Exposición final de Álgebra Lineal

Mínimos Cuadrados

Alumnas:
Natalia Maién Gomez
Milagros Vera

Profesora: Victoria Vampa
JTP: Lucia Rizzo

¿A**x**=**b** sin solución?

Aproximación

Cuando menor es la distancia entre A**x** y **b**, mucho mejor será la aproximación

<u>Definición</u>:

Si A es una matriz de m x n y b pertenece a R^m, una solución de mínimos cuadrados de A x = b es un x tal que:

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

Para toda **x** en R^n

Se busca un vector **x** que haga que A**x** sea el punto en Col A más cercano a **b**.

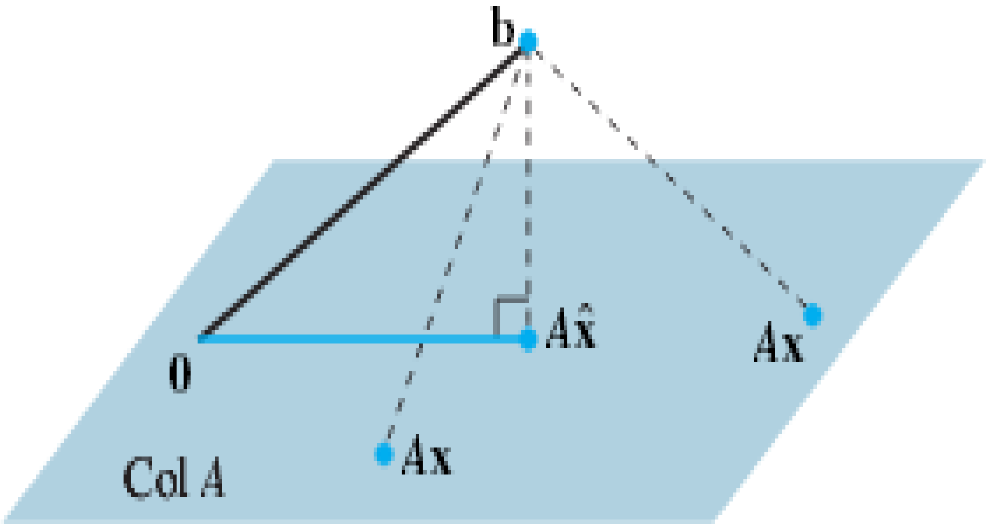
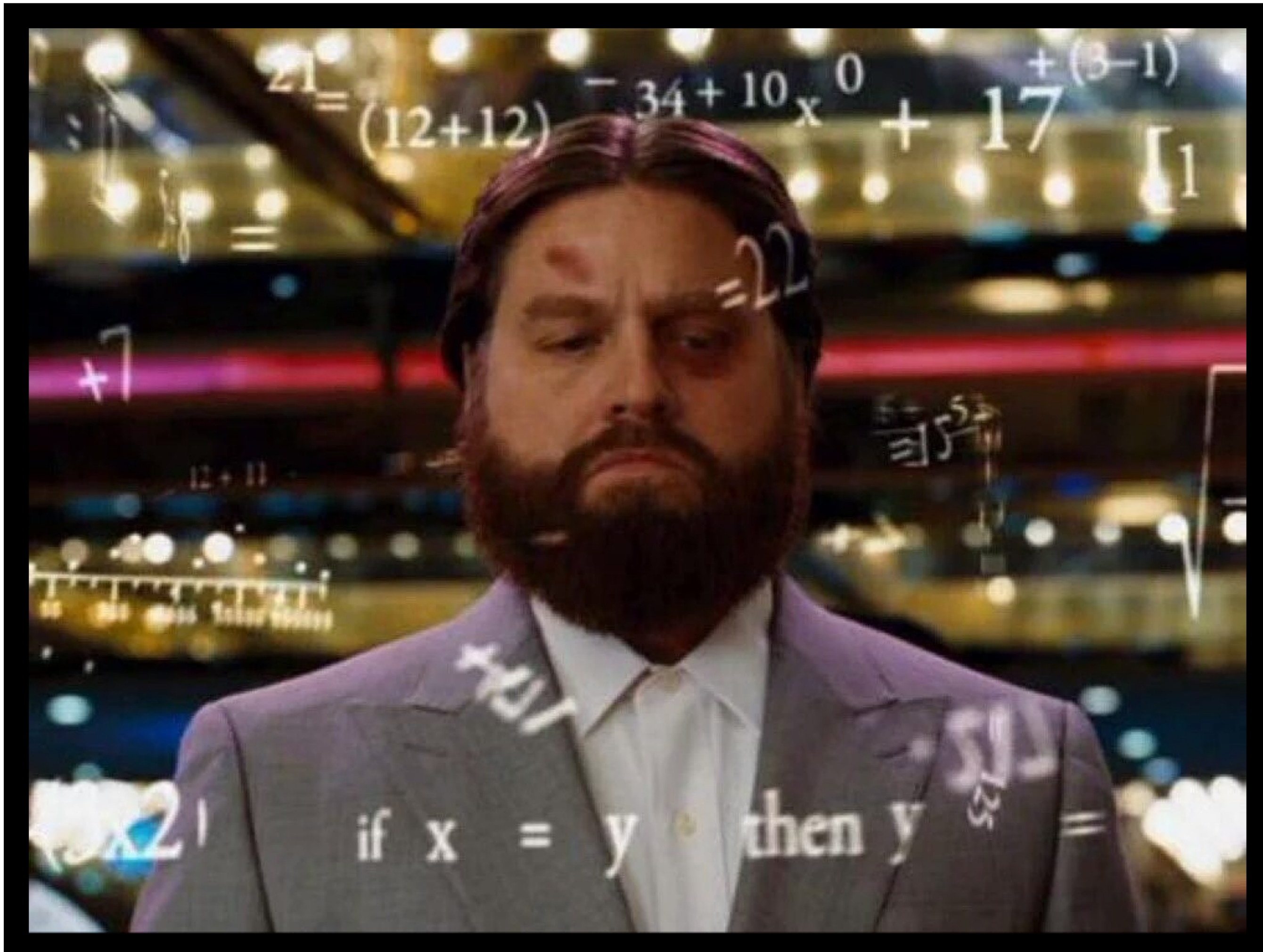


