in \mathbb{R}^n$ sono le colonne delle coordinate relative alla base \mathcal{B} rispettivamente di u e w , si ha:

$$(u, w) = \langle x, y \rangle_A .$$

In particolare allora se $(\cdot, \cdot): \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ è un prodotto scalare e e_1, \dots, e_n sono 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 V uno spazio vettoriale complesso di dimensione n e sia $(\cdot, \cdot): V \times V \rightarrow \mathbb{R}$ un prodotto hermitiano. Fissata una base $\mathcal{B} = \{v_1, \dots, v_n\}$ di V , sia $A = (a_{kj})$ la matrice hermitiana definita da $a_{kj} = (v_j, v_k)$. Allora, se per $u, w \in V$ si ha $x = F_{\mathcal{B}}(u)$ e $y = F_{\mathcal{B}}(w)$, ossia $x, y \in \mathbb{C}^n$ sono le colonne delle coordinate relative alla base \mathcal{B} rispettivamente di u e w , si ha:

$$(u, w) = \langle x, y \rangle_A .$$

In particolare allora se $(\cdot, \cdot): \mathbb{C}^n \times \mathbb{C}^n \rightarrow \mathbb{C}$ è un prodotto hermitiano e e_1, \dots, e_n sono 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 $F_{\mathcal{B}}(u) = x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, $F_{\mathcal{B}}(w) = y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{R}^n$. Allora

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

Il resto del primo enunciato non è altro che un caso particolare nel quale $V = \mathbb{R}^n$ e la base scelta è quella canonica.

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

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

Il resto del secondo enunciato non è altro che un caso particolare nel quale $V = \mathbb{C}^n$ e la base scelta è quella canonica. \square

Fissata dunque una base \mathcal{B} per uno spazio vettoriale finitamente generato reale (risp. complesso) la matrice simmetrica (risp. hermitiana) A definita nella Proposizione 1.1 permette dunque di descrivere in modo completo il prodotto scalare (risp. hermitiano) (\cdot, \cdot) . Chiameremo A matrice associata al (o che rappresenta il) prodotto scalare (risp. hermitiano) (\cdot, \cdot) rispetto alla base \mathcal{B} . È interessante chiarire come la matrice A dipende dalla scelta della base. A tal riguardo abbiamo la seguente:

Proposizione 1.2:

- (i) Sia V uno spazio vettoriale reale di dimensione n e sia $(\cdot, \cdot): V \times V \rightarrow \mathbb{R}$ prodotto scalare. Siano $\mathcal{B} = \{v_1, \dots, v_n\}$ e $\mathcal{B}' = \{v'_1, \dots, v'_n\}$ basi di V e $A = (a_{kj})$, $A' = (a'_{kj})$ le matrici associate al prodotto scalare (\cdot, \cdot) . Allora, se B è la matrice del cambio di base da \mathcal{B} a \mathcal{B}' , si ha:

$$A' = B^T A B.$$

(ii) Sia V uno spazio vettoriale complesso di dimensione n e sia $(\cdot, \cdot): V \times V \rightarrow \mathbb{R}$ prodotto hermitiano. Siano $\mathcal{B} = \{v_1, \dots, v_n\}$ e $\mathcal{B}' = \{v'_1, \dots, v'_n\}$ basi di V e $A = (a_{kj})$, $A' = (a'_{kj})$ le matrici associate al prodotto hermitiano (\cdot, \cdot) . Allora, se B è la matrice del cambio di base da \mathcal{B} a \mathcal{B}' , si ha:

$$A' = B^H AB.$$

Dimostrazione: (i): Per $u, w \in V$, Se $x = F_{\mathcal{B}}(u)$ e $y = F_{\mathcal{B}}(w)$ e $x' = F_{\mathcal{B}'}(u)$ e $y' = F_{\mathcal{B}'}(w)$ le matrici A e A' sono le uniche matrici tali che

$$y^T Ax = (u, w) = (y')^T A' x'.$$

D'altra parte, dato che $x = Bx'$ e $y = By'$, abbiamo che

$$(y')^T A' x' = (u, w) = y^T Ax = (By')^T A(Bx') = (y')^T B^T ABx' = (y')^T (B^T AB)x'$$

e quindi la tesi segue. Per (ii) si procede allo stesso modo.

