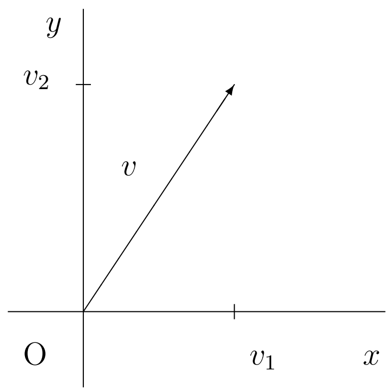
Matematica II 16.12.09

Norma.

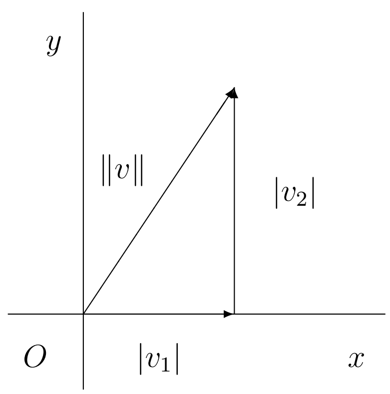
1. **Norma di un vettore.**

   Sia fissato nel piano un sistema di riferimento cartesiano ortogonale monometrico con origine in un punto O.

   Possiamo interpretare un vettore $v = [v_i]_{1=1}^2$ di $\mathbb{R}^2$ come il vettore con origine in O e punta nel punto di coordinate $(v_i)_{1=1}^2$:

   Ora, un punto che partendo da O si sposta di v_1 unita' nella direzione dell'asse x e poi si sposta di v_2 unita' nella direzione dell'asse y descrive i due cateti di un triangolo rettangolo che ammette v come ipotenusa.

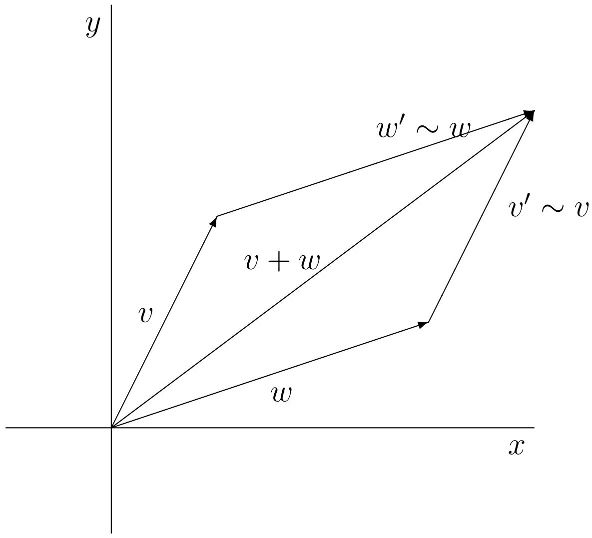
   

   Dunque, per il Teorema di Pitagora, si ha che la lunghezza $\|v\|$ del vettore $v = [v_i]_{1=1}^2$ e' data, nell'unita' di misura scelta, da

   $$\|v\| = \sqrt{|v_1|^2 + |v_2|^2} = \sqrt{v_1^2 + v_2^2}.$$

   La lunghezza dei vettori e' legata alle operazioni sui vettori nel modo seguente:

   - Consideriamo due vettori v, w e il vettore $v + w$ loro somma.



Osserviamo che un punto che partendo da O si sposta prima lungo il vettore v e poi si sposta lungo il vettore $w' \sim w$ descrive due lati di un triangolo che ammette il vettore $v + w$ come terzo lato. Ora, la lunghezza di un lato di un triangolo non supera la somma delle lunghezze degli altri due. Si ha cosi' si ha disuguaglianza

$$\|v + w\| \leq \|v\| + \|w\|,$$

detta appunto *disuguaglianza triangolare*.

