Matematica II 12.05.06

1. Consideriamo un sistema lineare di m equazioni in n incognite
$$\begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ \vdots \\ b_m \end{bmatrix},$$
sinteticamente

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

   Ricordiamo che una soluzione esatta del sistema e' un vettore $s \in R^n$ tale che la sostituzione $x = s$ renda il primo membro uguale al secondo:

   $As = b.$

   Se non ci sono soluzioni esatte, possiamo comunque, per ciascun $t \in R^n$, valutare il primo membro Ax in $x = t$, ottenendo un vettore di $At \in R^m$; avremo $At \neq b$, cioe' $At - b \neq 0$. Diciamo che il vettore

   $At - b \in R^m$

   e' l'errore che si commette assumendo il vettore t come soluzione, in breve, l'errore corrispondente a t.

   Un vettore $s \in R^n$ cui corrisponde un errore di norma minima, cioe' tale che

   $\|As - b\| \leq \|At - b\|, \qquad \forall t \in R^n$

   viene detto *soluzione ai minimi quadrati del sistema* $Ax = b$.

   Questa condizione puo' essere scritta anche nella forma

   $d(b, As) \leq d(b, At), \qquad \forall t \in R^n$

   e interpretata dicendo che il vettore $s \in R^n$ rende il primo membro Ax il piu' vicino possibile al secondo membro b.

   La ricerca delle soluzioni ai minimi quadrati equivale alla ricerca dei punti di minimo della funzione di n variabili $\|Ax - b\|$, o, che e' lo stesso, dei punti di minimo della funzione $\|Ax - b\|^2$:

   $(a_{11}x_1 + \dots + a_{1n}x_n - b_1)^2 + \dots + (a_{m1}x_1 + \dots + a_{mn}x_n - b_1)^2.$

2. **Esempi**

   - Il sistema lineare di tre equazioni in una incognita
   $$\begin{cases} x = & 4 \\ 2x = & 5 \\ 3x = & 6 \end{cases}, \qquad \textit{sinteticamente} \quad ax = b,$$

non ha soluzioni; l'errore che si commette assumendo che $t \in R$ sia una soluzione, in breve, l'errore corrispondente a t e'

$$at - b = \begin{bmatrix} t-4 \\ 2t-5 \\ 3t-6 \end{bmatrix},$$

che ha norma

$$\|at - b\| = \sqrt{(t-4)^2 + (2t-5)^2 + (3t-6)^2}.$$

Le soluzioni ai minimi quadrati di $ax = b$ sono i punti di minimo della funzione di una variabile

$$\|ax - b\|^2 = (x-4)^2 + (2x-5)^2 + (3x-6)^2.$$

- Il sistema lineare di tre equazioni in due incognite

$$\begin{cases} x_1 + x_2 = & 3 \\ x_1 + 2x_2 = & 4 \\ 2x_2 = & 5 \end{cases}, \qquad \textit{sinteticamente} \quad Ax = b,$$

non ha soluzioni; l'errore che si commette assumendo che $t = \begin{bmatrix} t_1 \\ t_2 \end{bmatrix} \in R^2$ sia una soluzione, in breve, l'errore corrispondente a t e'

$$At - b = \begin{bmatrix} t_1 + t_2 - 3 \\ t_1 + 2t_2 - 4 \\ 2t_2 - 5 \end{bmatrix},$$

che ha norma

$$\|At - b\| = \sqrt{(t_1 + t_2 - 3)^2 + (t_1 + 2t_2 - 4)^2 + (2t_2 - 5)^2}.$$

Le soluzioni ai minimi quadrati di $Ax = b$ sono i punti di minimo della funzione di due variabili

$$\|ax - b\|^2 = (x_1 + x_2 - 3)^2 + (x_1 + 2x_2 - 4)^2 + (2x_2 - 5)^2.$$

La ricerca delle soluzioni ai minimi quadrati di un sistema lineare puo' essere condotta coi metodi dell'Analisi per lo studio delle funzioni di piu' variabili; i realta', i risultati illustrati nelle lezioni scorse sulle proiezioni ortogonali permettono di assicurare l'esistenza di soluzioni ai minimi quadrati e di dare formule o procedimenti per la loro determinazione.

