<u>Método de los mínimos cuadrados</u>

(Ver página 226 del libro de teoría)

Supongamos un sistema, $A\mathbf{x} = \mathbf{b}$, que no tiene solución.

Por ejemplo, este es un sistema que no tiene solución.

$$\begin{cases} 2x+3y=5 \\ 3x+2y=-2 \\ 5x+8y=9 \\ 4x-y=10 \end{cases} \quad \leftrightarrow \quad \begin{pmatrix} 2 & 3 \\ 3 & 2 \\ 5 & 8 \\ 4 & -1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 5 \\ -2 \\ 9 \\ 10 \end{pmatrix} \quad \leftrightarrow$$

$$\leftrightarrow \quad x\begin{pmatrix} 2 \\ 3 \\ 5 \\ 4 \end{pmatrix} + y\begin{pmatrix} 3 \\ 3 \\ 8 \\ -1 \end{pmatrix} = \begin{pmatrix} 5 \\ -2 \\ 9 \\ 10 \end{pmatrix} \quad \leftrightarrow \quad xC_1 + yC_2 = B$$

C_1 y C_2 son las columnas de la matriz, A, del sistema. No hay ningún (x, y) que verifique la igualdad.

El problema de resolver el sistema $A\mathbf{x} = \mathbf{b}$ por el método de los mínimos cuadrados consiste en encontrar el vector $\mathbf{x_0}$ que mejor aproxima una solución del sistema, en el sentido de que hace mínima la distancia entre $A\mathbf{x_0}$ y $\mathbf{b}$ .Es decir, se trata de encontrar el valor $\mathbf{x_0}$ que hace mínimo

$$\|A\mathbf{x} - \mathbf{b}\|$$

En nuestro ejemplo, el problema consiste en halla un vector $\begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$ que haga que

$$A\begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = x_0C_1 + y_0C_2 \approx B \ \text{(lo más aproximado posible)}$$

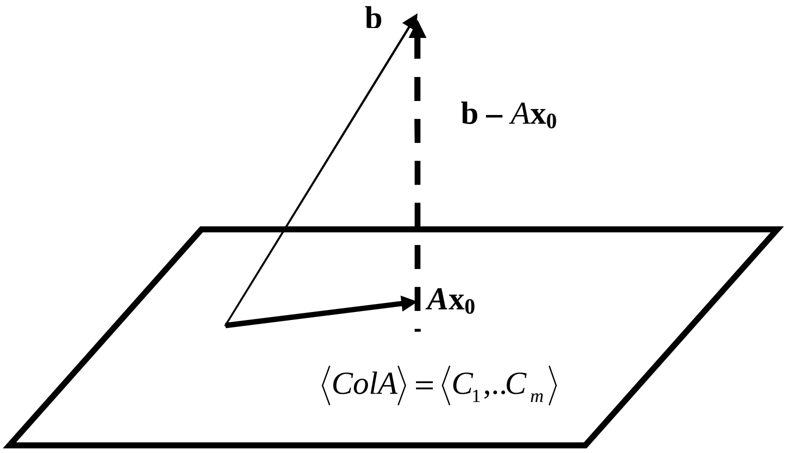
Obsérvese que el vector del primer miembro es una combinación lineal de las columnas de la matriz A.

$$x_0C_1 + y_0C_2 \in \langle C_1, C_2 \rangle$$

Así pues, el problema de hallar la solución de los mínimos cuadrados se puede formular como hallar el vector $A\mathbf{x_0}$ contenido en el espacio generado por las columnas de A que mejor aproxima al vector $\mathbf{b}$.

En nuestro ejemplo, se trata de hallar el vector $x_0C_1 + y_0C_2 \in \langle C_1, C_2 \rangle$ que mejor aproxima al vector B.