- Consideriamo un vettore v, uno scalre r e il vettore vr multiplo di v secondo r. Allora:

$$\|vr\| = \|v\||r|,$$

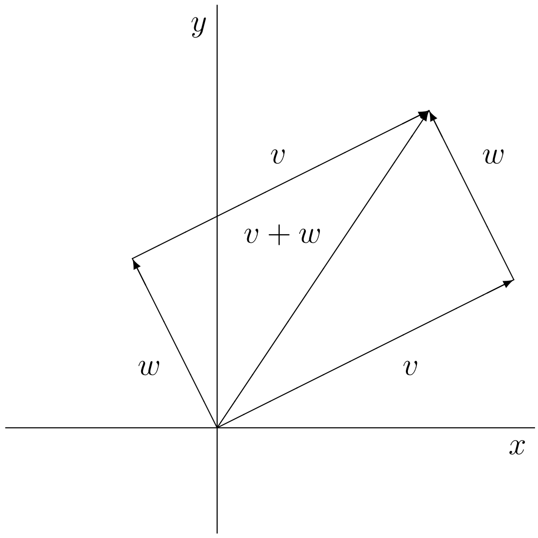
dove $|r|$ e' il valore assoluto di r.

Utilizzando il prodotto interno, possiamo riscrivere la lunghezza di un vettore a di $\mathbb{R}^2$ nella forma

$$\|a\| = \sqrt{a'a}.$$

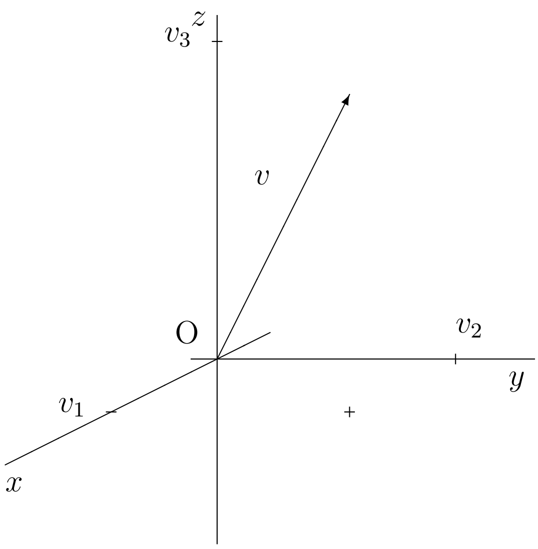
Il teorema di Pitagora puo' essere espresso algebricamente nella forma: se due vettori v e w sono fra loro ortogonali, allora il quadrato della lunghezza del vettore somma $v+w$ e' uguale alla somma dei quadrati delle lunghezze dei vettori addendi v, w :

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

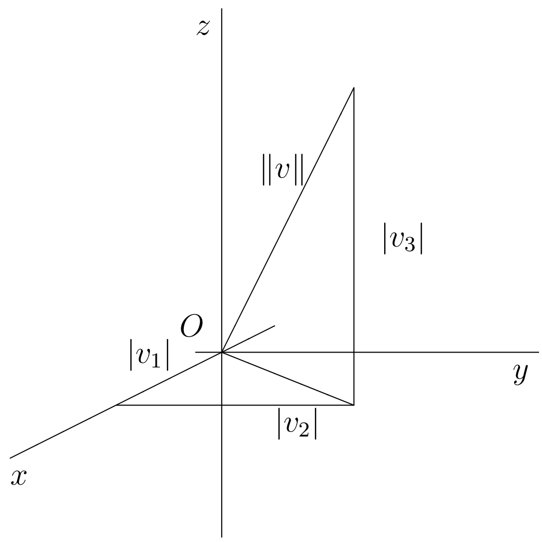


Sia fissato nello spazio un sistema di riferimento cartesiano ortogonale monometrico con origine in un punto O.

Possiamo interpretare un vettore $v = [v_i]_{i=1}^3$ di $\mathbb{R}^3$ come il vettore con origine nel punto origine O avente punta nel punto di coordinate $(v_i)_{i=1}^3$:



Ora, un punto che partendo da O si sposta prima nel punto di coordinate $(v_1, v_2, 0)$ e poi si sposta di v_3 unita' nella direzione dell'asse z, descrive i due cateti di un triangolo rettangolo che ammette v come ipotenusa.



