Spazi vettoriali Euclidei

Prodotto scalare di due vettori nel piano, lunghezza ed ortogonalita'. Diamo per nota la nozione di "lunghezza" di un vettore (rispetto ad una prefissata unita') e la relazione di "ortogonalita'" fra due vettori non nulli; assumiamo per convenzione che il vettore nullo sia ortogonale ad ogni altro vettore. Indichiamo con $\|\mathbf{a}\|$ la lunghezza di un vettore $\mathbf{a}$ e scriviamo $\mathbf{a} \perp \mathbf{b}$ per indicare che un vettore $\mathbf{a}$ e' ortogonale a un vettore $\mathbf{b}$.

Fissato nel piano un punto O, rappresentiamo ciascun vettore del piano con un segmento orientato con origine O. Identifichiamo i vettori canonici $\mathbf{e}_1$ ed $\mathbf{e}_2$ di $\mathbb{R}^2$ con due vettori del piano aventi lunghezza uno e fra loro ortogonali, e di conseguenza identifichiamo ciascun vettore $\mathbf{a} = (a_1, a_2)$ di $\mathbb{R}^2$ con un vettore $\mathbf{a} = a_1\mathbf{e}_1 + a_2\mathbf{e}_2$ del piano.

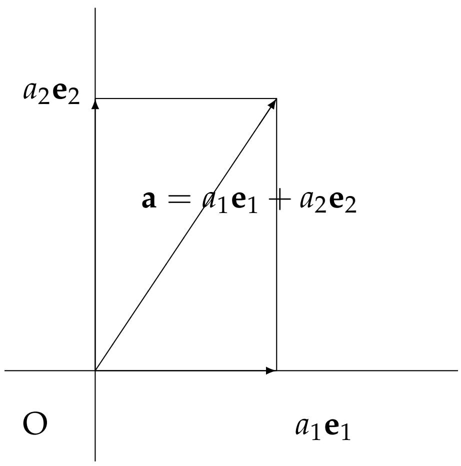
Per ogni due vettori $\mathbf{a} = (a_1, a_2)$ e $\mathbf{b} = (b_1, b_2)$ di $\mathbb{R}^2$ lo scalare

$$\mathbf{a}^T\mathbf{b} = \left[\, a_1 \; a_2 \,\right] \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = a_1b_1 + a_2b_2$$

viene detto "prodotto scalare" di $\mathbf{a}$ per $\mathbf{b}$.

La lunghezza di un vettore si puo' esprimere in funzione del prodotto scalare. Infatti un qualsiasi vettore $\mathbf{a} = a_1\mathbf{e}_1 + a_2\mathbf{e}_2$ e' la diagonale del rettangolo avente per lati i vettori $a_1\mathbf{e}_1$ e $a_2\mathbf{e}_2$, cosi' dal teorema di Pitagora si ha

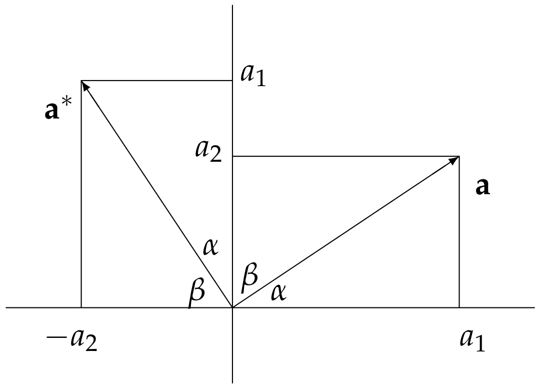
$$\|\mathbf{a}\| = \sqrt{\|a_1\mathbf{e}_1\|^2 + \|a_2\mathbf{e}_2\|^2} = \sqrt{a_1^2 + a_2^2} = \sqrt{\mathbf{a}^T\mathbf{a}}.$$



Anche la relazione di ortogonalita' si puo' esprimere in funzione del prodotto scalare. Precisamente, per ogni coppia di vettori $\mathbf{a} = (a_1, a_2)$ e $\mathbf{b} = (b_1, b_2)$ di $\mathbb{R}^2$, si ha che

$$\mathbf{a} \perp \mathbf{b} \quad \text{se e solo se} \quad a_1b_1 + a_2b_2 = \mathbf{a}^T\mathbf{b} = 0.$$

Ci si puo' rendere conto di una parte di questa affermazione, cioe' che vettori ortogonali abbiano prodotto scalare zero, nel modo seguente. Dato un vettore non nullo $\mathbf{a} = (a_1, a_2)$, ci sono un paio di scelte psicologicamente naturali per un vettore ortogonale ad $\mathbf{a}$, una delle quali e' $\mathbf{a}^* = (-a_2, a_1)$.



