# Algebra/Algebra Lineare, 12.03.08

1. **Soluzioni ai minimi quadrati, primo esempio.**

   Consideriamo il sistema lineare di tre equazioni in una incognita

   $$\begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} x = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \qquad \textit{in sintesi}, \qquad ax = b.$$

   Questo sistema e' impossibile.

   Possiamo associare ad ogni valore dell'incognita $x$ una terna $E = [\epsilon_i]_1^3$ di 'errori', ponendo

   $$\begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} x + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \qquad \textit{in sintesi}, \qquad ax + E = b,$$

   e possiamo misurare l'errore $E = b - ax = E(x)$ con la sua norma:

   $$\|E(x)\| = \|b - ax\| = \| \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} x\| = \sqrt{(1-2x)^2 + (1-3x)^2 + (1-4x)^2}.$$

   Un valore $\overline{x}$ dell'incognita $x$ che minimizzi questa quantita':

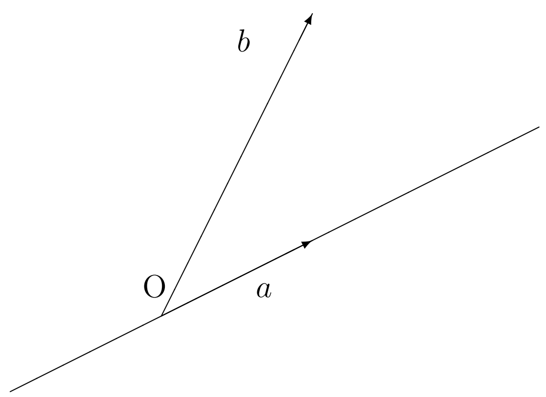
   $$\|E(\overline{x})\| = min_{x \in \mathbb{R}} \; \|E(x)\|,$$

   viene detto *soluzione ai minimi quadrati del sistema.*

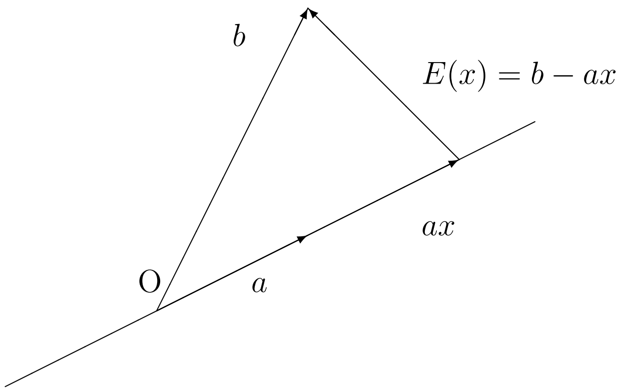
   Geometricamente, il sistema

   $$ax = b$$

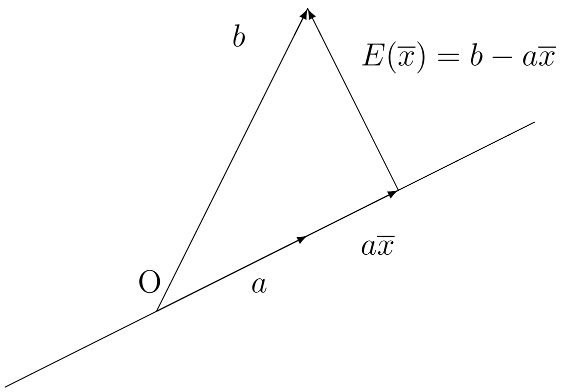
   puo' essere letto come la posizione del problema di rappresentare il vettore $b$ come un multiplo scalare del vettore $a$. Questo problema ha soluzione se e solo se il vettore $b$ sta sulla retta generata da $a$. In generale ci si aspetta che, come nel nostro esempio, cio' non accada.



L'errore $E(x) = b-ax$ puo' essere rappresentato come il vettore con punto iniziale nel punto $ax$ e con punto finale nel punto $b$; la norma $\|E(x)\| = \|b-ax\|$ di questo errore e' la distanza fra il punto $ax$ e il punto $b$; un valore $\overline{x}$ dell'incognita $x$ e' una soluzione ai minimi quadrati del sistema se e solo se il punto $a\overline{x}$ ha, fra i punti della retta generata da $a$, distanza minima dal punto $b$.