3. Consideriamo il generico sistema lineare di m equazioni in una incognita x

$$\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix} \begin{bmatrix} x \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix},$$

sinteticamente

$$ax = b, \qquad a \in R^m, \quad x \in R, \quad b \in R^m,$$

e supponiamo che $a \neq 0$.

L'espressione a primo membro

ax

descrive, al variare di $x \in R$, una retta l passante per il vettore nullo 0. Il sistema avra' soluzioni se e solo se il vettore b dei termini noti giace sulla retta l; in questo caso, b si potra' scrivere in uno ed un solo modo come multiplo

$as = b$

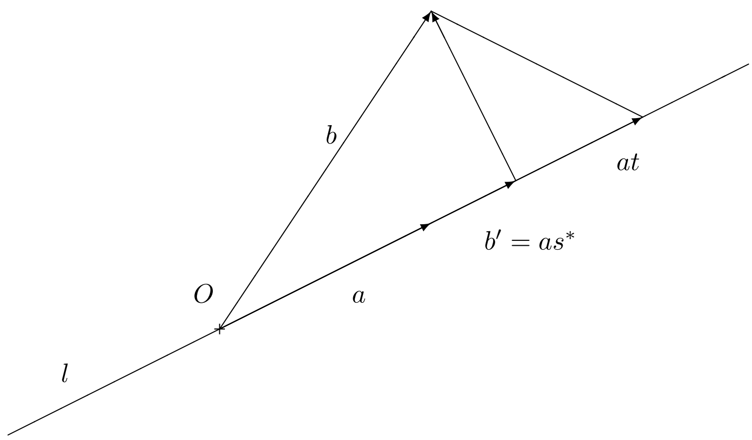
di a, e il peso s dara' la soluzione del sistema. In generale, per $m > 1$, dobbiamo attenderci che $b \notin l$.

Possiamo allora considerare, fra i vettori di l, la proiezione ortogonale

$b' = as^*, \qquad s^* = \frac{\langle a,b \rangle}{\langle a,a \rangle}$

di b su l, definita dalla condizione

$b - b' \perp a$.



Ricordiamo che $b' = as^*$ e', fra i vettori della retta l, quello piu' vicino a b, nel senso che

$d(b, as^*) \leq d(b, at), \qquad \forall t \in R$

$d(b, as^*) = d(b, at) \qquad \textit{solo se} \quad as^* = at.$

Possiamo allora riconoscere che

- *Il sistema lineare*

$$ax = b, \qquad a \in R^m, \quad x \in R, \quad b \in R^m,$$

sotto la condizione $a \neq 0$, ha una ed una sola soluzione ai minimi quadrati: il coefficiente di Fourier

$$s^* = \frac{\langle a, b\rangle}{\langle a, a\rangle}$$

del vettore b dei termini noti rispetto al vettore a dei coefficienti, che a sua volta e' l'unico scalare cui corrisponde un errore ortogonale alla colonna dei coefficienti:

$$as^* - b \perp a.$$

4. Consideriamo ora il generico sistema lineare di m equazioni in n incognite

$$\begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ \vdots \\ b_m \end{bmatrix},$$

sinteticamente

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

e supponiamo che le colonne $a_1, a_2, \dots, a_n \in R^m$ di A siano linearmente indipendenti, e dunque che $n \leq m$.

L'espressione a primo membro

$Ax = a_1x_1 + \cdots + a_nx_n$

descrive, al variare di $x \in R^n$, lo spazio S generato in R^m dalle n colonne a_i della matrice A.

Il sistema avra' soluzioni se e solo se il vettore b dei termini noti appartiene allo spazio S; in questo caso, b si potra' scrivere in uno ed un solo modo come combinazione lineare

$a_1s_1 + \cdots + a_ns_n = b$

delle colonne a_i, e i pesi s_i di questa combinazione lineare daranno le componenti della soluzione s del sistema.

In generale, per $m > n$, dobbiamo attenderci che $b \notin S$.

