Matematica per Analisi dei Dati, 16.03.09

1. **Lunghezza di vettori in $\mathbb{R}^n$**

   Fissato nel piano un sistema di riferimento cartesiano ortogonale monometrico, si ha che la lunghezza del vettore $v = [v_i]_{i=1}^2$ rispetto all'unita' di misura scelta e' data da
   $$\|v\| = \sqrt{v_1^2 + v_2^2}.$$
   Fissato nello spazio un sistema di riferimento cartesiano ortogonale monometrico, si ha che la lunghezza del vettore $v = [v_i]_{i=1}^3$ rispetto all'unita' di misura scelta e' data da
   $$\|v\| = \sqrt{v_1^2 + v_2^2 + v_3^2}.$$
   Questi fatti suggeriscono di definire la lunghezza di un vettore $v = [v_i]_{i=1}^n$ dello spazio vettoriale $\mathbb{R}^n$ come
   $$\|v\| = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}.$$
   Usando il prodotto interno, possiamo scrivere sinteticamente
   $$\|v\| = \sqrt{v^T v}.$$
   Osserviamo che i vettori $e_1, \ldots, e_n$ della base canonica di $\mathbb{R}^n$ hanno tutti lunghezza 1. Un vettore di lunghezza 1 viene detto *versore*.

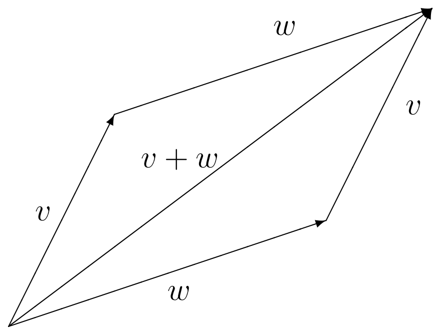
2. Osserviamo che, per ogni $a, b \in \mathbb{R}^n$, la lunghezza del vettore somma $a+b$ e' legata alle lunghezze dei vettori addenti a, b dalla relazione
   $$\begin{aligned} \|a+b\|^2 &= (a+b)^T(a+b) \\ &= \left(a^T + b^T\right)(a+b) \\ &= a^T a + a^T b + b^T a + b^T b \\ &= \|a\|^2 + 2a^T b + \|b\|^2. \end{aligned}$$

3. **Proprieta', norma**

   La lunghezza di vettori di $\mathbb{R}^n$ possiede le seguenti proprieta', che hanno un sapore geometrico, ma la cui validita' si fonda su fatti algebrici:

   - per ogni $v \in \mathbb{R}^n$, si ha
     $$\|v\| \geq 0; \qquad \textit{inoltre}, \qquad \|v\| = 0 \quad \textit{se e solo se } v = 0_n;$$
   - per ogni $r \in \mathbb{R}$ ed ogni $v \in \mathbb{R}^n$, si ha
     $$\|rv\| = |r|\|v\|;$$
   - per ogni $v, w \in \mathbb{R}^n$, si ha
     $$\|v + w\| \leq \|v\| + \|w\|.$$

L'ultima proprieta' viene detta *disuguaglianza triangolare*, per il significato che acquista nel caso $n = 2$:



Una funzione $\mathbb{R}^n \to \mathbb{R}$ che soddisfi le proprieta' sopra elencate viene detta *norma*. Quella che noi abbiamo definito e' una particolare norma, detta *norma euclidea*.

4. **Disuguaglianza di Cauchy-Schwarz**

La norma euclidea e' legata al prodotto interno dalla seguente proprieta':

Per ogni due vettori $a = [a_i]_{i=1}^n$ e $b = [b_i]_{i=1}^n$ in $\mathbb{R}^n$, si ha

$$|a_1 b_1 + \cdots + a_n b_n| \leq \sqrt{a_1^2 + \cdots + a_n^2}\sqrt{b_1^2 + \cdots + b_n^2},$$

in breve

$$|a^T b| \leq \|a\|\|b\|.$$

Questa disuguaglianza e' nota come *disuguaglianza di Cauchy-Schwarz.*

Sotto la condizione $a, b \neq 0_n$, la disuguaglianza di Cauchy-Schwarz si puo' riscrivere nella forma

$$-1 \leq \frac{a^T b}{\|a\|\|b\|} \leq 1;$$

esiste allora uno ed un solo angolo ϑ, con $0 \leq \vartheta \leq \pi$, tale che $\cos\vartheta = \frac{a^T b}{\|a\|\|b\|}$. Si dice che ϑ e' *l'angolo formato dai vettori* a e b; questo angolo si indica con $\hat{ab}$, e cosi' si ha

$$\cos \hat{ab} = \frac{a^T b}{\|a\|\|b\|},$$

esplicitamente

$$\cos \hat{ab} = \frac{a_1 b_1 + \cdots + a_n b_n}{\sqrt{a_1^2 + \cdots + a_n^2}\sqrt{b_1^2 + \cdots + b_n^2}}.$$

In realta', l'angolo $\hat{ab}$ dipende dai due vettori a, b solo attraverso le due semirette da essi generate. Infatti, se sostituiamo a con un suo multiplo scalare ra, con

