Capitolo 10

Prodotti scalari reali e complessi.

10.1 Definizione di prodotto scalare

La nozione di prodotto scalare in uno spazio vettoriale su $\mathbb{R}$ o su $\mathbb{C}$ è alla base di tutte le nozioni di carattere metrico, in particolare la lunghezza (norma), angolo, ortogonalità. La definizione che daremo è diversa nei due casi reale e complesso. In entrambi i casi c'è un esempio fondamentale di prodotto scalare, detto canonico o standard, rispettivamente in $\mathbb{R}^n$ e in $\mathbb{C}^n$.

Esempi 10.1.1. Il prodotto scalare standard in $\mathbb{R}^n$ è l'applicazione $\langle \cdot, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ tale che $\langle v, w \rangle = a_1 b_1 + \cdots + a_n b_n$, dove $v = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$ e $w = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$.

Osserviamo che si può scrivere in maniera concisa $\langle v, w \rangle = {}^t v \cdot w$, dove ora v, w denotano le relative n-uple scritte in colonna.

Il prodotto scalare standard in $\mathbb{C}^n$ è l'applicazione $\langle \cdot, \cdot \rangle : \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$ tale che $\langle z, w \rangle = \bar{z_1} w_1 + \cdots + \bar{z_n} w_n$, dove $z = \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix}$ e $w = \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix}$.

Anche questa volta si può scrivere in maniera concisa $\langle z, w \rangle = {}^t \bar{z} \cdot y$, dove z, w denotano le relative n-uple scritte in colonna, e dove $\bar{z}$ denota il vettore che ha come componenti i complesso coniugato delle componenti di z.

Daremo ora la definizione generale di prodotto scalare nei due casi. Premettiamo un richiamo sul coniugio nel campo complesso.

Il **coniugio** è l'applicazione $J : \mathbb{C} \to \mathbb{C}$, tale che $J(z) = \bar{z}$, dove se $z = a + bi$, il coniugato è $\bar{z} = a - bi$. Si ha $z = \bar{z}$ se e solo se $z \in \mathbb{R}$. Valgono le proprietà:

$$\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}, \quad \overline{\lambda z} = \bar{\lambda}\bar{z},$$

ovvero

$$J(\lambda z) = \bar{\lambda} J(z).$$

Dunque J non è un endomorfismo lineare di $\mathbb{C}$ come $\mathbb{C}$-spazio vettoriale, infatti conserva la somma ma non il prodotto per scalari; è invece un'applicazione $\mathbb{R}$-lineare. La proprietà $J(\lambda z) = \bar{\lambda} J(z)$ si esprime dicendo che J è anti-lineare.

Ricordiamo inoltre che:

- $z + \bar{z} = 2\Re(z)$, dove $\Re(z)$ denota la parte reale di z;
- $z - \bar{z} = 2i\Im(z)$, dove $\Im(z)$ denota la parte immaginaria di z;
- $z\bar{z} = (a + bi)(a - bi) = a^2 + b^2 =\mid z \mid^2$, il modulo di z al quadrato.

Definizione 10.1.2. Sia V un $\mathbb{R}$-spazio vettoriale. Un **prodotto scalare** in V è un'applicazione

$$b : V \times V \to \mathbb{R}$$

verificante le proprietà:

1. b è bilineare, ovvero lineare in ciascuno dei due argomenti;
2. b è simmetrica, ovvero $b(v, v') = b(v', v)$, per ogni $v, v' \in V$;
3. b è definita positiva, ovvero $b(v, v) \geq 0$ per ogni $v \in V$, e inoltre $b(v, v) = 0$ se e solo se $v = 0$.

Le tre propretà si riassumono dicendo che b è una **forma bilineare simmetrica definita positiva**.

Definizione 10.1.3. Sia V un $\mathbb{C}$-spazio vettoriale. **Un prodotto scalare in** V è un'applicazione

$$b : V \times V \to \mathbb{C}$$

verificante le proprietà:

1. b è sesquilineare, ovvero lineare nel secondo argomento ma antilineare nel primo;

2. b è hermitiana, ovvero $b(w, z) = \overline{b(z, w)}$, per ogni $z, w \in V$;

3. b è definita positiva, ovvero $b(z, z) \geq 0$ per ogni $z \in V$, e inoltre $b(z, z) = 0$ se e solo se $z = 0$.

Osserviamo che l'ultima condizione ha senso, perché dalla seconda segue che $b(z, z) \in \mathbb{R}$.

Le tre propretà si riassumono dicendo che b è una **forma sesquilineare hermitiana definita positiva**.

Osservazione 42. I prodotti scalari standard su $\mathbb{R}^n$ e su $\mathbb{C}^n$ soddisfano le tre condizioni richieste per essere un prodotto scalare.

Osservazione 43. Le tre proprietà che si richiedono a un prodotto scalare complesso si particolarizzano a quelle dei prodotti scalari reali quando si lavora con elementi reali, in quanto il coniugato di un numero reale x è x stesso.

Definizione 10.1.4. Uno **spazio vettoriale euclideo** è uno spazio vettoriale reale in cui è stato fissato un prodotto scalare reale. Uno **spazio vettoriale unitario** è uno spazio vettoriale complesso in cui è stato fissato un prodotto scalare complesso.

10.2 Rappresentazione matriciale

Le forme bilineari sugli spazi vettoriali su $\mathbb{R}$ e quelle sesquilineari sugli spazi vettoriali su $\mathbb{C}$ si possono rappresentare in coordinate mediante opportune matrici, una volta fissata una base.

Sia dunque $b : V \times V \to \mathbb{K}$ una tale forma, con $\mathbb{K} = \mathbb{R}$ o $\mathbb{C}$, sia $\mathcal{B} = (v_1, \ldots, v_n)$ una base di V.

Definizione 10.2.1. La matrice di b rispetto a $\mathcal{B}$ è la matrice $n \times n$ data da

$$M_{\mathcal{B}}(b) = (b(v_i, v_j))_{i,j=1,\ldots,n};$$

al posto di indici i, j c'è lo scalare $b(v_i, v_j)$.

Siano v, w due vettori di V, di coordinate rispetto a $\mathcal{B}$ rispettivamente $x_1, \ldots, x_n$, e $y_1, \ldots, y_n$.