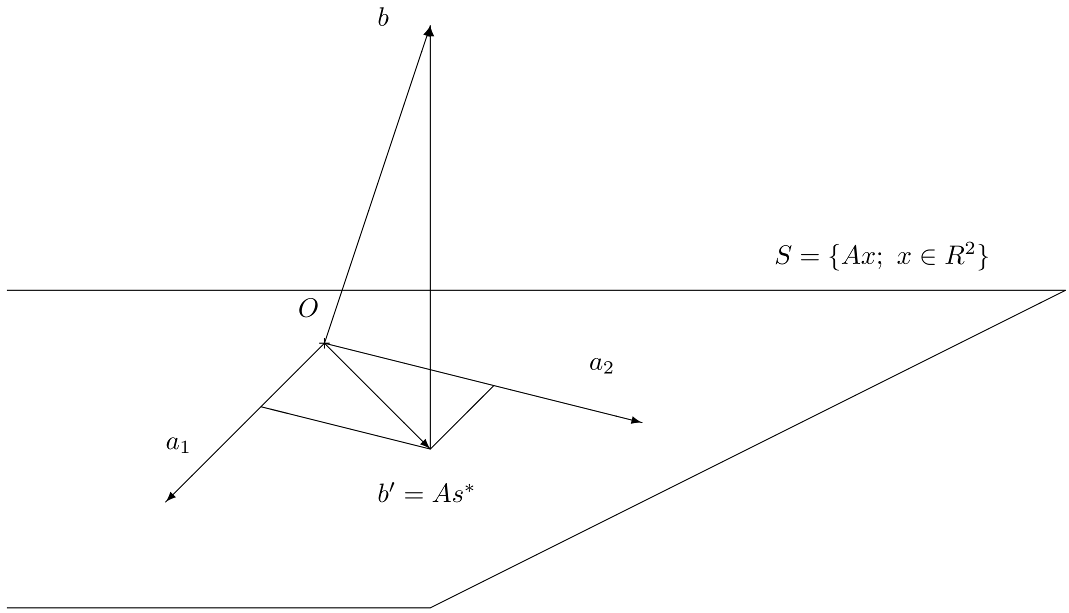
Possiamo allora considerare, fra i vettori di S, la proiezione ortogonale

$b' = As^*, \qquad s^* = (A^TA)^{-1}A^Tb$

di b su l, definita dalla condizione

$b - b' \perp a_i, \qquad \forall i = 1, \dots, n$

Caso $m = 3, n = 2$:



Ricordiamo che $b' = As^*$ e', fra i vettori dello spazio S, quello piu' vicino a b, nel senso che

$d(b, As^*) \leq d(b, At), \qquad \forall t \in R^n$

$d(b, As^*) = d(b, At) \qquad$ *solo se* $\quad As^* = At.$

Possiamo allora riconoscere che

- *Il sistema lineare*

$$Ax = b, \qquad A = [a_1 \dots a_n] \in R^{m\times n} \quad x \in R^n, \quad b \in R^m,$$

  sotto la condizione $a_1, \dots, a_n$ linearmente indipendenti, ha una ed una sola soluzione ai minimi quadrati: il coefficiente di Fourier

$$s^* = (A^T A)^{-1} A^T b$$

  del vettore b dei termini noti rispetto alla matrice A dei coefficienti, cioe' l'unico vettore di R^n cui corrisponde un errore ortogonale alle colonne dei coefficienti:

$$As^* - b \perp a_1, \dots, a_n.$$

In generale si ha

- *Ogni sistema lineare*

$$Ax = b, \qquad A = [a_1 \dots a_n] \in R^{m \times n} \quad x \in R^n, \quad b \in R^m,$$

ha soluzioni ai minimi quadrati, e sono i vettori s^* di R^n cui corrisponde un errore ortogonale alle colonne dei coefficienti:

$$As^* - b \perp a_1, \dots, a_n,$$

in altri termini, le soluzioni esatte del sistema lineare

$$A^T A x = A^T b.$$

5. Per esercizio, per ciascuno dei sistemi lineari del punto 2, si determinino le soluzioni ai minimo quadrati, si controlli la correttezza del risultato con la condizione di ortogonalita', e si determini l'errore corrispondente.