Teniendo en cuenta que la mejor aproximación es la proyección ortogonal



Hay que hallar el vector $\mathbf{x_0}$ tal que

$$\text{Proy}_{<\text{Col } A>}\, \mathbf{b} = A\mathbf{x_0}$$

Lo que significa que el vector ($\mathbf{b} - A\mathbf{x_0}$) es ortogonal a todas las columnas de la matriz A

Es decir, el producto escalar de **b** por cada una de las columnas de A es cero

$$C_i^{\mathrm{t}} \cdot (B - AX_0) = 0$$

Todas estas igualdades escritas en una sola fórmula es

$$A^t\left(B - AX_0\right) = 0 \text{ (vector columna 0)} \quad \leftrightarrow \quad \boxed{A^t AX_0 = A^t B}$$

A estas ecuaciones se denominan *ecuaciones normales*

En el ejemplo del principio, resulta esto

$$\begin{pmatrix} 2 & 3 & 5 & 4 \\ 3 & 2 & 8 & -1 \end{pmatrix}\begin{pmatrix} 2 & 3 \\ 3 & 2 \\ 5 & 8 \\ 4 & -1 \end{pmatrix}\begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \begin{pmatrix} 2 & 3 & 5 & 4 \\ 3 & 2 & 8 & -1 \end{pmatrix}\begin{pmatrix} 5 \\ -2 \\ 9 \\ 10 \end{pmatrix}$$

$$\begin{pmatrix} 54 & 48 \\ 48 & 78 \end{pmatrix}\begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \begin{pmatrix} 89 \\ 73 \end{pmatrix}$$

La solución es x_0 = 191/106 = 1,80; y_0 = –55/318 = –0,17

En este ejemplo, el error cometido es

$$\|A\mathbf{x_0} - \mathbf{b}\| = \left\| \begin{pmatrix} 2 & 3 \\ 3 & 2 \\ 5 & 8 \\ 4 & -1 \end{pmatrix} \begin{pmatrix} 1{,}80 \\ -0{,}17 \end{pmatrix} - \begin{pmatrix} 5 \\ -2 \\ 9 \\ 10 \end{pmatrix} \right\| = \left\| \begin{pmatrix} 3{,}08 \\ 5{,}06 \\ 7{,}64 \\ 7{,}37 \end{pmatrix} - \begin{pmatrix} 5 \\ -2 \\ 9 \\ 10 \end{pmatrix} \right\| = \left\| \begin{pmatrix} -1{,}92 \\ 7{,}06 \\ -1{,}36 \\ -2{,}63 \end{pmatrix} \right\| = =$$

$$= \sqrt{(-1{,}92)^2 + (7{,}06)^2 + (-1{,}36)^2 + (-2{,}63)^2} = 7{,}8$$

Archivo Editar Celda Maxima Ecuaciones Álgebra Análisis Simplificar Gráficos Numérico Ayuda

(%i1) A:matrix([2,3],[3,2],[5,8],[4,-1]);

(%o1) $\begin{bmatrix} 2 & 3 \\ 3 & 2 \\ 5 & 8 \\ 4 & -1 \end{bmatrix}$

(%i2) AT:transpose(A);

(%o2) $\begin{bmatrix} 2 & 3 & 5 & 4 \\ 3 & 2 & 8 & -1 \end{bmatrix}$

(%i3) AT.A;

(%o3) $\begin{bmatrix} 54 & 48 \\ 48 & 78 \end{bmatrix}$

(%i4) AT.matrix([5],[-2],[9],[10]);

(%o4) $\begin{bmatrix} 89 \\ 73 \end{bmatrix}$

(%i5) solve([54*x+48*y=89, 48*x+78*y=73],[x,y]);

(%o5) $[[x=\frac{191}{106}, y=-\frac{55}{318}]]$

(%i10) float (%o5);

(%o10) $[[x=1.80188679245283, y=-0.17295597484277]]$

Problema de ajustar una nube de puntos a una recta usando el método de los mínimos cuadrados (Gauss)

Comentario: Gauss utilizó el método de los mínimos cuadrados para ajustar la órbita de Ceres a partir de unas cuantas observaciones y así predecir cuando y dónde podría volver a observarse el planetoide.