$r \in \mathbb{R}^+$, e sostituiamo b con un suo multiplo scalare sb, con $s \in \mathbb{R}^+$, l'espressione che definisce il coseno dell'angolo rimane invariata:

$$\frac{(ra)^T(sb)}{\|ra\|\|sb\|} = \frac{rsa^Tb}{|r||s|\|a\|\|b\|} = \frac{a^Tb}{\|a\|\|b\|}.$$

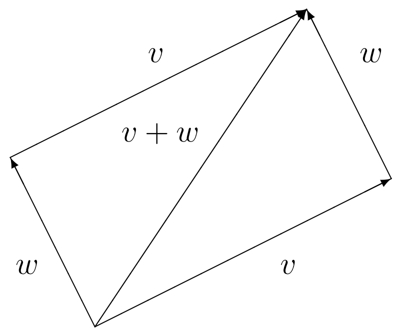
5. **Teorema di Pitagora**

La norma euclidea soddisfa ulteriori proprieta', ad esempio:

Per ogni $v, w \in \mathbb{R}^n$, con $v \perp w$, si ha

$$\|v + w\|^2 = \|v\|^2 + \|w\|^2.$$

Questa proprieta' viene detta *Teorema di Pitagora*, per il significato che acquista nel caso $n = 2$:



6. **Basi ortogonali e basi ortonormali**

Nello spazio ordinario $\mathbb{R}^3$, si puo' osservare che due vettori non nulli fra loro ortogonali sono linearmente indipendenti, e che tre vettori non nulli a due a due ortogonali sono linearmente indipendenti, e dunque formano una base di $\mathbb{R}^3$. In generale si ha

- **Proposizione** *Se i vettori $v_1, v_2, \ldots, v_p$ di $\mathbb{R}^n$ sono non nulli e a due a due ortogonali, allora i vettori $v_1, v_2, \ldots, v_p$ sono linearmente indipendenti.*

  Infatti, se

  $$r_1v_1 + r_2v_2 + \cdots + r_pv_p = 0_n, \qquad r_i \in \mathbb{R},$$

  e' una combinazione lineare dei vettori v_i il cui risultato e' il vettore nullo, allora moltiplicando entrambe i membri a sinistra per v_1^T si ha

  $$v_1^T\left(r_1v_1 + r_2v_2 + \cdots + r_pv_p\right) = v_1^T0_n,$$

  da cui

  $$r_1\left(v_1^Tv_1\right) + r_2\left(v_1^Tv_2\right) + \cdots + r_p\left(v_1^Tv_p\right) = 0;$$

ora, per l'ipotesi che v_1 sia ortogonale a tutti gli altri vettori, si ha

$$r_1 \left(v_1^T v_1 \right) = 0,$$

e, per l'potesi che $v_1 \neq 0_n$, si ha

$$r_1 = 0.$$

In modo analogo si mostra che tutti gli altri coefficienti r_i devono essere nulli.

- **Proposizione** *Se i vettori $v_1, v_2, \ldots, v_n$ di $\mathbb{R}^n$ sono non nulli e a due a due ortogonali, allora i vettori $v_1, v_2, \ldots, v_n$ formano una base di $\mathbb{R}^n$.*

  Cio' deriva direttamente dalla proposizione precedente.

  Osserviamo che e' facile determinare le coordinate di un vettore b rispetto alla base dei vettori $v_1, v_2, \ldots, v_n$; infatti se

  $$b = r_1 v_1 + r_2 v_2 + \cdots + r_p v_p \qquad r_i \in \mathbb{R},$$

  e' una scrittura di b come combinazione lineare dei vettori v_i, allora moltiplicando entrambe i membri a sinistra per v_1^T si ha

  $$v_1^T b = v_1^T \left(r_1 v_1 + r_2 v_2 + \cdots + r_p v_p \right),$$

  da cui

  $$v_1^T b = r_1 \left(v_1^T v_1 \right) + r_2 \left(v_1^T v_2 \right) + \cdots + r_p \left(v_1^T v_p \right);$$

  ora, per l'ipotesi che v_1 sia ortogonale a tutti gli altri vettori, si ha

  $$v_1^T b = r_1 \left(v_1^T v_1 \right),$$

  e, per l'potesi che $v_1 \neq 0_n$, possiamo ricavare

  $$r_1 = \frac{v_1^T b}{v_1^T v_1}.$$

  In modo analogo si ricavano tutti le altre coordinate, e si ha

  $$r_i = \frac{v_i^T b}{v_i^T v_i}, \qquad i = 1, 2, \ldots, n.$$

  Osserviamo infine che, se i vettori v_i hanno tutti norma 1, allora la $i-$ma coordinata di b rispetto alla base $v_1, \ldots, v_n$ e' semplicemnte il prodotto interno di b col vettore v_i :

  $$r_i = v_i^T b \qquad i = 1, 2, \ldots, n.$$

Sia $\{v_1, \ldots, v_n\}$ un insieme di vettori di R^n; se i vettori v_i sono non nulli e a due a due ortogonali, diciamo che $\{v_1, \ldots, v_n\}$ e' una *base ortogonale* di $\mathbb{R}^n$; se i vettori v_i sono versori a due a due ortogonali, diciamo che $\{v_1, \ldots, v_n\}$ e' una *base ortonormale* di $\mathbb{R}^n$.