La scelta e' corretta. Informalmente, si puo' osservare che si vengono a formare quattro triangoli rettangoli uguali, l'angolo formato dai vettori $\mathbf{a}$ e $\mathbf{a}^*$ e' $\alpha + \beta$, ma $\alpha + \beta$ e' anche l'angolo formato dai due assi coordinati, che e' retto.

I vettori ortogonali al vettore $\mathbf{a}$ sono tutti e soli quelli del tipo $\mathbf{b} = r\mathbf{a}^* = (-ra_2, ra_1)$, dove r e' uno scalare qualsiasi. E in effetti si ha

$$\mathbf{a}^T\mathbf{b} = a_1(-ra_2) + a_2(ra_1) = 0.$$

Prodotto scalare di due vettori nello spazio, lunghezza ed ortogonalita'. Fissato nello spazio un punto O, rappresentiamo ciascun vettore dello spazio con un segmento orientato con origine O. Identifichiamo i vettori canonici $\mathbf{e}_1$, $\mathbf{e}_2$, ed $\mathbf{e}_3$ di $\mathbb{R}^3$ con tre vettori dello spazio aventi ciascuno lunghezza uno e a due a due ortogonali, e di conseguenza identifichiamo ciascun vettore $\mathbf{a} = (a_1, a_2, a_3)$ di $\mathbb{R}^3$ con un vettore $\mathbf{a} = a_1\mathbf{e}_1 + a_2\mathbf{e}_2 + a_3\mathbf{e}_3$ dello spazio.

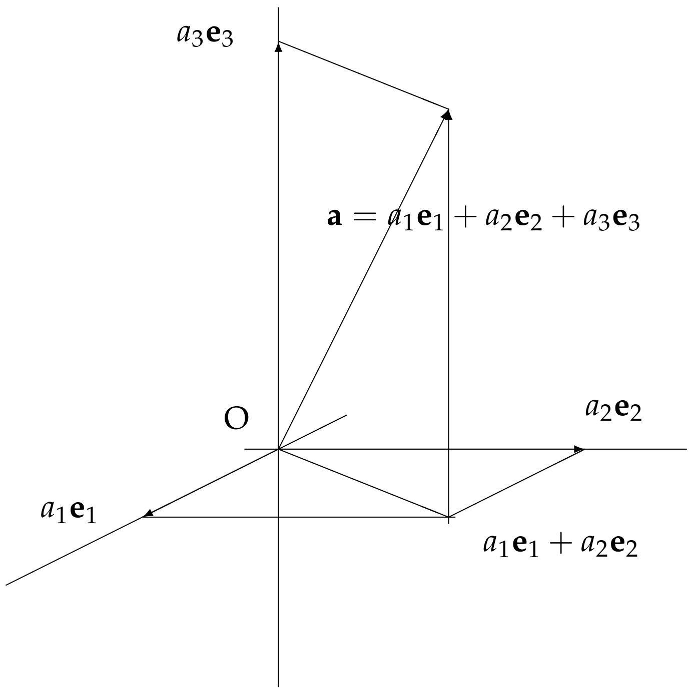
Per ogni due vettori $\mathbf{a} = (a_1, a_2, a_3)$ e $\mathbf{b} = (b_1, b_2, b_3)$ di $\mathbb{R}^3$ lo scalare

$$\mathbf{a}^T\mathbf{b} = [\, a_1 \; a_2 \; a_3 \,] \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = a_1b_1 + a_2b_2 + a_3b_3$$

viene detto "prodotto scalare" di $\mathbf{a}$ per $\mathbf{b}$.