Denotiamo con x la colonna delle coordinate x_i, e con y quella delle y_j. Allora si ha:

$$\text{caso reale}: \; b(v, w) = b(\sum_{i=1}^{n} x_i v_i, \sum_{j=1}^{n} y_j v_j) = \text{ bilinearità } = \sum_{i,j=1}^{n} x_i y_j b(v_i, v_j) = \sum_{i,j=1}^{n} x_i b(v_i, v_j) y_j,$$

caso complesso : $b(v,w) = b(\sum_{i=1}^{n} x_i v_i, \sum_{j=1}^{n} y_j v_j) =$ sesquilinearità $= \sum_{i,j=1}^{n} \overline{x_i} b(v_i, v_j) y_j.$

Allora, ponendo $M = M_{\mathcal{B}}(b)$, si ha nel caso reale

$$ {}^t x \cdot M \cdot y = (x_1, \dots, x_n) \cdot M \cdot \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = (\sum_{i=1}^{n} x_i b(v_i, v_1), \dots, \sum_{i=1}^{n} x_i b(v_i, v_n)) \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = b(v,w), $$

e nel caso complesso

$$ {}^t \bar{x} \cdot M \cdot y = (\bar{x_1}, \dots, \bar{x_n}) \cdot M \cdot \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = (\sum_{i=1}^{n} \bar{x}_i b(v_i, v_1), \dots, \sum_{i=1}^{n} \bar{x}_i b(v_i, v_n)) \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = b(v,w). $$

Abbiamo dunque $b(v,w) = {}^t\bar{x} \cdot M_{\mathcal{B}}(b) \cdot y$, dove x, y sono i vettori colonna delle coordinate di v, w rispetto a $\mathcal{B}$. Nel caso reale questo si particolarizza in $b(v,w) = {}^t x \cdot M_{\mathcal{B}}(b) \cdot y$.

Osserviamo che $b(v_i, v_j) = {}^t\bar{e_i} \cdot M \cdot e_j$, perché le coordinate dei vettori di base sono date dai vettori della base canonica.

Ricordiamo che una matrice quadrata A è detta simmetrica se $A = {}^tA$, coincide con la sua trasposta.

Definizione 10.2.2. Una matrice quadrata complessa A è detta **hermitiana** se ${}^tA = \bar{A}$, cioè la trasposta di A coincide con la coniugata di A, ottenuta coniugando tutti gli elementi di A.

Sulla diagonale principale di una matrice hermitiana tutti gli elementi sono numeri reali.

Se la forma b considerata è simmetrica reale (rispettivamente hermitiana complessa), la matrice $M_{\mathcal{B}}(b)$ risulta simmetrica, cioè $M_{\mathcal{B}}(b) = {}^tM_{\mathcal{B}}(b)$ (rispettivamente hermitiana, cioè ${}^tM_{\mathcal{B}}(b) = \overline{M_{\mathcal{B}}(b)}$). Infatti $b(v_j, v_i) = b(v_i, v_j)$ (rispettivamente $b(v_j, v_i) = \overline{b(v_i, v_j)}$).

Abbiamo visto come, data una base di V, a ogni forma bilineare simmetrica (risp. sesquilineare hermitiana) si può associare una matrice. Vedremo ora che il procedimento si può invertire.

Proposizione 10.2.3. *Sia A una matrice quadrata $n \times n$.*

1. *A è simmetrica reale se e solo se l'applicazione $b : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ definita da $b(x,y) = {}^txAy$ è una forma bilineare simmetrica.*

2. *A è hermitiana complessa se e solo se l'applicazione $b : \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$ definita da $b(x,y) = {}^t\bar{x}Ay$ è una forma sesquilineare hermitiana.*

Dimostrazione. Un'implicazione è già stata vista. Osserviamo inoltre che il caso 2. è un caso particolare del caso 1., dunque supponiamo che A sia hermitiana e definiamo b ponendo $b(x,y) = {}^t\bar{x}Ay$. Verifichiamo la linearità di b nel secondo argomento:

$$b(x, \lambda y + \mu y') = {}^t\bar{x}A(\lambda y + \mu y') = \text{ proprietà distributiva} = {}^t\bar{x}A(\lambda y) + {}^t\bar{x}A(\mu y') =$$

$$= \text{proprietà del prodotto di matrici} = \lambda({}^t\bar{x}Ay) + \mu({}^t\bar{x}Ay') = \lambda b(x,y) + \mu b(x,y').$$

Verifichiamo l'antilinearità nel primo argomento:

$$b(\lambda x + \mu x', y) = {}^t(\overline{\lambda x + \mu x'})Ay = \text{proprietà del coniugio} = {}^t(\bar{\lambda}\bar{x} + \bar{\mu}\bar{x}')Ay =$$

$$= \text{proprietà del prodotto di matrici} = \bar{\lambda}{}^t\bar{x}Ay + \bar{\mu}{}^t\bar{x}'Ay = \bar{\lambda}b(x,y) + \bar{\mu}b(x',y).$$

Verifichiamo l'hermitianità: $b(y,x) = {}^t\bar{y}Ax =$ una matrice 1×1 coincide con la sua trasposta $= {}^t({}^t\bar{y}Ax) = {}^tx{}^tA\bar{y} = \overline{{}^t\bar{x}{}^t\bar{A}y} = A$ è hermitiana $= \overline{{}^t\bar{x}Ay} = \overline{b(x,y)}$. □

Osservazione 44. Osserviamo che se $A = (a_{ij}), B = (b_{ij})$ sono matrici simmetriche reali (risp. hermitiane complesse) che definiscono la stessa forma b, ovvero se ${}^txAy = {}^txBy$ per ogni $x, y \in \mathbb{R}^n$ (risp. ${}^t\bar{x}Ay = {}^t\bar{x}By$ per ogni $x, y \in \mathbb{C}^n$) allora $A = B$. Infatti se come vettori x, y prendiamo due vettori della base canonica $\mathcal{E}$ di $\mathbb{R}^n$ (risp. di $\mathbb{C}^n$), abbiamo ${}^t\bar{e}_iAe_j = a_{ij} = {}^t\bar{e}_iBe_j = b_{ij}$.