Dunque, per il Teorema di Pitagora, si ha che la lunghezza $\|v\|$ del vettore $v = [v_i]_{i=1}^3$ e' data, nell'unita' di misura scelta, da

$$\|v\| = \sqrt{\sqrt{|v_1|^2 + |v_2|^2}^2 + |v_3|^2} = \sqrt{v_1^2 + v_2^2 + v_3^2}.$$

Utilizzando il prodotto interno, possiamo riscrivere la lunghezza di un vettore a di $\mathbb{R}^3$ nella forma

$$\|a\| = \sqrt{a'a}.$$

Nello spazio vettoriale $\mathbb{R}^n$ definiamo la lunghezza $\|v\|$ di un vettore $v = [v_i]_{i=1}^n$ di $\mathbb{R}^n$ ponendo

$$\|v\| = \sqrt{\sum_{i=1}^n v_i^2} = \sqrt{v'v}.$$

Solitamente, specialmente nel caso $n > 3$, al termine *lunghezza* si preferisce il termine *norma.*

2. Proprieta'.

   Il prodotto interno di due vettori u, v e' legato alle loro norme dalla *disuguaglianza di Cauchy-Schwarz*:

   $$|u'v| \leq \|u\|\|v\|,$$

   o, in altri termini,

   $$-1 \leq \frac{u'v}{\|u\|\|v\|}. \leq 1$$

   Questa disuguaglianza permette di definire l'angolo formato da due vettori: il coseno dell'angolo dei vettori u e v viene definito ponendo

   $$\cos \widehat{uv} = \frac{u'v}{\|u\|\|v\|}.$$

Si osservi che, in base a questa definizione, si ha:

$$\cos \widehat{uv} = \begin{array}{rlll} 1 & se & u = tv, & con\ t > 0 \\ 0 & se & u \perp v & \\ -1 & se & u = tv, & con\ t < 0 \end{array} .$$

Un semplice conto particolarmente interessante:

$$\begin{aligned} \|u + v\|^2 &= (u + v)'(u + v) \\ &= (u' + v')(u + v) \\ &= u'u + u'v + v'u + v'v \\ &= \|u\|^2 + 2u^T v + \|v\|^2. \end{aligned}$$

Le principali proprieta' della norma dei vettori in $\mathbb{R}^n$ sono:

- la norma di un vettore e' sempre maggiore-uguale a zero e vale zero se e solo se il vettore e' nullo:
  $\|v\| \geq 0 \quad \forall v \in \mathbb{R}^n$;
  $\|v\| = 0 \quad se\ e\ solo\ se \quad v = 0_n$.
- La norma del vettore prodotto di un vettore per uno scalare e' uguale al prodotto della norma del vettore per il valore assoluto dello scalare:
  $$\|vr\| = \|v\||r|, \qquad r \in \mathbb{R}.$$
- La norma del vettore somma e' minore o uguale alla somma delle norme dei vettori addendi:
  $$\|v + w\| \leq \|v\| + \|w\|.$$
  Questa disuguaglianza viene detta *disuguaglianza triangolare*, per il significato che assume nel piano e nello spazio.

Le prime due proprieta' seguono direttamente dalle proprieta' del prodotto interno. Mostriamo come la disuguaglianza triangolare si possa dimostrare usando la disuuaglianza di Cauchy-Schwarz.

Infatti si ha

$$\begin{aligned} \|u + v\|^2 &= \|u\|^2 + 2u'v + \|v\|^2 \\ &\leq \|u\|^2 + 2|u'v| + \|v\|^2 \\ &\leq \|u\|^2 + 2\|u\|\|v\| + \|v\|^2 \\ &= (\|u\| + \|v\|)^2 . \end{aligned}$$

Dunque si ha la disuguaglianza

$\|u + v\|^2 \leq (\|u\| + \|v\|)^2$,

che e' equivalente alla disuguaglianza

$\|u + v\| \leq \|u\| + \|v\|$,

in quanto sia $\|u+v\|$ che $\|u\|+\|v\|$ sono ≥ 0.