Definizione. Due matrici $A, A' \in M_n(\mathbb{K})$ quadrate di ordine n con entrate nel campo \mathbb{K} si dicono *congruenti* se esiste una matrice invertibile $B \in GL_n(\mathbb{K})$ tale che $A' = B^T AB$. L'insieme di tutte le matrici congruenti a una matrice $A \in M_n(\mathbb{K})$ si dice classe di congruenza definita da A .

Dunque due matrici simmetriche rappresentano lo stesso prodotto scalare su uno spazio vettoriale reale finitamente generato rispetto a basi diverse se e solo se sono congruenti.

2. Spazi vettoriali metrici

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 $\langle w, w \rangle < 0$.

Esempio 3. Siano $x_0, \dots, x_n \in \mathbb{R}$. Un prodotto scalare $\langle \cdot, \cdot \rangle$ sullo spazio $\mathbb{R}_n[x]$ dei polinomi reali di grado al più n è definito da

$$\langle p(x), q(x) \rangle = \sum_{j=0}^n p(x_j)q(x_j)$$

per ogni $p(x), q(x) \in \mathbb{R}_n[x]$. Si vede subito che questo prodotto scalare è definito positivo.

Esempio 4. 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.

Esempio 4. Sia $M_n(\mathbb{R})$ lo spazio delle matrici quadrate reali di ordine n . Ricordiamo che la traccia $Tr(A)$ di una matrice $A = (a_{ij}) \in M_n(\mathbb{R})$ è definita da $Tr(A) = \sum_{j=1}^n a_{jj}$. Il prodotto scalare definito su $M_n(\mathbb{R})$ per $A = (a_{ij}), B = (b_{ij}) \in M_n(\mathbb{R})$

$$\langle A, B \rangle = Tr(B^T A) = \sum_{j=1}^n \sum_{k=1}^n b_{kj} a_{kj}$$

definisce un prodotto scalare definito positivo.

Esempio 6. 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. Questo prodotto scalare 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 2.1: 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 disuguaglianza di Cauchy-Schwarz:

$$|\langle v, w \rangle| \leq \|v\| \|w\| \quad (2.1)$$

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

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

$$\|v + w\| \leq \|v\| + \|w\|. \quad (2.2)$$

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

$$\langle v, w \rangle = \frac{1}{4} (\|v + w\|^2 - \|v - w\|^2) \quad \text{per } V \text{ su } \mathbb{R} \quad (2.3)$$

$$\operatorname{Re} \langle v, w \rangle = \frac{1}{4} (\|v + w\|^2 - \|v - w\|^2) \quad \text{e} \quad \operatorname{Im} \langle v, w \rangle = \frac{1}{4} (\|v + iw\|^2 - \|v - iw\|^2) \quad \text{per } V \text{ su } \mathbb{C}. \quad (2.4)$$

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 disuguaglianza 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 disuguaglianza 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 (2.2) 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

$$\cos \theta = \frac{\langle v, w \rangle}{\|v\|\|w\|}. \quad (2.5)$$

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 (2.5).

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 2.2: 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 ortogonali, e quelle ortonormali in particolare, sono strumenti molto utiliper studiare gli spazi vettoriali metrici perché possono essere utilizzate per esprimere in modo semplice le proprietà metriche. Questo fatto è illustrato ad esempio nella seguente

Proposizione 2.3: Sia $\mathcal{B} = \{v_1, \dots, v_n\}$ una base ortogonale di V . Allora valgono le seguenti formule:

(i) Per ogni $v \in V$ si ha $v = \sum_{i=1}^n \frac{\langle v, v_i \rangle}{\|v_i\|^2} v_i$ e, per \mathcal{B} ortonormale,

$$v = \sum_{i=1}^n \langle v, v_i \rangle v_i \quad (\text{somma di Fourier}) \quad (2.6)$$

