Matematica II 22.12.10

Indipendenza lineare ed ortogonalita'

1. Nel piano $\mathbb{R}^2$ si ha che due vettori non nulli fra loro ortogonali sono necessariamente non allineati, e dunque sono linearmente indipendenti.

   Nello spazio $\mathbb{R}^3$ si ha ancora che due vettori non nulli fra loro ortogonali sono necessariamente non allineati, e dunque sono linearmente indipendenti; inoltre, si ha che tre vettori non nulli fra loro ortogonali sono necessariamente non complanari, e dunque sono linearmente indipendenti.

   Piu' in generale, si ha

   Teorema 1. *Siano $v_1, \ldots, v_m$ vettori non nulli in $\mathbb{R}^n$. Se $v_i \perp v_j$ per ogni $i \neq j$, allora $v_1, \ldots, v_m$ sono linearmente indipendenti.*

   Dim. Consideriamo l'uguaglianza

   $$x_1v_1 + x_2v_2 + \cdots + x_mv_m = 0_n, \qquad x_i \in \mathbb{R}.$$

   Da questa uguaglianza deduciamo l'uguaglianza

   $$v_1'(x_1v_1 + x_2v_2 + \cdots + x_mv_m) = v_1'0_n,$$

   cioe'

   $$x_1(v_1'v_1) + x_2(v_1'v_2) + \cdots + x_m(v_1'v_m) = 0;$$

   poiche' $v_1 \perp v_j$ per ogni $j = 2, \ldots, m$, si ha

   $$x_1(v_1'v_1) = 0,$$

   ed essendo $v_1 \neq 0_n$, si ha

   $$x_1 = 0.$$

   In modo analogo si prova che

   $$x_2 = 0, \ \ldots, \ x_m = 0.$$

2. Nello spazio $\mathbb{R}^3$ e' chiaro che ogni sottospazio V di dimensione 2 (piano per O) si puo' rappresentare come il sottospazio generato da due vettori non nulli fra loro ortogonali.

   Supponiamo che $V = \langle a_1, a_2 \rangle$, il sottospazio generato da due vettori linearmente indipendenti a_1, a_2. Possiamo costruire due vettori non nulli b_1, b_2 fra loro ortogonali che generino V prendendo

   $$b_1 = a_1$$
   $$b_2 = a_2 - a_1\frac{a_1'a_2}{a_1'a_1}.$$

Osserviamo che certamente $b_1, b_2 \in V$, inoltre

$$\begin{aligned} b_1' b_2 &= a_1' \left(a_2 - a_1 \frac{a_1' a_2}{a_1' a_1} \right) \\ &= a_1' a_2 - a_1' a_1 \frac{a_1' a_2}{a_1' a_1} \\ &= a_1' a_2 - a_1' a_2 \ = 0, \end{aligned}$$

cioe' $b_1 \perp b_2$.

3. Sappiamo che ogni sottospazio V di dimensione m in $\mathbb{R}^n$ si puo' rappesentare come il sottospazio $V = \langle a_1, \ldots, a_m \rangle$ generato da un insieme di vettori $a_1, \ldots, a_m$ linearmente indipendenti. Possiamo costruire un insieme di m vettori non nulli $b_1, \ldots, b_m$ fra loro ortogonali che generino V prendendo

$$\begin{aligned} b_1 &= a_1 \\ b_2 &= a_2 - b_1 \frac{b_1' a_2}{b_1' b_1} \\ b_3 &= a_3 - b_1 \frac{b_1' a_3}{b_1' b_1} - b_2 \frac{b_2' a_3}{b_2' b_2} \\ &\vdots \\ b_m &= a_m - b_1 \frac{b_1' a_m}{b_1' b_1} - b_2 \frac{b_2' a_m}{b_2' b_2} - \cdots - b_{m-1} \frac{b_{m-1}' a_m}{b_{m-1}' b_{m-1}} \end{aligned}$$

Osserviamo che $b_1, \ldots, b_m \in V$; come fatto sopra, si prova che $b_1 \perp b_2$; inoltre

$$\begin{aligned} b_1' b_3 &= b_1' \left(a_3 - b_1 \frac{b_1' a_3}{b_1' b_1} - b_2 \frac{b_2' a_3}{b_2' b_2} \right) \\ &= b_1' a_3 - b_1' b_1 \frac{b_1' a_3}{b_1' b_1} - b_1' b_2 \frac{b_2' a_3}{b_2' b_2} \\ &= b_1' a_3 - b_1' a_3 - 0 \ \frac{b_2' a_3}{b_2' b_2} \ = 0, \end{aligned}$$

cosi' $b_1 \perp b_3$; in modo analogo si prova che $b_i \perp b_j$ per ogni $i \neq j$.

Il processo sopra descritto viene detto *processo di ortogonalizzazione di Gram-Schmidt.*

Minimi quadrati

1. Consideriamo il sistema lineare

$$\begin{cases} 2x = 3 \\ 3x = 5 \\ 4x = 9 \end{cases} .$$

Questo sistema non ha soluzioni.

Per ogni valore $s \in \mathbb{R}$ dell'incognita x, possiamo considerare il vettore degli scarti fra i primi e i secondi membri

$$E(s) = \begin{bmatrix} 2s - 3 \\ 3s - 5 \\ 4s - 9 \end{bmatrix}$$

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

$$\|E(s)\| = \sqrt{(2s-3)^2 + (3s-5)^2 + (4s-9)^2}.$$

Ad esempio, si ha

$$\|E(\frac{3}{2})\| = \sqrt{(2\frac{3}{2} - 3)^2 + (3\frac{3}{2} - 5)^2 + (4\frac{3}{2} - 9)^2} = \sqrt{\frac{37}{4}},$$

$$\|E(\frac{5}{3})\| = \sqrt{(2\frac{5}{3} - 3)^2 + (3\frac{5}{3} - 5)^2 + (4\frac{5}{3} - 9)^2} = \sqrt{\frac{50}{9}}.$$

Poiche'

$$\sqrt{\frac{50}{9}} < \sqrt{\frac{37}{4}},$$

siamo condotti a preferire $s = \frac{5}{3}$ a $s = \frac{3}{2}$. 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 $s \in \mathbb{R}^n$ dell'incognita x, 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 $s^* \in \mathbb{R}^n$ cui corrisponde un errore di norma minima, cioe' tale che

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

si dice *soluzione ai minimi quadrati* del sistema.

Teorema 2. *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**

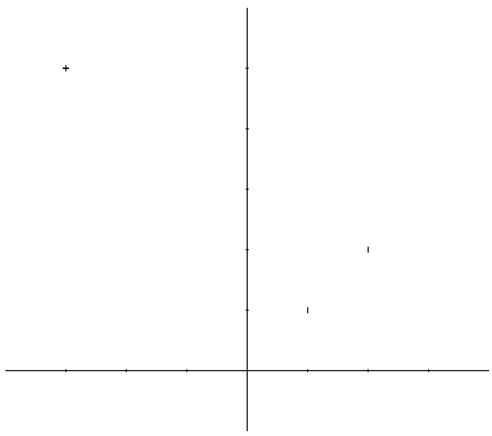
Determinare la retta

$$y = a + bx$$

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

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