Si ha così una biiezione fra forme bilineari simmetriche su $\mathbb{R}^n$ e matrici simmetriche reali di ordine n (risp. tra forme sesquilineari hermitiane su $\mathbb{C}^n$ e matrici hermitiane complesse di ordine n).

Una biiezione analoga si ha su uno spazio vettoriale V di dimensione n, reale o complesso, se è stata fissata una sua base.

Esempi 10.2.4. 1. Se $\mathcal{E}$ è la base canonica di $\mathbb{R}^n$ (risp. di $\mathbb{C}^n$) e b è il prodotto scalare standard, si ha $b(e_i, e_j) = \delta_{ij}$ e dunque $M_{\mathcal{E}}(b) = \mathbb{I}_n$.

2. Sia $A = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix}$ matrice simmetrica reale di ordine 3. La forma bilineare simmetrica corrispondente rispetto a $\mathcal{E}$ su $\mathbb{R}^3$ è tale che $b(e_1, e_1) = 2$, $b(e_1, e_2) = 1 = b(e_2, e_1)$, $b(e_1, e_3) = 0 = b(e_3, e_1)$, $b(e_2, e_2) = 1$, $b(e_2, e_3) = 0 = b(e_3, e_2)$, $b(e_3, e_3) = 3$. Dunque

$$b(x,y) = \sum_{ij=1}^{3} x_i y_j b(e_i, e_j) = 2x_1y_1 + x_1y_2 + x_2y_1 + x_2y_2 + 3x_3y_3.$$

Per vedere se b è o meno un prodotto scalare consideriamo $b(x,x)$ e vediamo se risulta sempre positivo o nullo, e se si annulla solo nella terna nulla.

$$b(x,x) = 2x_1^2 + 2x_1x_2 + x_2^2 + 3x_3^2 = x_1^2 + (x_1^2 + 2x_1x_2 + x_2^2) + 3x_3^2 = x_1^2 + (x_1+x_2)^2 + 3x_3^2;$$

essendo una somma di quadrati è sempre ≥ 0, e si annulla solo se

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

se e solo se $x_1 = x_2 = x_3 = 0$. Quindi b è un prodotto scalare.

3. Sia $S = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. La forma bilineare simmetrica corrispondente su $\mathbb{R}^3$ è

$$b(x,y) = {}^t xSy = x_1y_1 + x_1y_2 + x_2y_1 + x_2y_2 + x_3y_3.$$

$$b(x,x) = {}^t xSx = x_1^2 + 2x_1x_2 + x_2^2 + x_3^2 = (x_1+x_2)^2 + x_3^2 \geq 0$$

per ogni $x \in \mathbb{R}^3$; inoltre $b(x,x) = 0$ se e solo se

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

Questa volta dunque $b(x,x) = 0$ per ogni $x \in \langle (1,-1,0) \rangle$, dunque b è una forma bilineare simmetrica non definita positiva, non è un prodotto scalare.

4. Sia $B = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. B definisce $b(x,y) = x_1y_2 + x_2y_1 + x_3y_3$, $b(x,x) = 2x_1x_2 + x_3^2$: non è sempre ≥ 0. Per esempio $b(e_1, e_1) = 0$, $b(\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}) = -2$. Non è definita positiva.

5. Sia $H = \begin{pmatrix} 2 & 1-i \\ 1+i & 2 \end{pmatrix}$, matrice hermitiana 2×2. La forma sesquilineare hermitiana h associata ad H rispetto alla base canonica di $\mathbb{C}^2$ opera come segue su $z = (z_1, z_2)$, $w = (w_1, w_2)$:

$$h(z,w) = (\bar{z}_1 \ \bar{z}_2) H \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = 2\bar{z}_1w_1 + (1-i)\bar{z}_1w_2 + (1+i)\bar{z}_2w_1 + 2\bar{z}_2w_2.$$

In particolare

$$h(z,z) = 2\bar{z}_1 z_1 + (1-i)\bar{z}_1 z_2 + (1+i)\bar{z}_2 z_1 + 2\bar{z}_2 z_2 = 2|z_1|^2 + 2|z_2|^2 + 2\Re(\bar{z}_1 z_2) + 2\Im(\bar{z}_1 z_2).$$

Scriviamo ora $z_1 = a_1 + b_1 i$, $z_2 = a_2 + b_2 i$, e sostituiamo:

$$h(z,z) = 2(a_1^2+b_1^2+a_2^2+b_2^2+(a_1a_2+b_1b_2)+(a_1b_2-a_2b_1)) = (a_1+a_2)^2+(b_1+b_2)^2+(a_1+b_2)^2+(a_2-b_1)^2.$$

Chiaramente questa somma è sempre ≥ 0 e si annulla se e solo se $a_1+a_2 = b_1+b_2 = a_1+b_2 = a_2-b_1 = 0$, che equivale a $a_1 = b_1 = a_2 = b_2 = 0$ ovvero a $z = (z_1, z_2) = 0$. Quindi h è definita positiva, e risulta essere un prodotto scalare su $\mathbb{C}^2$.

6. Un esempio di carattere diverso di prodotto scalare su $V = \mathbb{R}[t]$ è definito ponendo

$$\langle p, q\rangle = \int_0^1 p(t)q(t)\mathrm{d}t.$$

E` una forma bilineare simmetrica per le proprietà dell'integrale; è definita positiva perché

$$\langle p, p\rangle = \int_0^1 p^2(t)\mathrm{d}t \geq 0$$

per ogni polinomio $p(t)$ e si annulla solo se p è il polinomio nullo.

7. Sia $V = M_{m,n}(\mathbb{R})$. Definiamo $b : V\times V \to \mathbb{R}$ ponendo $b(A,B) = \mathrm{tr}({}^tB\cdot A)$, la traccia della matrice prodotto. E` immediato verificare che b è bilineare e simmetrica; inoltre $b(A,A) = \mathrm{tr}({}^tAA)$. La matrice ${}^tAA$ al posto d'indici ii ha $(a_{1i}\dots a_{ni})\begin{pmatrix} a_{1i} \\ \vdots \\ a_{ni}\end{pmatrix} = a_{1i}^2 + a_{2i}^2 + \cdots + a_{ni}^2$.

Dunque $b(A,A)$ è la somma dei quadrati di tutti gli elementi della matrice A, perciò b è un prodotto scalare.