(ii) Per ogni $v, w \in V$ si ha $\langle v, w \rangle = \sum_{i=1}^n \frac{\langle v, v_i \rangle \langle v_i, w \rangle}{\|v_i\|^2}$ e, per \mathcal{B} ortonormale,

$$\langle v, w \rangle = \sum_{i=1}^n \langle v, v_i \rangle \langle v_i, w \rangle \quad (\text{formula di Parseval}) \quad (2.7)$$

(iii) Per ogni $v \in V$ si ha: $\|v\|^2 = \sum_{i=1}^n \frac{|\langle v, v_i \rangle|^2}{\|v_i\|^2}$ e, per \mathcal{B} ortonormale,

$$\|v\|^2 = \sum_{i=1}^n |\langle v, v_i \rangle|^2 \quad (\text{teorema di Pitagora}) \quad (2.8)$$

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 $i_0 \in \{1, \dots, n\}$. Allora utilizzando il fatto che $\{v_1, \dots, v_n\}$ è una base ortogonale:

$$\langle v, v_{i_0} \rangle = \left\langle \sum_{i=1}^n \alpha_i v_i, v_{i_0} \right\rangle = \sum_{i=1}^n \alpha_i \langle v_i, v_{i_0} \rangle = \alpha_{i_0} \langle v_{i_0}, v_{i_0} \rangle$$

e quindi (i) segue.

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

$$\begin{aligned} \langle v, w \rangle &= \left\langle \sum_{i=1}^n \frac{\langle v, v_i \rangle}{\|v_i\|^2} v_i, \sum_{j=1}^n \frac{\langle w, v_j \rangle}{\|v_j\|^2} v_j \right\rangle = \sum_{i,j=1}^n \frac{\langle v, v_i \rangle \overline{\langle w, v_j \rangle}}{\|v_i\|^2 \|v_j\|^2} \langle v_i, v_j \rangle \\ &= \sum_{i=1}^n \frac{\langle v, v_i \rangle \overline{\langle w, v_i \rangle}}{\|v_i\|^2} = \sum_{i=1}^n \frac{\langle v, v_i \rangle \langle v_i, w \rangle}{\|v_i\|^2} \end{aligned}$$

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

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 2.4: (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_r definiti per ricorrenza da*

$$w_1 = v_1 \text{ e } w_j = v_j - \sum_{h=1}^{j-1} \frac{\langle v_j, w_h \rangle}{\langle w_h, w_h \rangle} w_h \quad \forall j = 2, \dots, r \quad (2.9)$$

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_r sono definiti da (2.9), 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

$$\begin{aligned} \langle w_j, w_k \rangle &= \left\langle v_j - \sum_{h=1}^{j-1} \frac{\langle v_j, w_h \rangle}{\langle w_h, w_h \rangle} w_h, w_k \right\rangle = \langle v_j, w_k \rangle - \sum_{h=1}^{j-1} \frac{\langle v_j, w_h \rangle}{\langle w_h, w_h \rangle} \langle w_h, w_k \rangle, \\ &= \langle v_j, w_k \rangle - \frac{\langle v_j, w_k \rangle}{\langle w_k, w_k \rangle} \langle w_k, w_k \rangle = 0 \end{aligned} \quad (2.10)$$

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 (2.10) 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 2.5: *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 2.4. \square

Corollario 2.6: Sia V uno spazio vettoriale metrico di dimensione n e sia $U \subset V$ un sottospazio di dimensione r di V . Allora

$$U^\perp = \{v \in V \mid \langle v, u \rangle = 0 \quad \forall v \in V\},$$

si dice complemento ortogonale di U ed è l'unico sottospazio di V tale che $U \perp U^\perp$ e $V = U \oplus U^\perp$. 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 $U^\perp = \text{Span}(v_{r+1}, \dots, v_n)$.