7. **Esempio**

I vettori

$$v_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \quad v_2 = \begin{bmatrix} -3 \\ 6 \end{bmatrix}$$

formano una base ortogonale di $\mathbb{R}^2$. Le coordinate del vettore $b = \begin{bmatrix} 7 \\ 8 \end{bmatrix}$ rispetto a questa base, cioe' i coefficienti della scrittura di b come combinazione lineare

$$\begin{bmatrix} 7 \\ 8 \end{bmatrix} = r_1 \begin{bmatrix} 2 \\ 1 \end{bmatrix} + r_2 \begin{bmatrix} -3 \\ 6 \end{bmatrix}$$

dei vettori v_1, v_2, sono dati da

$$\begin{aligned} r_1 &= \frac{v_1^T b}{v_1^T v_1} = \frac{22}{5}, \\ r_2 &= \frac{v_2^T b}{v_2^T v_2} = \frac{27}{45}. \end{aligned}$$

I vettori

$$w_1 = \begin{bmatrix} \frac{\sqrt{3}}{2} \\ \\ \frac{1}{2} \end{bmatrix}, \qquad w_2 = \begin{bmatrix} -\frac{1}{2} \\ \\ \frac{\sqrt{3}}{2} \end{bmatrix}$$

formano una base ortonormale di $\mathbb{R}^2$. Le coordinate del vettore $b = \begin{bmatrix} 7 \\ \\ 8 \end{bmatrix}$ rispetto a questa base, cioe' i coefficienti della scrittura di b come combinazione lineare

$$\begin{bmatrix} 7 \\ \\ 8 \end{bmatrix} = r_1 \begin{bmatrix} \frac{\sqrt{3}}{2} \\ \\ \frac{1}{2} \end{bmatrix} + r_2 \begin{bmatrix} -\frac{1}{2} \\ \\ \frac{\sqrt{3}}{2} \end{bmatrix}$$

dei vettori w_1, w_2, sono dati da

$$\begin{aligned} r_1 &= w_1^T b = \frac{7\sqrt{3} + 8}{2}, \\ r_2 &= w_2^T b = \frac{-7 + 8\sqrt{3}}{2}. \end{aligned}$$

8. **Matrici ortogonali**

Sia $\{v_1, v_2, \ldots, v_n\}$ una base ortonormale di $\mathbb{R}^n$, siano soddisfatte cioe' le condizioni

$$\begin{aligned} v_i^T v_j &= 0, \qquad \forall i \neq j \\ v_i^T v_i &= 1, \qquad \forall i. \end{aligned}$$

Ora, queste condizioni possono essere riassunte nell'unica condizione

$$\begin{bmatrix} v_1^T v_1 & v_1^T v_2 & \dots & v_1^T v_n \\ v_2^T v_1 & v_2^T v_2 & \dots & v_2^T v_n \\ \vdots & \vdots & \ddots & \vdots \\ v_n^T v_1 & v_n^T v_2 & \dots & v_n^T v_n \end{bmatrix} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix},$$

che puo' essere espressa nella forma

$$\left[\begin{array}{c} v_1^T \\ \hline v_2^T \\ \hline \vdots \\ \hline v_n^T \end{array}\right] \left[\begin{array}{c|c|c|c} & & & \\ v_1 & v_2 & \cdots & v_n \\ & & & \end{array}\right] = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix},$$

e infine, indicata con A la matrice avente come colonne i vettori v_i, nella forma

$$A^T A = I_n.$$

Si verifica che vale anche la

$$AA^T = I_n.$$

Cosi' la matrice A e' invertibile, e la sua inversa coincide con la sua trasposta:

$$A^{-1} = A^T.$$

Ripercorrendo i passi al contrario, si arriva a vedere che vale anche il viceversa: se una matrice quadrata di ordine n e' invertibile e la sua inversa coincide con la sua trasposta, allora le colonne della matrice formano una base ortonormale di $\mathbb{R}^n$, e anche le righe della matrice formano una base ortonormale di $\mathbb{R}^n$.

Una matrice che soddisfa questa condizione si dice *matrice ortogonale*.

9. Data una matrice Q quadrata di ordine n, associando a ciascun vettore $v \in \mathbb{R}^n$ il vettore $f_Q(v) = Qv \in \mathbb{R}^n$ otteniamo una funzione

$$f_Q : \mathbb{R}^n \to \mathbb{R}^n.$$

Se la matrice Q e' ortogonale, si ha che la funzione f_Q conserva il prodotto interno (e dunque la norma e gli angoli). Infatti, per ogni due vettori v e w di $\mathbb{R}^n$, si ha

$$(f_Q(v))^T f_Q(w) = (Bv)^T Bw = v^T B^T Bw = v^T I_n w = v^T w.$$

Viceversa, si puo' provare che se una matrice possiede questa proprieta', allora deve essere ortogonale.