La lunghezza di un vettore si puo' esprimere in funzione del prodotto scalare. Infatti un qualsiasi vettore $\mathbf{a} = a_1\mathbf{e}_1 + a_2\mathbf{e}_2 + a_3\mathbf{e}_3$ e' la diagonale del rettangolo avente per lati i vettori $a_1\mathbf{e}_1 + a_2\mathbf{e}_2$ e $a_3\mathbf{e}_3$, e il vettore $a_1\mathbf{e}_1 + a_2\mathbf{e}_2$ e' la diagonale del rettangolo avente per lati i vettori $a_1\mathbf{e}_1$ e $a_2\mathbf{e}_2$, cosi' dal teorema di Pitagora si ha

$$\begin{aligned}
\|a\| &= \sqrt{\|a_1\mathbf{e}_1 + a_2\mathbf{e}_2\|^2 + \|a_3\mathbf{e}_3\|^2} \\
&= \sqrt{\left(\sqrt{\|a_1\mathbf{e}_1\|^2 + \|a_2\mathbf{e}_2\|^2}\right)^2 + \|a_3\mathbf{e}_3\|^2} \\
&= \sqrt{a_1^2 + a_2^2 + a_3^2} = \sqrt{\mathbf{a}^T\mathbf{a}}.
\end{aligned}$$



Anche la relazione di ortogonalita' si puo' esprimere in funzione del prodotto scalare. Precisamente, per ogni coppia di vettori $\mathbf{a} = (a_1, a_2, a_3)$ e $\mathbf{b} = (b_1, b_2, b_3)$ di $\mathbb{R}^3$, si ha

$$\mathbf{a} \perp \mathbf{b} \quad \text{se e solo se} \quad a_1b_1 + a_2b_2 + a_3b_3 = \mathbf{a}^T\mathbf{b} = 0.$$

Non proviamo questa affermazione; osserviamo solo che e' coerente con il fatto che i tre vettori canonici di $\mathbb{R}^3$ siano stati identificati con tre vettori a due a due ortogonali. Infatti si ha

$$\begin{aligned}
\mathbf{e}_1^T\mathbf{e}_2 &= 1 \cdot 0 + 0 \cdot 1 + 0 \cdot 0 = 0, \\
\mathbf{e}_1^T\mathbf{e}_3 &= 1 \cdot 0 + 0 \cdot 0 + 0 \cdot 1 = 0, \\
\mathbf{e}_2^T\mathbf{e}_3 &= 0 \cdot 0 + 1 \cdot 0 + 0 \cdot 1 = 0.
\end{aligned}$$

Prodotto scalare standard di due vettori in $\mathbb{R}^n$; definizione di lunghezza e di ortogonalita'. Dati due vettori $\mathbf{a} = (a_i)_1^n$ e $\mathbf{b} = (b_i)_1^n$ di $\mathbb{R}^n$, lo scalare

$$\mathbf{a}^T\mathbf{b} = \begin{bmatrix} a_1 & \dots & a_n \end{bmatrix} \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} = \sum_1^n a_i b_i$$

viene detto *prodotto scalare* di $\mathbf{a}$ per $\mathbf{b}$.

Dalle proprieta' delle operazioni nell'algebra delle matrici discendono le seguenti proprieta' del prodotto scalare:

$$\begin{aligned} (\mathbf{a}+\mathbf{c})^T\mathbf{b} &= \mathbf{a}^T\mathbf{b} + \mathbf{c}^T\mathbf{b} \\ \mathbf{a}^T(\mathbf{b}+\mathbf{c}) &= \mathbf{a}^T\mathbf{b} + \mathbf{a}^T\mathbf{c} \\ (r\mathbf{a})^T\mathbf{b} &= r(\mathbf{a}^T\mathbf{b}) = \mathbf{a}^T(r\mathbf{b}), \end{aligned}$$

per ogni $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{R}^n$ ed ogni $r \in \mathbb{R}$. Inoltre, scambiando i fattori il prodotto scalare non cambia:

$$\mathbf{b}^T\mathbf{a} = \sum_1^n b_i a_i = \sum_1^n a_i b_i = \mathbf{a}^T\mathbf{b};$$

percio' potremo dire "prodotto scalare fra $\mathbf{a}$ e $\mathbf{b}$," ed usare indifferentemente la forma $\mathbf{a}^T\mathbf{b}$ o la forma $\mathbf{b}^T\mathbf{a}$.