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 se $\tilde{U} = \text{Span}(v_{r+1}, \dots, v_n)$, si vede subito che $\tilde{U} = U^\perp$. Infatti $v \in \tilde{U}$ se e solo se $\langle v, v_j \rangle = 0$ per ogni $j = 1, \dots, r$ e quindi se e solo se $\langle v, u \rangle = 0$ per ogni $u \in U$ ossia se e solo se $u \in U^\perp$. Questo dimostra che U^\perp è un sottospazio supplementare di U . L'unicità è ovvia perchè se $V = U \oplus U'$ con $U \perp U'$, allora $v \in U'$ se e solo se $v \perp v_j$ per ogni $j = 1, \dots, r$ ossia se e solo se $v \in \tilde{U}$. \square

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

Proposizione 2.7: 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

$$P_U(v) = \sum_{i=1}^r \langle v, v_i \rangle v_i; \quad (2.11)$$

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

$$\|v - u\| \geq \|v - P_U(v)\| \quad (2.12)$$

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

Dimostrazione: È immediato convincersi che la formula (2.11) 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 (2.12). 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 (2.9):

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

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

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 più semplice per trovare un vettore che completi \mathcal{B}' a una base ortonormale di \mathbb{R}^4 è il seguente. Per trovare un vettore ortogonale a U basta osservare che

$$U = \left\{ x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \mid x_1 + x_2 = x_3 + x_4 \right\} = \{x \mid 0 = -x_1 - x_2 + x_3 + x_4\},$$

e quindi se $z = \begin{pmatrix} -1 \\ -1 \\ 1 \\ 1 \end{pmatrix}$, allora $U = \{x \mid \langle x, z \rangle = 0\}$ e quindi z è ortogonale a U e quindi a ciascuno dei

w_1, w_2, w_3 . Aggiungendo il vettore w_4 a $\{w_1, w_2, w_3\}$ si ottiene una base ortogonale di \mathbb{R}^4 e quindi \mathcal{B}' si completa a una base ortonormale aggiungendo $\frac{w_4}{\|w_4\|}$.

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

$$P_U \left(\frac{w_1}{\|w_1\|} \right) = \frac{w_1}{\|w_1\|}, \quad P_U \left(\frac{w_2}{\|w_2\|} \right) = \frac{w_2}{\|w_2\|}, \quad P_U \left(\frac{w_3}{\|w_3\|} \right) = \frac{w_3}{\|w_3\|}, \quad P_U \left(\frac{w_4}{\|w_4\|} \right) = 0,$$

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

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

3. Aggiunta e endomorfismi autoaggiunti. Isometrie, matrici ortogonali e unitarie.

Da questo punto in poi consideriamo esclusivamente spazi vettoriali metrici di dimensione finita. Vogliamo ora considerare classi di applicazioni lineari su spazi vettoriali metrici che hanno proprietà di compatibilità con la struttura metrica.

Ci saranno utili due semplici e ovvie osservazioni.

Osservazione 1. Sia V uno spazio vettoriale metrico e $w \in V$. Allora $\langle v, w \rangle = 0$ per ogni $v \in V \iff w = O$. infatti se $w = O$ allora ovviamente $\langle v, w \rangle = 0$ per ogni $v \in V$. Viceversa se $\langle v, w \rangle = 0$ per ogni $v \in V$ allora in particolare $\langle w, w \rangle = 0$ e quindi necessariamente $w = O$.

Osservazione 2. Fissata una base ortonormale \mathcal{B} su uno spazio vettoriale metrico V di dimensione n su $\mathbb{K} = \mathbb{R}$ (o $\mathbb{K} = \mathbb{C}$), è molto semplice esprimere il prodotto scalare (o hermitiano) in termini delle coordinate. Infatti se $x = F_{\mathcal{B}}(u), y = F_{\mathcal{B}}(v)$ sono le coordinate di vettori $u, v \in V$, allora

$$\langle u, v \rangle_V = \langle x, y \rangle_{\mathbb{K}^n}$$

dove $\langle \cdot, \cdot \rangle_{\mathbb{K}^n}$ è il prodotto scalare (o hermitiano) canonico di \mathbb{K}^n . Infatti, dato che la base \mathcal{B} è ortonormale, questo fatto è conseguenza immediata della Proposizione 1.1.