Nello spazio $\mathbb{R}^n$ vale il teorema di Pitagora: se due vettori a e b sono ortogonali, cioe' se $a'b = 0 = b'a$, allora

$$\|a+b\|^2 = \|a\|^2 + \|b\|^2.$$

Infatti

$$\|a+b\|^2 = \|a\|^2 + 2a'b + \|b\|^2 = \|a\|^2 + \|b\|^2.$$

Minimi quadrati

1. Consideriamo il sistema lineare

$$\begin{cases} x_1 + \ \ x_2 = 2 \\ x_1 + 2x_2 = 3 \\ 2x_1 + 3x_2 = 4 \end{cases}.$$

Questo sistema non ha soluzioni. Per ogni $s = (s_1, s_2) \in \mathbb{R}^2$, possiamo considerare il vettore degli scarti fra i primi e i secondi membri

$$E(s) = \begin{bmatrix} s_1 + \ \ s_2 - 2 \\ s_1 + 2s_2 - 3 \\ 2s_1 + 3s_2 - 4 \end{bmatrix}$$

come *l'errore* che si commette assumendo che s sia una soluzione del sistema; questo errore e' un vettore di $\mathbb{R}^3$, che possiamo misurare con la sua norma

$$\|E(s)\|.$$

Ad esempio, per $s = (1, 0)$, si ha

$$\|E(s)\| = \left\| \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} \right\| = \sqrt{5},$$

e per $s = (0, 1)$, si ha

$$\|E(s)\| = \left\| \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\| = \sqrt{3},$$

Siammo cosi' condotti a preferire la coppia $(0, 1)$ alla coppia $(1, 0)$. Ci possiamo chiedere se c'e' una scelta ottima per s.

2. Soluzioni ai minimi quadrati

Consideriamo un sistema lineare di m equazioni in n incognite

$$Ax = b, \qquad A \in \mathbb{R}^{m \times n}, \quad b \in \mathbb{R}^m.$$

Per ogni valore s di $x \in \mathbb{R}^n$, possiamo considerare il vettore scarto fra il primo ed il secondo membro

$$E(s) = As - b$$

come *l'errore* che si commette assumendo che s sia una soluzione del sistema; questo errore e' un vettore di $\mathbb{R}^m$, che possiamo misurare con la sua norma

$$\|E(s)\| = \|As - b\|.$$

Si noti che s e' una soluzione del sistema se e solo se $\|E(s)\| = 0$. Un vettore $x^* \in \mathbb{R}^n$ cui corrisponde un errore di norma minima, cioe' tale che

$$\|E(x^*)\| \leq \|E(x)\| \qquad \forall x \in \mathbb{R}^n,$$

si dice *soluzione ai minimi quadrati* del sistema.

Teorema 1. *Sia $Ax = b$, un sistema lineare di m equazioni in n incognite.*

- *Le soluzioni ai minimi quadrati del sistema $Ax = b$ sono tutte e sole le soluzioni del sistema di n equazioni in n incognite*

$$A^T A\, x = A^T b;$$

- *questo sistema ha sempre qualche soluzione; la soluzione e' unica se e solo se le colonne di A sono linearmente indipendenti; in questo caso, l'unica soluzione e' data da*

$$x = \left(A^T A\right)^{-1} A^T b.$$

Dunque abbiamo che

> *un sistema lineare $Ax = b$ ha sempre qualche soluzione ai minimi quadrati; la soluzione ai minimi quadrati e' unica se e solo se le colonne di A sono linearmente indipendenti; in questo caso, l'unica soluzione ai minimi quadrati e' data dal coefficiente di Fourier di b rispetto ad A.*