Ejemplo (Ver página 234)

Imaginemos la nube de puntos del plano (2, 3), (4, 2), (5,1) y (7, 0,5).

x_i	2	4	5	7
y_i	3	2	1	0,5

Evidentemente no hay una recta $y = mx + b$ que pase por estos cuatro puntos. Es decir, no hay valores m y b que verifiquen simultáneamente

$$\begin{cases} 2m+b=3 \\ 4m+b=2 \\ 5m+b=1 \\ 7m+b=0{,}5 \end{cases} \quad \leftrightarrow \quad \begin{pmatrix} 2 & 1 \\ 4 & 1 \\ 5 & 1 \\ 7 & 1 \end{pmatrix} \begin{pmatrix} m \\ b \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \\ 1 \\ 0{,}5 \end{pmatrix}$$

Las ecuaciones normales

$$A^t A X_0 = A^t B$$

En este caso, las ecuaciones normales son

$$\begin{pmatrix} 2 & 4 & 5 & 7 \\ 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 4 & 1 \\ 5 & 1 \\ 7 & 1 \end{pmatrix} \begin{pmatrix} m_0 \\ b_0 \end{pmatrix} = \begin{pmatrix} 2 & 4 & 5 & 7 \\ 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 3 \\ 2 \\ 1 \\ 0{,}5 \end{pmatrix}$$

$$\begin{pmatrix} 94 & 18 \\ 18 & 4 \end{pmatrix} \begin{pmatrix} m_0 \\ b_0 \end{pmatrix} = \begin{pmatrix} 22{,}5 \\ 6{,}5 \end{pmatrix}$$

Si se despeja el vector de las incógnitas

$$\begin{pmatrix} m_0 \\ b_0 \end{pmatrix} = \begin{pmatrix} 94 & 18 \\ 18 & 4 \end{pmatrix}^{-1} \begin{pmatrix} 22{,}5 \\ 6{,}5 \end{pmatrix} = \begin{pmatrix} -0{,}52 \\ 3{,}96 \end{pmatrix}$$

Así pues, la ecuación de la recta de regresión, ajustada a la nube de puntos es

$$y = 3{,}96 - 0{,}52x$$

Comentario: Cuando se hace un enfoque analítico, para el cálculo de la recta de regresión, $y = mx + b$, que ajusta una nube de N puntos $\{(x_i, y_i)\}$, se busca hallar el mínimo de la función, $E(m, b)$, que es suma de los errores cuadráticos.

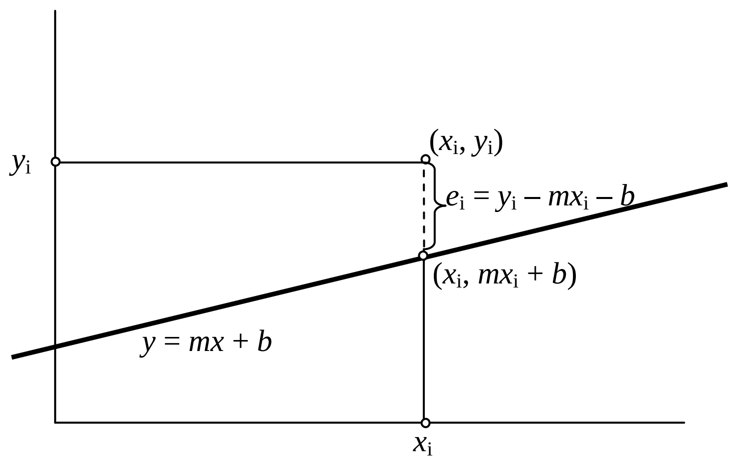
$$E(m,b) = \sum_1^N e_i^{\,2} = \sum_1^N \left(y_i - mx_i - b\right)^2$$