Proposizione 3.1: 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:

$$\langle T(v), w \rangle_W = \langle v, T^*(w) \rangle_V . \quad (3.1)$$

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

$$A^* = \begin{cases} A^T & \text{se } \mathbb{K} = \mathbb{R} \\ A^H = \overline{A}^T & \text{se } \mathbb{K} = \mathbb{C} \end{cases} . \quad (3.2)$$

Dimostrazione: Esistenza di T^* : Siano $\mathcal{B} = \{v_1, \dots, v_n\}$ e $\mathcal{C} = \{w_1, \dots, w_m\}$ rispettivamente basi ortonormali di V e W . Se A è la matrice associata a T relativa a \mathcal{B} e \mathcal{C} , vogliamo provare che l'applicazione $T^*: W \rightarrow V$ che, relativamente a \mathcal{C} e \mathcal{B} ha matrice $A^* = (a_{ij}^*)$ definita da (3.2) effettivamente verifica (3.1). Siano

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

Dunque $x = F_{\mathcal{B}}(v) \in \mathbb{K}^n$ e $y = F_{\mathcal{C}}(w) \in \mathbb{K}^m$. Allora, dato che $T(v)$ è l'unico vettore di W che ha coordinate Ax rispetto alla base \mathcal{C} , si ha:

$$\langle T(v), w \rangle_W = \langle Ax, y \rangle_{\mathbb{K}^m} = y^* Ax .$$

dove $y^* = y^T$ se $\mathbb{K} = \mathbb{R}$ e $y^* = y^H$ se $\mathbb{K} = \mathbb{C}$. D'altra parte dato che $T^*(w)$ è l'unico vettore di V che ha coordinate A^*y rispetto alla base \mathcal{B} , si ha:

$$\langle v, T^*(w) \rangle_V = \langle x, A^*y \rangle_{\mathbb{K}^n} = (A^*y)^* x = y^* (A^*)^* x = y^* Ax$$

e quindi la (3.1) è verificata.

Unicità di T^* : Se $S: V \rightarrow W$ è un'altra applicazione lineare che soddisfa (3.1), 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 l'Osservazione 1 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 3.1 si dice *aggiunta* di T .

Osservazione 3. 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 3.1

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 (3.2).

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 3.1. 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 hermitiano 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 3.1. Anche in questo caso è facile dimostrare direttamente che $L_A^* = L_{A^H}$ dato che per ogni $v \in \mathbb{C}^n$ e $w \in \mathbb{C}^m$

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

Prima di andare avanti diamo un'applicazione dell'esistenza dell'aggiunta per dimostrare un risultato che fornisce un isomorfismo canonico tra uno spazio vettoriale metrico e il suo duale:

Teorema 3.2: (di rappresentazione di Riesz.) *Sia V uno spazio vettoriale metrico e $\phi \in V^*$ un funzionale lineare su V . Allora esiste un unico vettore $v_\phi \in V$ tale che per ogni $v \in V$ si ha $\phi(v) = \langle v, v_\phi \rangle$. L'applicazione $\Phi: V^* \rightarrow V$ definita da $\Phi(\phi) = v_\phi$ è un isomorfismo.*

Dimostrazione: L'unicità di v_ϕ è immediata perché se ci fosse $v'_\phi \in V$ con la stessa proprietà, allora per ogni $v \in V$ si avrebbe $\langle v, v_\phi - v'_\phi \rangle = 0$ e quindi necessariamente $v_\phi = v'_\phi$. Si consideri su \mathbb{R} (\mathbb{C}) il prodotto scalare (hermitiano) canonico (\cdot, \cdot) . Allora per ogni $\phi \in V^*$ esiste l'aggiunta $\phi^*: \mathbb{R} \rightarrow V$ ($\phi^*: \mathbb{C} \rightarrow V$). Si definisca $v_\phi = \phi^*(1)$. Allora

$$\phi(v) = \phi(v) \cdot 1 = (\phi(v), 1) = \langle v, \phi^*(1) \rangle = \langle v, v_\phi \rangle.$$