10.3 Norma

Sia V un $\mathbb{K}$-spazio vettoriale, con $\mathbb{K} = \mathbb{R}$ o $\mathbb{K} = \mathbb{C}$.

Definizione 10.3.1. Un'applicazione $N : V \to \mathbb{R}$ è una **norma** su V se:

1. $N(\lambda v) = |\lambda| N(v)$, per ogni $v \in V$, dove λ denota il valore assoluto di λ nel caso reale e il modulo di λ nel caso complesso;

2. $N(v+w) \leq N(v)+N(w)$, per ogni $v, w \in V$; questa è detta **disuguaglianza triangolare**;

3. $N(v) = 0$ se e solo se $v = 0$.

Osserviamo che dalla definizione segue subito che $N(v) \geq 0$ per ogni vettore v; infatti $N(v-v) = N(0) = 0$ per la 3.; ma $0 = N(v-v) \leq N(v)+N(-v) = N(v)+|-1|N(v) = 2N(v)$; perciò $N(v) \geq 0$.

Un'applicazione norma è spesso indicata con il simbolo $\|\cdot\|$. In tal caso le tre proprietà si scrivono $\|\lambda v\| = |\lambda|\|v\|$, $\|v+w\| \leq \|v\| + \|w\|$, $\|v\| = 0$ se e solo se $v = 0$.

Esempi 10.3.2. Esempi in $\mathbb{R}^n$.

1. Norma euclidea: $\|(x_1, \dots, x_n)\| = \sqrt{x_1^2 + \cdots + x_n^2}$. Le proprietà 1. e 2. di una norma sono immediate, la disuguaglianza triangolare sarà dimostrata in seguito.

2. Norma 1: $\|(x_1, \dots, x_n)\|_1 = |x_1| + \cdots + |x_n|$, somma dei valori assoluti. Per dimostrare la disuguaglianza traingolare si usa il fatto che in $\mathbb{R}$ $|x+y| \leq |x| + |y|$.

3. Norma ∞: $\|(x_1, \dots, x_n)\|_\infty = \max_{i=1,\dots,n}\{|x_i|\}$: massimo dei valori assoluti delle componenti.

Esempio in $\mathbb{C}^n$, norma standard: $\|(z_1, \dots, z_n)\| = \sqrt{\bar{z}_1 z_1 + \cdots + \bar{z}_n z_n}$.

Vedremo ora che a ogni prodotto scalare si può associare una norma. Denoteremo spesso un prodotto scalare con il simbolo $\langle \cdot, \cdot \rangle$ anzichè con b.

Proposizione 10.3.3. *Sia $\langle \cdot, \cdot \rangle$ un prodotto scalare su V, spazio vettoriale reale o complesso. Allora l'applicazione $\|\cdot\| : V \to \mathbb{R}$ tale che $\|v\| = \sqrt{\langle v, v \rangle}$ è una norma su V.*

Osserviamo che la definizione è ben posta perché il prodotto scalare su V è definito positivo.

Dimostrazione. La 3. vale perché il prodotto scalare è definito positivo. Verifichiamo la 1. nel caso complesso:

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

La dimostrazione della 2. sarà data dopo il Teorema 10.4.1. □

Si può dimostrare che non ogni funzione norma su uno spazio vettoriale proviene da un prodotto scalare, ma che, perché lo sia, occorre che sia verificata per ogni coppia di vettori v, w anche la " Legge del parallelogramma": $\|v+w\|^2 + \|v-w\|^2 = 2\|v\|^2 + 2\|w\|^2$.

10.4 Disuguaglianza di Cauchy-Schwarz e sue conseguenze

Teorema 10.4.1 (Disuguaglianza di Cauchy-Schwarz)**.** *Sia* $(V, \langle\cdot,\cdot\rangle)$ *uno spazio vettoriale euclideo o unitario. Allora per ogni coppia di vettori* $v, w \in V$ *vale la disuguaglianza*

$$|\langle v, w\rangle| \leq \|v\|\|w\| \tag{10.1}$$

dove $|\cdot|$ *denota il valore assoluto nel caso reale e il modulo nel caso complesso. Inoltre vale l'uguaglianza in (10.1) se e solo se* v, w *sono linearmente dipendenti.*

Dimostrazione. Vediamo la dimostrazione nel caso complesso, quello reale è un caso particolare.

Se $w = 0$, $|\langle v, 0\rangle| = 0 = \|v\|\|0\|$, quindi vale la (10.1) con segno di uguaglianza.

Se $w \neq 0$, consideriamo lo scalare

$$\lambda := \frac{\overline{\langle v, w\rangle}}{\langle w, w\rangle}.$$

Si ha:

$$0 \leq \langle v - \lambda w, v - \lambda w\rangle = \langle v, v\rangle - \lambda\langle v, w\rangle - \bar{\lambda}\langle w, v\rangle + \bar{\lambda}\lambda\langle w, w\rangle =$$

sostituiamo il valore di λ

$$= \langle v, v\rangle - \frac{\overline{\langle v, w\rangle}}{\langle w, w\rangle}\langle v, w\rangle - \frac{\langle v, w\rangle}{\langle w, w\rangle}\overline{\langle v, w\rangle} + \frac{\langle v, w\rangle}{\langle w, w\rangle}\frac{\overline{\langle v, w\rangle}}{\langle w, w\rangle}\langle w, w\rangle =$$

abbiamo usato che il prodotto scalare è hermitiano e che $\langle w, w\rangle$ è reale; gli ultimi due addendi si cancellano:

$$= \|v\|^2 - \frac{|\langle v, w\rangle|^2}{\|w\|^2}.$$

Moltiplichiamo per $\|w\|^2$, che è positivo, e otteniamo

$$0 \leq \|v\|^2\|w\|^2 - |\langle v, w\rangle|^2$$

ovvero

$$|\langle v, w\rangle|^2 \leq \|v\|^2\|w\|^2.$$

Estraendo le radici otteniamo

$$|\langle v, w\rangle| \leq \|v\|\|w\|,$$

che è la disuguaglianza voluta.