Osserviamo che il prodotto scalare di un vettore con se' stesso e' la somma dei quadrati delle sue componenti:

$$\mathbf{a}^T\mathbf{a} = \sum_1^n a_i^2,$$

dunque esso e' sempre maggiore-uguale a zero, ed e' zero se e solo se tutte le compnenti sono nulle, cioe' il vettore e' il vettore nullo:

$$\begin{aligned} &\mathbf{a}^T\mathbf{a} \geq 0, && \forall \mathbf{a} \in \mathbb{R}^n \\ &\mathbf{a}^T\mathbf{a} = 0, \quad \text{se e solo se} && \mathbf{a} = \mathbf{0}. \end{aligned}$$

Definizione 1 *Definiamo la norma $\|\mathbf{a}\|$ di un vettore $\mathbf{a} = (a_i)_1^n$ di $\mathbb{R}^n$ pondendo*

$$\|\mathbf{a}\| = \sqrt{\mathbf{a}^T\mathbf{a}} = \sqrt{\sum_{i=1}^n a_i^2}.$$

Definiamo la relazione di ortogonalita' $\perp$ fra due vettori $\mathbf{a} = (a_i)_1^n$ e $\mathbf{b} = (b_i)_1^n$ di $\mathbb{R}^n$ ponendo

$$\mathbf{a} \perp \mathbf{b} \quad \textit{se e solo se} \quad \mathbf{a}^T\mathbf{b} = \sum_{i=1}^n a_i b_i = 0.$$

Osserviamo che per ogni due vettori della base acnonica di $\mathbb{R}^n$ si ha

$$\mathbf{e}_i^T\mathbf{e}_j = \begin{cases} 1 & \text{per } i=j \\ 0 & \text{per } i\neq j \end{cases},$$

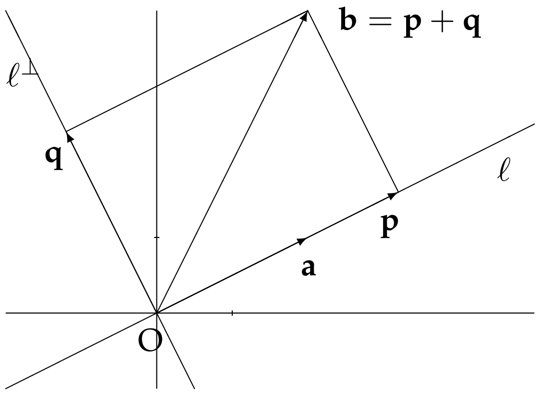
dunque i vettori $\mathbf{e}_1,\ldots,\mathbf{e}_n$ della base canonica di $\mathbb{R}^n$ hanno ciascuno norma 1 e sono a due a due ortogonali.

Proiezione ortogonale di un vettore su un vettore. Consideriamo ora di nuovo lo spazio vettoriale $\mathbb{R}^2$ identificato con lo spazio vettoriale dei vettori del piano, rappresentati da segmenti orientati con origine in un punto O, secondo le condizioni e convenzioni stabilite sopra.

Dato nel piano un vettore $\mathbf{a}$ non nullo, siano ℓ la retta per O generata da $\mathbf{a}$ e $\ell^\perp$ la retta per O ortogonale ad ℓ. Ogni vettore $\mathbf{b}$ del piano si puo' scomporre in uno ed un solo modo come somma di un vettore $\mathbf{p}$ sulla retta ℓ ed un vettore $\mathbf{q}$ sulla retta $\ell^\perp$

$$\begin{aligned} &\mathbf{b} = \mathbf{p}+\mathbf{q} \\ &\mathbf{p}\in\ell \\ &\mathbf{q}\in\ell^\perp \end{aligned}$$

Si dice che $\mathbf{p}$ e' la *proiezione ortogonale* di $\mathbf{b}$ su $\mathbf{a}$ e si pone $\mathbf{p} = \mathrm{pr}_{\mathbf{a}}(\mathbf{b})$.



Questo fatto e questa definizione si estendono ad $\mathbb{R}^n$ come segue.

Proposizione 1 *Dato in $\mathbb{R}^n$ un vettore $\mathbf{a}$ non nullo, siano $\langle\mathbf{a}\rangle$ il sottospazio generato da $\mathbf{a}$ e $\mathbf{a}^\perp$ l'insieme dei vettori ortogonali ad $\mathbf{a}$. Ogni vettore $\mathbf{b}$ di $\mathbb{R}^n$ si puo' scomporre in uno ed un solo modo come somma di un vettore $\mathbf{p}$ in $\langle\mathbf{a}\rangle$ ed un vettore $\mathbf{q}$ in $\mathbf{a}^\perp$*

$$\begin{aligned} &\mathbf{b} = \mathbf{p}+\mathbf{q} \\ &\mathbf{p}\in\langle\mathbf{a}\rangle \\ &\mathbf{q}\in\mathbf{a}^\perp; \end{aligned}$$

esplicitamente, si ha
$$\mathbf{p} = \mathbf{a}\frac{\mathbf{a}^T\mathbf{b}}{\mathbf{a}^T\mathbf{a}}, \quad \mathbf{q} = \mathbf{b} - \mathbf{p}.$$