È dunque ben definita un'applicazione $\Phi: V^* \rightarrow V$ ponendo $\Phi(\phi) = v_\phi$. L'applicazione Φ è lineare dato che dalla dimostrazione della Proposizione 3.1 segue immediatamente che l'aggiunta della somma di due applicazioni lineari è la somma delle aggiunte e l'aggiunta di un multiplo di un'applicazione lineare è il corrispondente multiplo dell'aggiunta. Dato che V e V^* hanno la stessa dimensione, per dimostrare che Φ è un isomorfismo, basta provare che è iniettiva. Ma questo è un fatto immediato: se $\Phi(\phi) = v_\phi = O$ allora per ogni $v \in V$ si ha

$$\phi(v) = \langle v, v_\phi \rangle = \langle v, O \rangle = 0,$$

dunque ϕ è l'applicazione nulla e quindi $\text{Ker}\Phi = \{O\}$: □

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 3.3: *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*

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

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 $v, w \in V$.

Proposizione 3.4: 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

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

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

$$\langle v, w - T^*(T(w)) \rangle = 0.$$

Ma allora

$$T^*(T(w)) = w \tag{3.3}$$

per ogni $w \in V$. Da (3.3) 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 (3.3) 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 $A^{-1} = A^*$. \square

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

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

ossia se e solo se $\{A^1, \dots, A^n\}$ 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. Sia A una matrice ortogonale (unitaria). Allora $|\det(A)| = 1$.

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 $T: V \rightarrow V$ un endomorfismo e siano A la matrice di T relativa a \mathcal{B} e A' la matrici di T relativa \mathcal{C} . Allora

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

Ricordiamo che due matrici A, A' reali quadrate di ordine n si dicono *congruenti* se esiste una matrice invertibile B tale che $A' = B^T A B$. Dunque matrici di uno stesso endomorfismo di uno spazio vettoriale metrico reale relative a basi ortonormali sono congruenti.

5. Appendice: Isometrie per la distanza euclidea.

Si consideri su \mathbb{R}^n la *distanza euclidea* $d: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ definita per $x, y \in \mathbb{R}^n$ da:

$$d(x, y) = \|x - y\| = \sqrt{\sum_{j=1}^n (x_j - y_j)^2}.$$

Un'*isometria per la distanza euclidea* o *movimento rigido* è un'applicazione biettiva $F: \mathbb{R}^n \rightarrow \mathbb{R}^n$ tale che

$$d(F(x), F(y)) = d(x, y) \quad \text{per ogni } x, y \in \mathbb{R}^n.$$

Esempi. Ci sono due classi di esempi ovvi. La prima è data dalle traslazioni. Dato $b \in \mathbb{R}^n$ la traslazione $t_b: \mathbb{R}^n \rightarrow \mathbb{R}^n$ mediante b è definita da $t_b(x) = x + b$. Per ogni $b \in \mathbb{R}^n$ la traslazione t_b è un'*isometria per la distanza euclidea*. Un'altra classe ovvia di isometrie per la distanza euclidea sono definite dalle isometrie lineari $L_A: \mathbb{R}^n \rightarrow \mathbb{R}^n$, dove A è una matrice ortogonale, che abbiamo studiato nel paragrafo precedente. Dato che è immediato osservare che la composizione di due isometrie per la distanza euclidea è un'*isometria per la distanza euclidea*, allora per ogni $b \in \mathbb{R}^n$ e matrice ortogonale A l'applicazione $\mathcal{I}_{b,A} = t_b \circ L_A$ definita da $\mathcal{I}_{b,A}(x) = Ax + b$ per $x \in \mathbb{R}^n$, è un'*isometria per la distanza euclidea*.