Se in (10.1) vale l'uguaglianza, o $w = 0$, oppure $\langle v - \lambda w, v - \lambda w\rangle = 0$, dunque $v - \lambda w = 0$, e quindi v, w sono linearmente dipendenti. Viceversa, se v, w sono linearmente dipendenti, con $w \neq 0$, esiste uno scalare λ tale che $v = \lambda w$; perciò $\langle v, w\rangle = \langle \lambda w, w\rangle = \bar{\lambda}\langle w, w\rangle = \bar{\lambda}\|w\|^2$. Perciò $|\langle v, w\rangle| = |\lambda|\|w\|^2$, e $\|v\|\|w\| = \|\lambda w\|\|w\| = |\lambda|\|w\|\|w\|$; segue l'uguaglianza in (10.1). □

Corollario 10.4.2. *La norma associata a un prodotto scalare soddisfa la disuguaglianza triangolare.*

Dimostrazione.

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

(Ora usiamo il fatto che per un numero complesso z si ha $\Re(z) \leq |z|$; infatti, posto $z = a + bi$, $\Re(z) = a \leq |a| = \sqrt{a^2} \leq \sqrt{a^2 + b^2} = |z|$. Inoltre vale l'uguaglianza se e solo se $a = |a| = \sqrt{a^2 + b^2}$, cioè z è reale positivo.)

$$\leq \|v\|^2 + 2\,|\langle v, w\rangle| + \|w\|^2 \leq$$

per la disuguaglianza di Cauchy-Schwarz

$$\leq \|v\|^2 + 2\|v\|\|w\| + \|w\|^2 = (\|v\| + \|w\|)^2;$$

estraendo le radici quadrate da ambo i membri della disuguaglianza si ha la tesi. □

Osservazione 45. Osserviamo che nella disuguaglianza triangolare vale l'uguaglianza se e solo se sono entrambe uguaglianze le due disuguaglianze che compaiono nella dimostrazione; questo equivale al fatto che v, w siano linearmente dipendenti e che il loro prodotto scalare $\langle v, w\rangle$ sia reale positivo, ovvero vale una relazione del tipo $v = \lambda w$ con λ reale ≥ 0.

Proposizione 10.4.3 (Formule di polarizzazione)**.** *Una funzione norma può essere indotta da un unico prodotto scalare, cioè se si conosce come opera la funzione norma associata a un prodotto scalare, si può ricostruire $\langle v, w\rangle$ per ogni coppia di vettori v, w.*

Dimostrazione. Nella dimostrazione del Corollario 10.4.2, abbiamo visto che

$$\|v + w\|^2 = \|v\|^2 + 2\Re\langle v, w\rangle + \|w\|^2. \tag{10.2}$$

Nel caso reale ne segue la seguente relazione, detta formula di polarizzazione reale:

$$\langle v, w\rangle = \frac{1}{2}(\|v + w\|^2 - \|v\|^2 - \|w\|^2) \tag{10.3}$$

che esprime $\langle v, w\rangle$ utilizzando la norma dei vettori $v, w, v + w$.

Nel caso complesso, dalla (10.2) si ottiene soltanto

$$\Re\langle v, w\rangle = \frac{1}{2}(\|v + w\|^2 - \|v\|^2 - \|w\|^2); \tag{10.4}$$

per ricostruire $\langle v, w\rangle$ è quindi necessario esprimere anche $\Im\langle v, w\rangle$ in funzione delle norme di opportuni vettori. In effetti si ha:

$$\Im\langle v, w\rangle = \frac{1}{2}(\|v - iw\|^2 - \|v\|^2 - \|w\|^2). \tag{10.5}$$

La verifica è lasciata per esercizio. Le relazioni (10.4) e (10.5) costituiscono la formula di polarizzazione complessa. □

10.5 Angoli e ortogonalità

Sia V uno spazio vettoriale reale euclideo. La disuguaglianza di Cauchy-Schwarz ci dice che il valore assoluto di $\langle v, w\rangle$ è limitato superiormente da $\|v\|\|w\|$, per ogni coppia di vettori v, w. Ciò significa che

$$-\|v\|\|w\| \leq \ \langle v, w\rangle \ \leq \|v\|\|w\|.$$

Se v, w sono entrambi non nulli, le norme $\|v\|, \|w\|$ sono strettamente positive, e otteniamo

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

Ricordiamo che la funzione coseno è continua strettamente decrescente nell'intervallo chiuso $[0, \pi]$, e assume tutti i valori compresi fra -1 e 1. Perciò per ogni $y \in [-1, 1]$ esiste uno ed un solo $x \in [0, \pi]$ tale che $y = \cos x$. Povveromo allora definire l'angolo convesso di due vettori non nulli.

Definizione 10.5.1. Siano $v, w \in V$, $v \neq 0, w \neq 0$. **L'angolo convesso** di v e w è l'unico $\alpha \in [0, \pi]$ tale che

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

Osserviamo che la definizione data ha senso soltanto nel caso reale, in quanto nel caso complesso in generale $\langle v, w\rangle \in \mathbb{C}$. Tuttavia, in entrambi i casi, reale e complesso, si dà la definizione di ortogonalità.

Definizione 10.5.2. Sia V uno spazio vettoriale euclideo o unitario. Due vettori $v, w \in V$ si dicono **ortogonali** se e solo se $\langle v, w\rangle = 0$. Per indicare che v, w sono ortogonali si scrive anche $v \perp w$.

Nel caso euclideo due vettori non nulli sono ortogonali se e solo se il loro angolo convesso vale $\frac{\pi}{2}$.

Definizione 10.5.3. 1. Sia V uno spazio vettoriale euclideo o unitario, siano $U, W \subset V$ sottospazi vettoriali. U, W si dicono ortogonali se $\langle u, w\rangle = 0$ per ogni $u \in U$ e per ogni $w \in W$. Si scrive $U \perp W$.

2. Sia $S \subset V$ un sottinsieme. L'ortogonale di S è il sottospazio vettoriale di V

$$S^{\perp} = \{v \in V \mid \langle v, s\rangle = 0 \text{ per ogni } s \in S\}.$$

(Verificare che è un sottospazio per esercizio)

Se $W \subset V$ è un sottospazio vettoriale, $W^{\perp}$ è detto **complemento ortogonale di** W.

Se $w \in V$ e $W = \langle w\rangle$ è il sottospazio generato da w, verificare che si ha $w^{\perp} = W^{\perp}$.