Lo scalare $\frac{\mathbf{a}^T\mathbf{b}}{\mathbf{a}^T\mathbf{a}}$ si dice *coefficiente di Fourier* di $\mathbf{b}$ rispetto ad $\mathbf{a}$; il vettore $\mathbf{p}$ si dice *proiezione ortogonale* di $\mathbf{b}$ su $\mathbf{a}$ e si indica con

$$\mathrm{pr}_{\mathbf{a}}(\mathbf{b}).$$

Vediamo di seguito come si possa ricavare la formula per la proiezione ortogonale di $\mathbf{b}$ su $\mathbf{a}$. Cerchiamo due vettori $\mathbf{p}, \mathbf{q} \in \mathbb{R}^n$ che soddisfino le condizioni

$$\begin{aligned} \mathbf{b} &= \mathbf{p} + \mathbf{q} \\ \mathbf{p} &= \mathbf{a}r, \qquad r \in \mathbb{R} \\ \mathbf{a}^T\mathbf{q} &= 0, \end{aligned}$$

dove r e' uno scalare incognito. Sostituendo l'espressione di $\mathbf{p}$ in funzione di r nella prima condizione
$$\mathbf{b} = \mathbf{a}r + \mathbf{q},$$
ed applicando and entrambe i membri il prodotto scalare con $\mathbf{a}$ si ha

$$\mathbf{a}^T\mathbf{b} = (\mathbf{a}^T\mathbf{a})r + \mathbf{a}^T\mathbf{q},$$

da cui, per la terza condizione, si ha

$$\mathbf{a}^T\mathbf{b} = (\mathbf{a}^T\mathbf{a})r.$$

Ora, questa e' un'equazione lineare nell'incognita r, e il coefficiente $\mathbf{a}^T\mathbf{a}$ e' diverso da 0 in quanto $\mathbf{a}$ e' diverso dal vettore nullo. Si ha cosi' una ed una sola soluzione:

$$r = \frac{\mathbf{a}^T\mathbf{b}}{\mathbf{a}^T\mathbf{a}},$$

dalla quale si ottiene
$$\mathbf{p} = \mathbf{a}r = \mathbf{a}\frac{\mathbf{a}^T\mathbf{b}}{\mathbf{a}^T\mathbf{a}}.$$

Esempio Nel caso concreto dei vettori del piano in cui $\mathbf{a} = (2,1)$, e $\mathbf{b} = (2,4)$, si ha

$$r = \frac{\mathbf{a}^T\mathbf{b}}{\mathbf{a}^T\mathbf{a}} = \frac{\begin{bmatrix} 2 & 1 \end{bmatrix}\begin{bmatrix} 2 \\ 4 \end{bmatrix}}{\begin{bmatrix} 2 & 1 \end{bmatrix}\begin{bmatrix} 2 \\ 1 \end{bmatrix}} = \frac{8}{5}.$$

da cui
$$\mathrm{pr}_a(\mathbf{b}) = \mathbf{a}r = (2,1)\,\frac{8}{5} = (\frac{16}{5}, \frac{8}{5}).$$

Ortogonalita' e indipendenza lineare. Nel piano, due vettori non nulli fra loro ortogonali sono sempre linearmente indipendenti; nello spazio, due vettori non nulli fra loro ortogonali sono sempre linearmente indipendenti, e tre vettori non nulli a due a due ortogonali sono sempre linearmente indipendenti. Questi fatti si estendono ad $\mathbb{R}^n$ come segue

Proposizione 2 *Se dei vettori* $\mathbf{a}_1, \ldots, \mathbf{a}_m$ *di* $\mathbb{R}^n$ *sono non nulli e a due a due ortogonali, allora sono linearmente indipendenti.*