Ci proponiamo ora di dimostrare che tutte le isometrie per la distanza euclidea sono di questo tipo:

Teorema 5.1: *Un'applicazione $F: \mathbb{R}^n \rightarrow \mathbb{R}^n$ è un'*isometria per la distanza euclidea* se e solo se per qualche $b \in \mathbb{R}^n$ e matrice ortogonale A si ha $F = \mathcal{I}_{b,A}$.*

Dimostrazione: Dobbiamo solo dimostrare che se $F: \mathbb{R}^n \rightarrow \mathbb{R}^n$ è un'*isometria per la distanza euclidea* allora $F = \mathcal{I}_{b,A}$ per qualche $b \in \mathbb{R}^n$ e matrice ortogonale A . Cominciamo esaminando il caso in cui $F(0) = 0$. In questo caso dimostreremo che $F = L_A$ per un'opportuna matrice ortogonale A . A tal fine bisogna dimostrare che F è lineare e per ogni $x, y \in \mathbb{R}^n$, si deve avere

$$\langle x, y \rangle = \langle F(x), F(y) \rangle \quad (5.1)$$

dove, come al solito \langle, \rangle denota il prodotto scalare canonico. Per provare (5.1) si osservi che dato che F è un'*isometria per la distanza euclidea*, si ha:

$$\begin{aligned} \|F(x)\|^2 &= (d(0, F(x)))^2 = (d(0, x))^2 = \|x\|^2, \\ \|F(y)\|^2 &= (d(0, F(y)))^2 = (d(0, y))^2 = \|y\|^2, \\ \|F(x) - F(y)\|^2 &= (d(F(x), F(y)))^2 = (d(x, y))^2 = \|x - y\|^2. \end{aligned}$$

Dunque

$$\begin{aligned} \|x\|^2 + \|y\|^2 - 2\langle x, y \rangle &= \langle x - y, x - y \rangle \\ &= \|x - y\|^2 = \|F(x) - F(y)\|^2 \\ &= \langle F(x) - F(y), F(x) - F(y) \rangle \\ &= \|F(x)\|^2 + \|F(y)\|^2 - 2\langle F(x), F(y) \rangle \end{aligned}$$

da cui segue immediatamente $\langle x, y \rangle = \langle F(x), F(y) \rangle$. Se ora dimostriamo che F è lineare allora segue che necessariamente deve essere $F = L_A$ per un'opportuna matrice ortogonale A . Si osservi che $F(x + y) = F(x) + F(y)$ se e solo se $\|F(x + y) - F(x) - F(y)\|^2 = 0$. Basta allora calcolare:

$$\begin{aligned} \|F(x + y) - F(x) - F(y)\|^2 &= \langle F(x + y) - F(x) - F(y), F(x + y) - F(x) - F(y) \rangle \\ &= \langle F(x + y), F(x + y) \rangle + \langle F(x), F(x) \rangle + \langle F(y), F(y) \rangle \\ &\quad - 2\langle F(x + y), F(x) \rangle - 2\langle F(x + y), F(y) \rangle + 2\langle F(x), F(y) \rangle \end{aligned}$$

e, usando (5.1), allora:

$$\begin{aligned} \|F(x + y) - F(x) - F(y)\|^2 &= \langle x + y, x + y \rangle + \langle x, x \rangle + \langle y, y \rangle \\ &\quad - 2\langle x + y, x \rangle - 2\langle x + y, y \rangle + 2\langle x, y \rangle = 0. \end{aligned}$$

In modo del tutto analogo, se $t \in \mathbb{R}$ e $x \in \mathbb{R}^n$, allora

$$\begin{aligned}\|F(tx) - tF(x)\|^2 &= \langle F(tx) - tF(x), F(tx) - tF(x) \rangle \\ &= \|F(tx)\|^2 + t^2\|F(x)\|^2 - 2t \langle F(tx), F(x) \rangle \\ &= \|tx\|^2 + t^2\|x\|^2 - 2t \langle tx, x \rangle = 0.\end{aligned}$$

Rimane da dimostrare ora il caso generale, senza assumere $F(0) = 0$. Sia allora $F: \mathbb{R}^n \rightarrow \mathbb{R}^n$ un'isometria per la distanza euclidea arbitraria e sia $F(0) = b$. Se $G: \mathbb{R}^n \rightarrow \mathbb{R}^n$ è definita da $G(x) = F(x) - b$, allora è immediato verificare che G è un'isometria per la distanza euclidea tale che $G(0) = F(0) - b = 0$. Per quanto dimostrato prima, esiste una matrice ortogonale A tale che $G = L_A$ e quindi $F = G + b = L_A + b$. \square