Ora, il punto $a\overline{x}$ ha, fra i punti della retta generata da $a$, distanza minima dal punto $b$ se e solo se il vettore $b-a\overline{x} = E(\overline{x})$ e' ortogonale ad $a$; la norma di questo errore e' la distanza del punto $b$ dalla retta, e $a\overline{x}$ e' la proiezione ortogonale del punto $b$ sulla retta.




Dunque si ha che $\overline{x}$ e' una soluzione ai minimi quadrati del sistema se e solo se

$$a^T(b - a\overline{x}) = 0, \qquad a^T a\overline{x} = a^T b.$$

Ora, questa e' un'equazione lineare in una incognita, che compare con coefficiente $a^T a \neq 0$; dunque si ha una ed una sola soluzione ai minimi quadrati, data da

$$\overline{x} = \frac{a^T b}{a^T a}.$$

Nel nostro caso si ha

$$\overline{x} = \frac{a^T b}{a^T a} = \frac{\begin{bmatrix} 2 & 3 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}}{\begin{bmatrix} 2 & 3 & 4 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}} = \frac{9}{29},$$

e l'errore corrispondente e' misurato da

$$\|E(\overline{x})\| = \|b - a\overline{x}\| = \| \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} \frac{9}{29} \| = \| \begin{bmatrix} 11 \\ 2 \\ -7 \end{bmatrix} \frac{1}{29} \| = \frac{\sqrt{174}}{29}.$$

2. **Soluzioni ai minimi quadrati, secondo esempio.**

Consideriamo il sistema lineare di tre equazioni in due incognite

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} x_1 + \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} x_2 = \begin{bmatrix} 1 \\ 2 \\ 8 \end{bmatrix}, \qquad \textit{in sintesi}, \qquad Ax = a_1x_1 + a_2x_2 = b.$$

Questo sistema e' impossibile.

Ad ogni valore dell'incognita $x \in \mathbb{R}^2$ possiamo associare una terna $E = [\epsilon_i]_1^3$ di errori, definita dalla relazione

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 8 \end{bmatrix}, \qquad \textit{in sintesi}, \qquad Ax + E = b,$$

e possiamo misurare l'errore $E(x) = b - Ax$ con la sua norma:

$$\|E(x)\| = \|b - Ax\| = \| \begin{bmatrix} 1 \\ 2 \\ 8 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \| =$$

$$= \sqrt{(1 - x_1)^2 + (2 - x_2)^2 + (8 - x_1 - x_2)^2}.$$

Un valore $\overline{x} \in \mathbb{R}^2$ che minimizzi questa quantita':

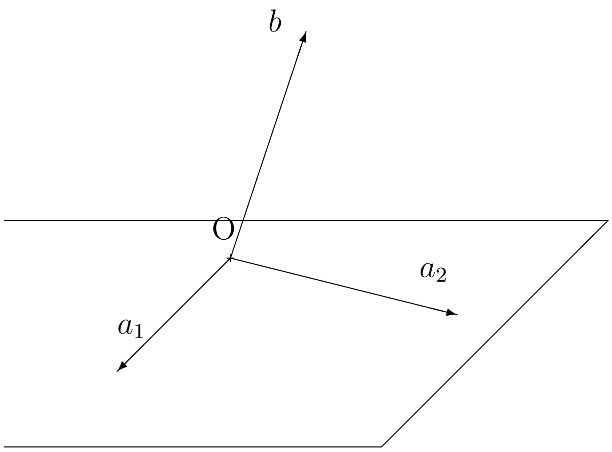
$$\|E(\overline{x})\| = min_{x \in \mathbb{R}^2} \ \|E(x)\|,$$