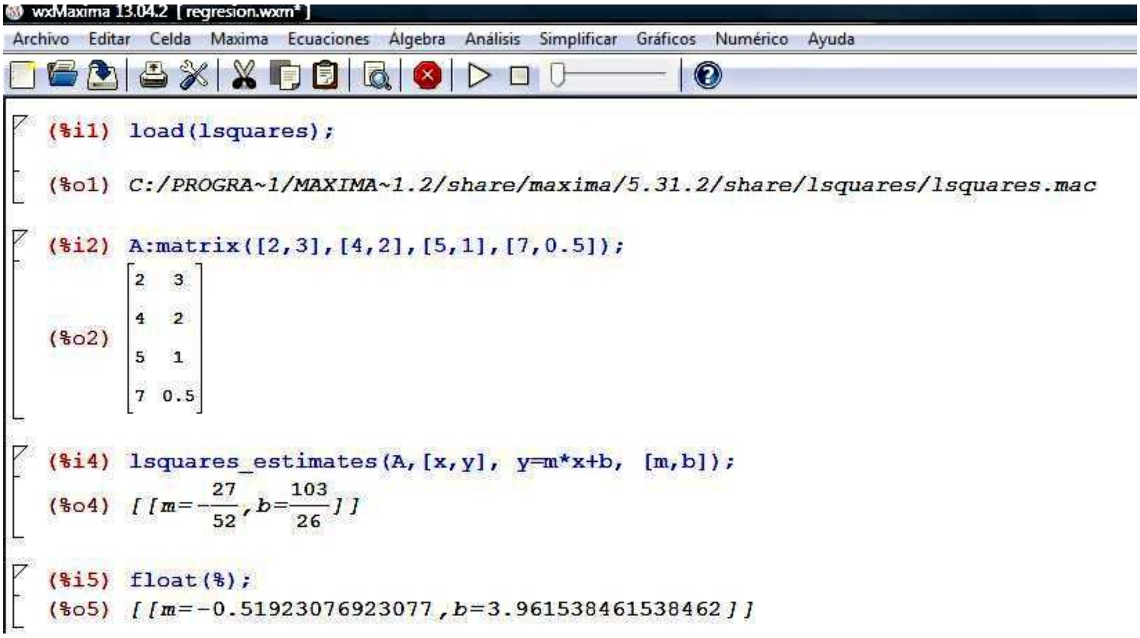
Los valores de m y b que minimizan $E(m, b)$, son las soluciones de las ecuaciones

$$\frac{\partial E}{\partial m} = 0 \qquad \rightarrow \qquad \left(\sum_1^N x_i^{\,2}\right) m + \left(\sum_1^N x_i\right) b = \sum_1^N x_i y_i$$

$$\frac{\partial E}{\partial b} = 0 \qquad \rightarrow \qquad \left(\sum_1^N x_i\right) m + Nb = \sum_1^N y_i$$

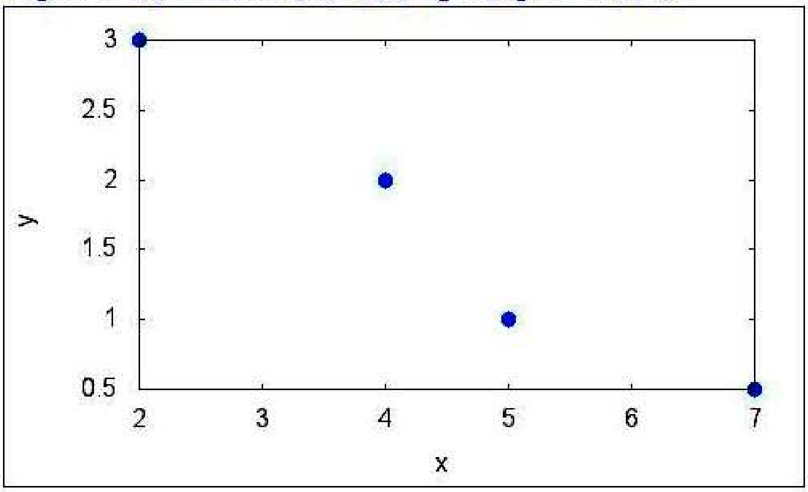
Como se puede ver, son las mismas que se han obtenido antes por métodos algebraicos.




wxMaxima 13.04.2 [regresion.wxm*]
Archivo Editar Celda Maxima Ecuaciones Álgebra Análisis Simplificar Gráficos Numérico Ayuda
(%i1) load(lsquares);
(%o1) C:/PROGRA~1/MAXIMA~1.2/share/maxima/5.31.2/share/lsquares/lsquares.mac
(%i2) A:matrix([2,3],[4,2],[5,1],[7,0.5]);
(%o2)
2 3
4 2
5 1
7 0.5
(%i4) lsquares_estimates(A,[x,y], y=m*x+b, [m,b]);
(%o4) [[m=-27/52,b=103/26]]
(%i5) float(%);
(%o5) [[m=-0.51923076923077,b=3.961538461538462]]