Definizione 10.5.4. Una famiglia $\{v_i\}_{i\in I}$ di vettori di V si dice **famiglia ortogonale** se $v_i \perp v_j$ per ogni $i, j \in I$, $i \neq j$, cioè sono a due a due ortogonali.

Un vettore di norma 1 è detto un **versore**.

Una famiglia ortogonale di vettori è detta **ortonormale** se in più $\|v_i\| = 1$ per ogni $i \in I$. Ciò si può anche esprimere dicendo che $\langle v_i, v_j\rangle = \delta_{ij}$ per ogni $i, j \in I$.

Una famiglia ortonormale è dunque una famiglia ortogonale di versori. Una **base ortonormale** di V è una base che risulta una famiglia ortonormale.

Se $v \neq 0$ è un vettore non nullo, **normalizzare** v significa passare da v al vettore

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

che ha chiaramente norma 1.

Esempi 10.5.5. 1. Per il prodotto scalare standard di $\mathbb{R}^n$ e di $\mathbb{C}^n$ la base canonica è una base ortonormale.

2. In $\mathbb{R}^2$ e in $\mathbb{C}^2$ con prodotto scalare standard, i vettori $(1, 1)$, $(1, -1)$ sono ortogonali; normalizzandoli si ottiene $v_1 = (\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2})$, $v_2 = (\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2})$: v_1, v_2 formano una base ortonormale.

Proposizione 10.5.6. *Sia $\mathcal{B} = (v_1, \ldots, v_n)$ una base ortonormale di uno spazio vettoriale ortogonale o euclideo V. Allora ogni vettore $v \in V$ si scrive nella forma*

$$v = \langle v_1, v\rangle v_1 + \cdots + \langle v_n, v\rangle v_n. \tag{10.6}$$

Dimostrazione. Siano $\lambda_1, \ldots, \lambda_n$ le coordinate di v rispetto a $\mathcal{B}$: $v = \lambda_1 v_1 + \cdots + \lambda_n v_n$. Allora, per ogni $j = 1, \ldots, n$ si ha:

$$\begin{aligned}\langle v_j, v\rangle &= \langle v_j, \lambda_1 v_1 + \cdots + \lambda_n v_n\rangle = \text{per la linearità nel secondo argomento}\\ &= \lambda_1\langle v_j, v_1\rangle + \cdots + \lambda_n\langle v_j, v_n\rangle =\\ &= \sum_{i=1}^{n} \lambda_i \langle v_j, v_i\rangle =\\ &= \sum_{i=1}^{n} \lambda_i \delta_{ji} = \lambda_j.\end{aligned}$$

□

La Proposizione 10.5.6 mostra che lavorare con basi ortonormali presenta notevoli vantaggi. La coordinata i-esima di $v \neq 0$ rispetto a una tale base è dunque $\langle v_i, v\rangle = \cos\theta_i \|v\|$, dove θ_i è l'angolo convesso tra v_i e v. La relazione (10.6) si può dunque riscrivere nella forma

$$v = \|v\|(\cos\theta_1 v_1 + \cdots + \cos\theta_n v_n), \tag{10.7}$$

dove il vettore in parentesi risulta un versore e precisamente il normalizzato di v.

Proposizione 10.5.7. *Sia $\{v_i\}_{i\in I}$ una famiglia ortogonale con $v_i \neq 0$ per ogni i. Allora:*

1. *i vettori v_i sono linearmente indipendenti;*
2. *$\{\frac{v_i}{\|v_i\|}\}_{i\in I}$ è una famiglia ortonormale.*

Dimostrazione. 1. Consideriamo una combinazione lineare nulla dei v_i; con ciò intendiamo una combinazione lineare nulla di una qualunque sottofamiglia finita della famiglia di partenza. Ricordiamo infatti che non ha senso considerare somme infinite. Questa combinazione lineare nulla si può scrivere nella forma $\sum_{i\in I} \lambda_i v_i = 0$ con l'avvertenza che i coefficienti λ_i sono supposti **quasi tutti nulli**, cioè soltanto un numero finito di scalari λ_i può essere diverso da 0. Allora $\langle v_j, \sum_{i\in I} \lambda_i v_i\rangle = \langle v_j, 0\rangle = 0$, e d'altra parte $\langle v_j, \sum_{i\in I} \lambda_i v_i\rangle = \sum_{i\in I} \lambda_i \langle v_j, v_i\rangle = \lambda_j$. Perciò la combinazione lineare considerata è banale.

2. Come abbiamo già osservato $\frac{v_i}{\|v_i\|}$ ha norma 1. Inoltre, se $i \neq j$ si ha

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

□

10.6 Proiezioni ortogonali e ortonomalizzazione

Sia V uno spazio vettoriale euclideo o unitario, sia $W \subset V$ un suo sottospazio vettoriale. Supponiamo che $(w_1, \ldots, w_m)$ sia una base **ortonormale** di W.

Definizione 10.6.1. Dato un vettore $v \in V$, la sua **proiezione ortogonale su** W è il vettore $\tilde{v} := \langle w_1, v\rangle w_1 + \cdots + \langle w_m, v\rangle w_m$.

Osserviamo che:
1. $\tilde{v} \in W$ perché è combinazione lineare di $w_1, \ldots, w_m$;
2. $v - \tilde{v}$ è ortogonale a W; infatti verifichiamo che $v - \tilde{v}$ risulta ortogonale a w_i per ogni $i = 1, \ldots, m$ ed è quindi ortogonale a ogni vettore di W:

$$\begin{aligned}\langle w_i, v - \tilde{v}\rangle &= \langle w_i, v\rangle - \langle w_i, \sum_{j=1}^{m} \langle w_j, v\rangle w_j\rangle = \text{ linearità nel secondo argomento}\\ &= \langle w_i, v\rangle - \sum_{j=1}^{m} \langle w_j, v\rangle \langle w_i, w_j\rangle = \\ &= \langle w_i, v\rangle - \sum_{j=1}^{m} \langle w_j, v\rangle \delta_{ij} = 0.\end{aligned}$$

Vedremo ora che basi ortonormali esistono in ogni spazio vettoriale euclideo o unitario di dimensione finita.