viene detto *soluzione ai minimi quadrati del sistema.*

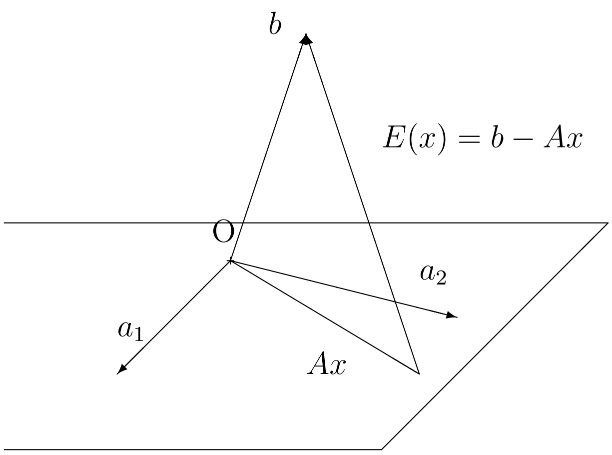
Geometricamente, il sistema

$$Ax = a_1x_1 + a_2x_2 = b$$

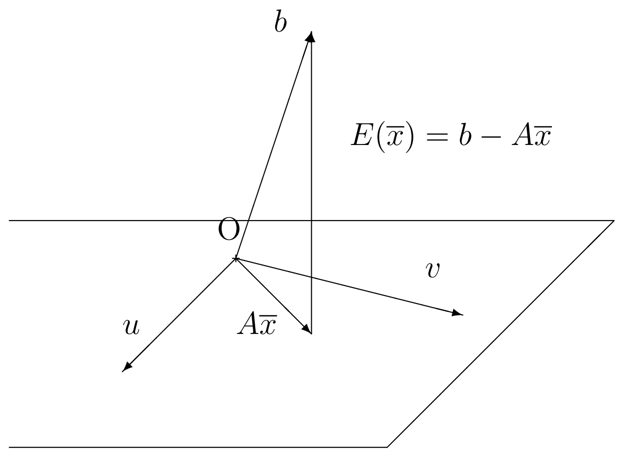
puo' essere letto come la posizione del problema di rappresentare il vettore $b$ come combinazione lineare dei vettori $a_1, a_2$. Questo problema ha soluzione se e solo se il vettore $b$ sta sul piano generato da $a_1, a_2$. In generale ci si aspetta che, come nel nostro esempio, cio' non accada.




L'errore $E(x) = b - Ax$ puo' essere rappresentato come il vettore con punto iniziale nel punto $Ax$ e con punto finale nel punto $b$; la norma $\|E(x)\| = \|b - Ax\|$ di questo errore e' la distanza fra il punto $Ax$ e il punto $b$; un valore $\overline{x}$ dell'incognita $x$ e' una soluzione ai minimi quadrati del sistema se e solo se il punto $A\overline{x}$ ha, fra i punti del piano generato da $a_1$ e $a_2$, distanza minima dal punto $b$.




Ora, il punto $A\overline{x}$ ha, fra i punti del piano generato da $a_1$ e $a_2$, distanza minima dal punto $b$ se e solo se il vettore $b - A\overline{x} = E(\overline{x})$ e' ortogonale ad $a_1$ e $a_2$; la norma di questo errore e' la distanza del punto $b$ dal piano, e $A\overline{x}$ e' la proiezione ortogonale del punto $b$ sul piano.




Dunque si ha che $\overline{x}$ e' una soluzione ai minimi quadrati del sistema se e solo se

$a_1^T(b - A\overline{x}) = 0, \qquad a_1^T A\overline{x} = a_1^T b.$

$a_2^T(b - A\overline{x}) = 0, \qquad a_2^T A\overline{x} = a_2^T b.$

cioe'

$A^T(b - A\overline{x}) = 0_{2\times 1}, \qquad A^T A\overline{x} = A^T b.$

Ora, essendo le colonne $a_1, a_2$ linearmente indipendenti, la matrice $A^T A$ e' invertibile, e cosi' si ha una ed una sola soluzione ai minimi quadrati, data da

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

Nel nostro caso si ha

$$\begin{aligned}
\overline{x} = \left(A^T A\right)^{-1} A^T b &= \left( \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix} \right)^{-1} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 8 \end{bmatrix} \\
&= \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}^{-1} \begin{bmatrix} 9 \\ 10 \end{bmatrix} \\
&= \frac{1}{3} \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 9 \\ 10 \end{bmatrix} \\
&= \frac{1}{3} \begin{bmatrix} 8 \\ 11 \end{bmatrix},
\end{aligned}$$

e l'errore corrispondente e' misurato da

$$\|\overline{E}(\overline{x})\| = \|b - A\overline{x}\| = \| \begin{bmatrix} 1 \\ 2 \\ 8 \end{bmatrix} - \frac{1}{3} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 8 \\ 11 \end{bmatrix} \| = \| \frac{1}{3} \begin{bmatrix} -5 \\ -5 \\ 5 \end{bmatrix} \| = \frac{5\sqrt{3}}{3}.$$