3. **Interpolazione**

Consideriamo un insieme "valori osservati"

x	y
x_1	y_1
x_2	y_2
$\vdots$	$\vdots$
x_m	y_m

di due "variabili" x, y ed una famiglia $\mathcal{F}$ di funzioni

$$y = f(x).$$

Possiamo allora cercare nella famiglia $\mathcal{F}$ una funzione f che "interpola" i valori osservati delle variabili:

$$y_i = f(x_i), \qquad i = 1, 2, \ldots, m.$$

Supponiamo che il numero m dei dati osservati sia "grande" e che la famiglia $\mathcal{F}$ sia costituita da "poche" funzioni, cosi' che non ci sia in $\mathcal{F}$ alcuna funzione che interpola i dati osseervati. Per ogni funzione f possiamo considerare il vettore degli scarti

$$E(f) = [f(x_i) - y_i]_1^m$$

come l'errore che si commette assumendo che f interpoli i dati osservati, e possiamo misurare questo errore con la sua norma

$$\|E(f)\| = \| \, [f(x_i) - y_i]_1^m \, \|.$$

Se $\|E(f)\| \leq \|E(q)\|$, diciamo che f interpola meglio di g l'insieme dei dati osservati, nel senso dei minimi quadrati.

Consideriamo il caso in cui

$$\mathcal{F} = \left\{ y = c_0 + c_1 x + c_2 x^2 + \ldots + c_{n-1} x^{n-1}; \; c_j \in \mathbb{R} \right\}$$

sia la famiglia dei polinomi di grado minore di n.

Un polinomio interpola i dati osservati se e solo se i suoi n coefficienti $c_0, c_1, \ldots, c_{n-1}$ sono soluzioni del sistema di m equazioni lineari

$$c_0 + c_1 x_i + c_2 x_i^2 + \ldots + c_{n-1} x_i^{n-1} = y_i, \qquad i = 1, 2, \ldots, m$$

In generale, se $m > n$, questo sistema non ha soluzioni. Le soluzioni ai minimi quadrati di questo sistema lineare forniscono i polinomi che approssimano meglio l'insieme dei dati osservati.

4. **Esempio**

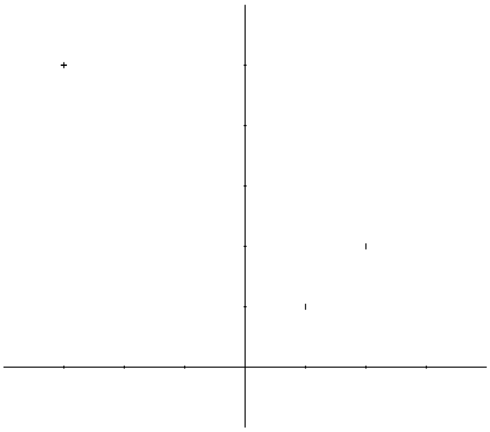
Determinare la retta

$$y = a + bx$$

che meglio approssima nel senso dei minimi quadrati l'insieme dei dati

x	y
1	1
2	2
-3	5

Il problema non ha una soluzione esatta, in quanto i punti $(1, 1), (2, 2), (-3, 5)$ non sono allineati.



Imponendo che la retta passi dai punti dati, si ha il sistema lineare

$$\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & -3 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 5 \end{bmatrix}.$$

Le due colonne della matrice dei coefficienti sono linearmente indipendenti, dunque il sistema ha una ed una sola soluzione ai minimi quadrati, data dal coefficiente di Fourier della colonna dei termini noti rispetto alla matrice dei coefficienti:

$$\begin{aligned} \begin{bmatrix} a^* \\ b^* \end{bmatrix} &= \left(\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & -3 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & -3 \end{bmatrix} \right)^{-1} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & -3 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 5 \end{bmatrix} \\ &= \begin{bmatrix} 3 & 0 \\ 0 & 14 \end{bmatrix}^{-1} \begin{bmatrix} 8 \\ -10 \end{bmatrix} \\ &= \begin{bmatrix} \frac{8}{3} \\ -\frac{5}{7} \end{bmatrix}. \end{aligned}$$

Cosi' c'e' una ed una sola retta che meglio approssima l'insieme dei tre punti dati nel senso dei minimi quadrati: la retta di equazione

$$y = \frac{8}{3} - \frac{5}{7}x.$$