Dimostrazione. Consideriamo l'equazione

$$\mathbf{a}_1 x_1 + \mathbf{a}_2 x_2 + \cdots + \mathbf{a}_m x_m = \mathbf{0}$$

nelle incognite scalari $x_1, x_2, \ldots, x_n$. Applicando ad entrame i membri il prodotto scalare con $\mathbf{a}_i$ si ottiene

$$(\mathbf{a}_i^T \mathbf{a}_1) x_1 + (\mathbf{a}_i^T \mathbf{a}_2) x_2 + \cdots + (\mathbf{a}_i^T \mathbf{a}_m) x_m = 0$$

da cui, essendo $\mathbf{a}_i$ ortogonale a tutti gli altri $\mathbf{a}_j$ si ha

$$(\mathbf{a}_i^T \mathbf{a}_i) x_i = 0,$$

da cui, essendo $\mathbf{a}_i \neq \mathbf{0}$, si ottiene $x_i = 0$. Dunque $x_1 = x_2 = \ldots = x_m = 0$, e cio' significa che i vettori $\mathbf{a}_1, \ldots, \mathbf{a}_m$ sono linearmente indipendenti.

Da questa proposizione segue che se n vettori $\mathbf{a}_1, \ldots, \mathbf{a}_n$ di $\mathbb{R}^n$ sono non nulli e a due a due ortogonali, allora sono una base di $\mathbb{R}^n$. Una tale base si dice "base ortogonale" di $\mathbb{R}^n$; se tutti i vettori a_i hanno norma uno, allora essa si dice "base ortonormale" di $\mathbb{R}^n$.

Le coordinate di un vettore rispetto ad una base ortogonale o ortonormale sono particolarmente semplici da determinare, come mostrato dalla seguente

Proposizione 3 *Le coordinate di un qualsiasi vettore* $\mathbf{b} \in \mathbb{R}^n$ *rispetto ad una base ortogonale* $\mathbf{a}_1, \ldots, \mathbf{a}_n$ *di* $\mathbb{R}^n$ *sono i coefficienti di Fourier di* $\mathbf{b}$ *rispetto ai vettori* $\mathbf{a}_1, \ldots, \mathbf{a}_n$:

$$\mathbf{b} = \sum_{i=1}^{n} \mathbf{a}_i \frac{\mathbf{a}_i^T \mathbf{b}}{\mathbf{a}_i^T \mathbf{a}_i};$$

se la base e' ortonormale, allora le coordinate di $\mathbf{b}$ *sono i prodotti scalari di* $\mathbf{b}$ *per i vettori* $\mathbf{a}_1, \ldots, \mathbf{a}_n$:

$$\mathbf{b} = \sum_{i=1}^{n} \mathbf{a}_i (\mathbf{a}_i^T \mathbf{b}).$$

Dimostrazione. Consideriamo l'equazione

$$\mathbf{a}_1 x_1 + \mathbf{a}_2 x_2 + \cdots + \mathbf{a}_m x_m = \mathbf{b}$$

nelle incognite scalari $x_1, x_2, \ldots, x_n$. Applicando ad entrame i membri il prodotto scalare con $\mathbf{a}_i$ si ottiene

$$(\mathbf{a}_i^T \mathbf{a}_1)x_1 + (\mathbf{a}_i^T \mathbf{a}_2)x_2 + \cdots + (\mathbf{a}_i^T \mathbf{a}_m)x_m = \mathbf{a}_i^T \mathbf{b}$$

da cui, essendo $\mathbf{a}_i$ ortogonale a tutti gli altri $\mathbf{a}_j$ si ha

$$(\mathbf{a}_i^T \mathbf{a}_i)x_i = \mathbf{a}_i^T \mathbf{b},$$

da cui, essendo $\mathbf{a}_i \neq \mathbf{0}$, si ottiene

$$x_i = \frac{\mathbf{a}_i^T \mathbf{b}}{\mathbf{a}_i^T \mathbf{a}_i};$$

se poi ogni vettore $\mathbf{a}_i$ ha norma uno, si ha $x_i = \mathbf{a}_i^T \mathbf{b}$.

Esempio. In $\mathbb{R}^2$ le coordinate del vettore $(7, 8)$ rispetto alla base ortogonale $(2, 1), (-2, 4)$ sono

$$\frac{\mathbf{a}_1^T \mathbf{b}}{\mathbf{a}_1^T \mathbf{a}_1} = \frac{22}{5}, \quad \frac{\mathbf{a}_2^T \mathbf{b}}{\mathbf{a}_2^T \mathbf{a}_2} = \frac{18}{20} = \frac{9}{10},$$

dunque si ha

$$(7, 8) = \frac{22}{5}(2, 1) + \frac{9}{10}(-2, 4).$$