3. **Soluzioni ai minimi quadrati.**

Consideriamo il generico sistema lineare di $m$ equazioni in $n$ incognite

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

in sintesi,

$$Ax = a_1 x_1 + \dots + a_n x_n = b.$$

Ad ogni valore dell'incognita $x \in \mathbb{R}^n$ possiamo associare una $n-$pla $E = [\epsilon_i]_1^n$ di errori, definita dalla relazione

$$\begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \vdots \\ \vdots \\ \epsilon_m \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ \vdots \\ b_m \end{bmatrix},$$

in sintesi,

$$Ax + E = b,$$

e possiamo misurare l'errore $E(x) = b - Ax$ con la sua norma:

$$\|E(x)\| = \|b - Ax\| = \| \begin{bmatrix} b_1 \\ \vdots \\ \vdots \\ b_m \end{bmatrix} - \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \| =$$

$$= \sqrt{(b_1 - \sum_{j=1}^n a_{1j} x_j)^2 + \dots + (b_m - \sum_{j=1}^n a_{mj} x_j)^2}.$$

Un valore $\overline{x} \in \mathbb{R}^n$ che minimizzi questa quantita':

$$\|E(\overline{x})\| = min_{x \in \mathbb{R}^n} \; \|E(x)\|,$$

viene detto *soluzione ai minimi quadrati del sistema.*

Si puo' provare che

> *Un valore $\overline{x} \in \mathbb{R}^n$ e' una soluzione ai minimi quadrati del sistema $Ax = b$ se e solo se l'errore $E(\overline{x}) = b - A\overline{x}$ e' ortogonale ad ogni colonna di $A$ :*
>
> $$A^T(b - A\overline{x}) = 0_{n \times 1}, \qquad A^T A \overline{x} = A^T b.$$

*Questo e' un sistema lineare di n equazioni in n incognite, dette equazioni normali del sistema dato, che ha sempre soluzioni.*

*Se le colonne $a_1, \ldots, a_n$ della matrice $A$ sono linearmente indipendenti, allora la matrice $A^T A$ e' invertibile, e*

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

*e' l'unica soluzione ai minimi quadrati del sistema.*

Si osservi che la matrice $A^T A$ e' la tabella dei prodotti interni delle colonne di $A$ :

$$A^T A = [a_1 \ \ldots \ a_n]^T [a_1 \ \ldots \ a_n] = \begin{bmatrix} a_1^T \\ \vdots \\ a_n^T \end{bmatrix} [a_1 \ \ldots \ a_n] = \begin{bmatrix} a_1^T a_1 & \ldots & a_1^T a_n \\ \vdots & & \\ a_n^T a_1 & \ldots & a_n^T a_n \end{bmatrix},$$

e che $A^T b$ e' la colonna dei prodotti interni delle colonne di $A$ con la colonna termine noto $b$ :

$$A^T b = [a_1 \ \ldots \ a_n]^T b = \begin{bmatrix} a_1^T \\ \vdots \\ a_n^T \end{bmatrix} b = \begin{bmatrix} a_1^T b \\ \vdots \\ a_n^T b \end{bmatrix}.$$

La formula per l'unica soluzione ai minimi quadrati, nel caso delle colonne linearmente indipendenti, si puo' allora espandere nella forma

$$\overline{x} = \begin{bmatrix} a_1^T a_1 & \ldots & a_1^T a_n \\ \vdots & & \\ a_n^T a_1 & \ldots & a_n^T a_n \end{bmatrix}^{-1} \begin{bmatrix} a_1^T b \\ \vdots \\ a_n^T b \end{bmatrix}.$$

Nel caso $n = 1$, si ha che il generico sistema lineare di $m$ equazioni in una incognita

$$ax = b, \qquad \textit{con } a \neq 0$$

ha una ed una sola soluzione ai minimi quadrati, data da

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