Teorema 10.6.2 (Teorema di ortonormalizzazione di Gram-Schmidt)**.** *Sia V uno spazio vettoriale euclideo o unitario di dimensione finita n; sia $W \subset V$ un suo sottospazio vettoriale di dimensione m. Allora ogni base ortonormale $(w_1, \ldots, w_m)$ di W si può prolungare a una base ortonormale di V.*

Dimostrazione. Per induzione su $n - m$, la codimensione di W in V. Se $n - m = 0$, risulta $W = V$ e la tesi del teorema è automaticamente verificata.

Sia dunque $n - m > 0$ e supponiamo vera la tesi del teorema per $n - m - 1$; ciò significa che ogni base ortonormale di un sottospazio vettoriale di dimensione $m + 1$ di V si può prolungare a una base ortonormale di V.

Essendo $n - m > 0$, $W \subsetneq V$ e dunque esiste un vettore $v \in V \setminus W$. Consideriamo la sua proiezione ortogonale su W: $\tilde{v} := \langle w_1, v\rangle w_1 + \cdots + \langle w_m, v\rangle w_m$. Abbiamo che $v - \tilde{v}$ è ortogonale a W, e $v - \tilde{v} \neq 0$, altrimenti $v = \tilde{v} \in W$. Allora lo povveromo normalizzare e definire $w_{m+1} := \frac{v-\tilde{v}}{\|v-\tilde{v}\|}$. Otteniamo che $\|w_{m+1}\| = 1$ e $w_{m+1} \perp w_i$ per ogni $i = 1, \ldots, m$, dunque $w_1, \ldots, w_m, w_{m+1}$ formano una famiglia ortonormale di vettori. Consideriamo allora il sottospazio $W' = \langle w_1, \ldots, w_m, w_{m+1}\rangle$: $w_1, \ldots, w_{m+1}$ è una sua base ortonormale, e $\dim W' =$

$m+1$. Allora $\dim V - \dim W' = n - m - 1$ e povveromo applicare l'ipotesi induttiva, quindi la base ortonormale $w_1, \ldots, w_{m+1}$ si può prolungare a una base ortonormale di V. Ciò conclude la dimostrazione, perché la base ortonormale di V appena ottenuta prolunga anche la base di W $(w_1, \ldots, w_m)$ fissata in partenza. $\square$

Corollario 10.6.3. *Ogni spazio vettoriale euclideo o unitario di dimensione finita $n > 0$ possiede una base ortonormale.*

Dimostrazione. Infatti preso in V un vettore non nullo v, povveromo normalizzarlo e porre $w_1 = \frac{v}{\|v\|}$. Consideriamo ora $W = \langle w_1 \rangle$ e applichiamo il Teorema 10.6.2: la sua base ortonormale di W costituita da w_1 può essere prolungata a una base ortonormale di V. $\square$

Osservazione 46. Sia $(V, \langle \cdot, \cdot \rangle)$ uno spazio vettoriale euclideo o unitario, e $\mathcal{B} = (v_1, \ldots, v_n)$ una sua base ortonormale. La matrice $M_{\mathcal{B}}(\langle , \rangle)$ del prodotto scalare rispetto a $\mathcal{B}$ risulta la matrice identica $\mathbb{I}_n$, infatti al posto di indici ij contiene $\langle v_i, v_j \rangle = \delta_{ij}$. Quindi l'espressione analitica del prodotto scalare, in coordinate rispetto a una base ortonormale, coincide con quella del prodotto scalare canonico di $\mathbb{R}^n$ o di $\mathbb{C}^n$ rispettivamente.

Definizione 10.6.4. Sia V uno spazio vettoriale euclideo o unitario. V è detto **somma ortogonale** dei suoi sottospazi vettoriali $V_1, \ldots, V_k$, e si scrive $V = V_1 \bigcirc\!\!\!\!\perp V_2 \bigcirc\!\!\!\!\perp \cdots \bigcirc\!\!\!\!\perp V_k$, se

1. $V = V_1 + \cdots + V_k$;
2. $V_i \perp V_j$ per ogni $i \neq j$.

Osservazione 47. Osserviamo che dalla condizione 2. segue che $V_i \cap V_j = (0)$ se $i \neq j$, in quanto se $v \in V_i \cap V_j$ dev'essere $\langle v, v \rangle = 0$ e dunque $v = 0$. Non solo, ma segue anche $V_i \perp (\sum_{j \neq i} V_j)$, e perciò $V_i \cap (\sum_{j \neq i} V_j) = (0)$. Concludiamo che **ogni somma ortogonale è una somma diretta.**

Proposizione 10.6.5. *Sia V uno spazio vettoriale euclideo o unitario di dimensione finita n. Sia $W \subset V$ un suo sottospazio vettoriale. Allora $V = W \bigcirc\!\!\!\!\perp W^\perp$. In particolare si ha $\dim V = \dim W + \dim W^\perp$.*

Dimostrazione. Sia $w_1, \ldots, w_m$ una base ortonormale di W, la prolunghiamo a una base ortonormale di V: $(w_1, \ldots, w_m, v_{m+1}, \ldots, v_n)$. Allora ogni vettore $v \in V$ si scrive

$$v = \langle w_1, v \rangle w_1 + \cdots + \langle w_m, v \rangle w_m + \langle v_{m+1}, v \rangle v_{m+1} + \cdots + \langle v_n, v \rangle v_n = \tilde{v} + v',$$

con $\tilde{v} \in W$ e $v' \in W^\perp$, il che prova che $V = W + W^\perp$; siccome $W \perp W^\perp$ la somma è ortogonale. $\square$

Osserviamo in particolare che i vettori aggiunti alla base di partenza $v_{m+1}, \dots, v_n$ costituiscono una base ortonormale di $W^\perp$.

Esempi 10.6.6. 1. **Ortonormalizzare una base con il metodo di Gram-Schmidt**.