(%i20) wxplot2d([discrete,a,b],[style,[points]]);

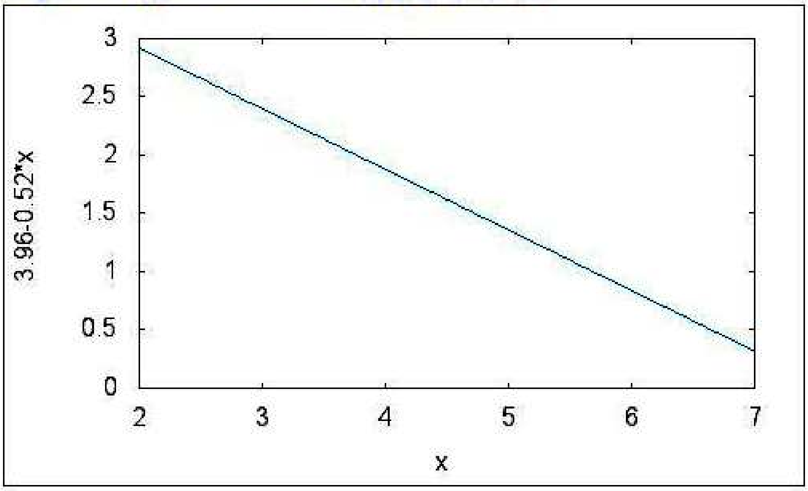
(%t20)



(%o20)

(%i24) wxplot2d([-0.52*x+3.96],[x,2,7]);

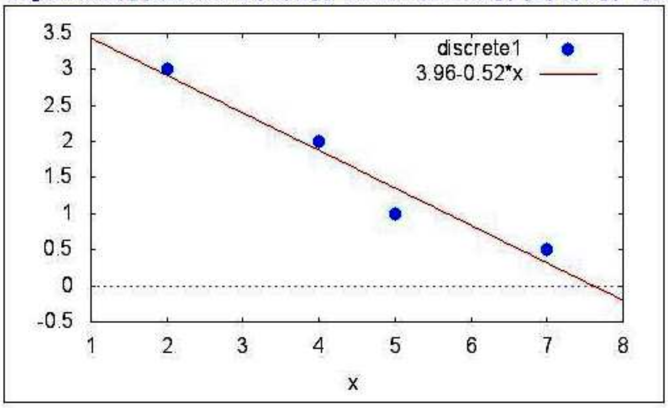
(%t24)



(%o24)

(%i33) wxplot2d([[discrete,a,b],-0.52*x+3.96],[x,1,8], [style,[points], [lines]]);

(%t33)



(%o33)

Solución del problema de los mínimos cuadrados a partir de la factorización *QR*

Recordatorio: **Factorizacion *QR***

Supongamos que la matriz A está factorizada

$$A = QR$$

Donde Q es la matriz ortonormal que se obtiene aplicando el proceso de Gram-Schmidt a las columnas de la matriz A.

Las columnas de la matriz Q forman una base ortonormal. Por tanto, $Q^tQ = I$. O lo que es lo mismo $Q^t = Q^{-1}$

La matriz R es una matriz triangular superior, con elementos positivos en la diagonal.

$$A = QR \;\;\rightarrow\;\; R = Q^{-1}A \;\;\rightarrow\;\; R = Q^tA$$

Supongamos que el sistema que hay que resolver por el método de los mínimos cuadrados es

$$AX = B$$

Supongamos $A = QR$ es su factorización QR

Las ecuaciones normales son

$$A^tAX = A^tB \;\;\rightarrow\;\; R^tQ^tQRX = R^tQ^tB \;\;\rightarrow\;\; R^tRX = R^tQ^tB \;\;\rightarrow$$

$$\boxed{RX = Q^tB}$$