Consideriamo la matrice simmetrica $S = \begin{pmatrix} 2 & 1 & -1 \\ 1 & 1 & 0 \\ -1 & 0 & 3 \end{pmatrix}$. La forma bilineare simmetrica su $\mathbb{R}^3$ ad essa associata rispetto alla base canonica opera come segue:

$$\langle x, y\rangle = 2x_1y_1 + x_1y_2 + x_2y_1 - x_1y_3 - x_3y_1 + x_2y_2 + 3x_3y_3,$$

$$\langle x, x\rangle = 2x_1^2 + 2x_1x_2 + x_2^2 - 2x_1x_3 + 3x_3^2 = (x_1 + x_2)^2 + (x_1 - x_3)^2 + 2x_3^2,$$

ed è perciò un prodotto scalare. Vogliamo costruire una base ortonormale di $\mathbb{R}^3$ rispetto a tale prodotto scalare, a partire dalla base canonica, applicando il teorema di Gram-Schmidt 10.6.2. Partiremo da e_1, lo normalizzeremo e otterremo v_1; poi considereremo e_2 che non appartiene a $W_1 = \langle v_1\rangle$; otterremo v_2 come $\frac{e_2-\tilde{e}_2}{\|e_2-\tilde{e}_2\|}$, dove $\tilde{e}_2$ è la proiezione ortogonale di e_2 su W_1; infine considereremo e_3 che non appartiene a W_2, il sottospazio generato da v_1, v_2, e otterremo v_3 come $\frac{e_3-\tilde{e}_3}{\|e_3-\tilde{e}_3\|}$, dove $\tilde{e}_3$ è la proiezione ortogonale di e_3 su W_2.

Osserviamo dunque che $\langle e_1, e_1\rangle = 2$, dunque $\|e_1\| = \sqrt{2}$ e $v_1 = (\frac{1}{\sqrt{2}}, 0, 0)$. Notiamo che e_2 non appartiene a $\langle v_1\rangle = \langle e_1\rangle$, e che $\langle e_1, e_2\rangle = 1 \neq 0$, dunque e_1, e_2 non sono ortogonali. Costruiamo la proiezione ortogonale di e_2 sul sottospazio $W_1 = \langle v_1\rangle$:

$$\tilde{e}_2 = \langle v_1, e_2\rangle v_1 = \langle(\frac{1}{\sqrt{2}}, 0, 0), (0, 1, 0)\rangle(\frac{1}{\sqrt{2}}, 0, 0) = \frac{1}{\sqrt{2}}(\frac{1}{\sqrt{2}}, 0, 0) = (\frac{1}{2}, 0, 0).$$

Allora

$$e_2 - \tilde{e}_2 = (0, 1, 0) - (\frac{1}{2}, 0, 0) = (-\frac{1}{2}, 1, 0),$$
$$\|e_2 - \tilde{e}_2\| = \sqrt{\frac{1}{2} - 1 + 1} = \frac{1}{\sqrt{2}}, \text{ quindi}$$
$$v_2 = \sqrt{2}(-\frac{1}{2}, 1, 0) = (-\frac{\sqrt{2}}{2}, \sqrt{2}, 0).$$

Osserviamo che e_3 non appartiene al sottospazio $W_2 = \langle v_1, v_2\rangle = \langle e_1, e_2\rangle$. Costruiamo la proiezione ortogonale di e_3 su W_2:

$$\begin{aligned}
\tilde{e}_3 &= \langle v_1, e_3\rangle v_1 + \langle v_2, e_3\rangle v_2 = \\
&= \langle\big(\frac{1}{\sqrt{2}}, 0, 0\big), (0,0,1)\rangle\big(\frac{1}{\sqrt{2}}, 0, 0\big) + \langle\big(-\frac{\sqrt{2}}{2}, \sqrt{2}, 0\big), (0,0,1)\rangle\big(-\frac{\sqrt{2}}{2}, \sqrt{2}, 0\big) = \\
&= -\frac{1}{\sqrt{2}}(-\frac{\sqrt{2}}{2}, \sqrt{2}, 0) + \frac{\sqrt{2}}{2}(-\frac{\sqrt{2}}{2}, \sqrt{2}, 0) = (-1, 1, 0).
\end{aligned}$$

Dunque

$$\begin{aligned}
e_3 - \tilde{e}_3 &= (1, -1, 1), \\
\|e_3 - \tilde{e}_3\| &= \sqrt{2}, \\
v_3 &= \frac{(1, -1, 1)}{\sqrt{2}}.
\end{aligned}$$

Abbiamo così costruito una base ortonormale (v_1, v_2, v_3) con sottospazi generati $\langle v_1\rangle = \langle e_1\rangle$, $\langle v_1, v_2\rangle = \langle e_1, e_2\rangle$.

2. Consideriamo su $\mathbb{R}^3$ il prodotto scalare canonico; vogliamo costruire una base ortogonale avente come primo vettore $v_1 = (1, 1, 1)$. Osserviamo che i due vettori rimanenti devono essere ortogonali a v_1, e quindi le loro coordinate x, y, z devono verificare l'equazione $x + y + z = 0$. Una soluzione non nulla è $v_2 = (1, -1, 0)$. Il terzo vettore dev'essere ortogonale sia a v_1 sia a v_2, quindi le sue coordinate devono essere una soluzione del sistema lineare $x + y + z = x - y = 0$, la cui matrice dei coefficienti è

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

Allora, applicando l'Esempio 8.15.2, 3, abbiamo immediatamente una soluzione non nulla $v_3 = (1, 1, -2)$. Povveromo ottenere una base ortonormale normalizzando i tre vettori, e otteniamo:

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

3. Il procedimento seguito nel precedente Esempio 1. porta al procedimento più generale di ortonormalizzazione di una base $\mathcal{B} = (v_1, \ldots, v_n)$. Poniamo

$$W_1 =< v_1 >,\ W_2 =< v_1, v_2 >, \ldots, W_n =< v_1, \ldots, v_n >= V.$$

- Normalizziamo v_1 e otteniamo $w_1 = \frac{v_1}{\|v_1\|}$. Osserviamo che $W_1 =< v_1 >=< w_1 >$ e w_1 costituisce una sua base ortonormale.

- $v_2 \notin W_1$, costruiamo la sua proiezione ortogonale $\tilde{v_2}$ su W_1 di cui abbiamo a disposizione una base ortonormale. Poniamo $w_2 = \frac{v_2 - \tilde{v_2}}{\|v_2 - \tilde{v_2}\|}$. Osserviamo che $W_2 =< v_1, v_2 >=< w_1, w_2 >$ e che quest'ultima è una sua base ortonormale.

- $v_3 \notin W_2$, ne facciamo la proiezione ortogonale $\tilde{v_3}$ su W_2 di cui conosciamo una base ortonormale. Poniamo w_3 la normalizzazione di $v_3 - \tilde{v_3}$, e così via.