FIGURA 1 El vector **b** está más cerca de $A\hat{\mathbf{x}}$ que de $A\mathbf{x}$ para otro **x**.

Como hacemos?!

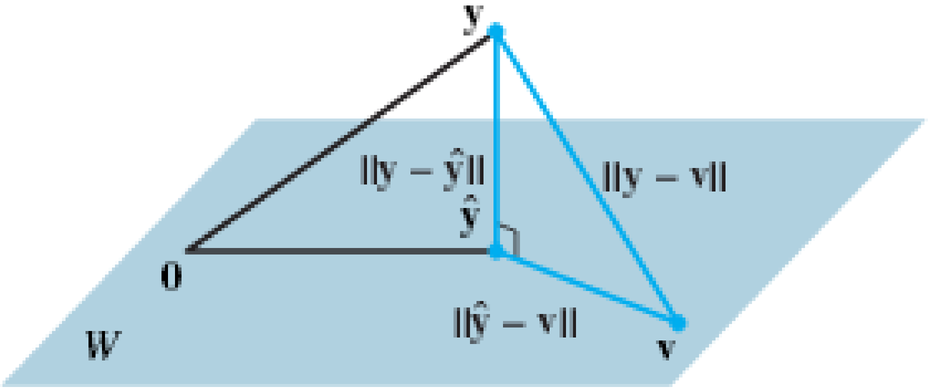


El teorema de la mejor aproximación

Sean W un subespacio de $\mathbb{R}^n$, $\mathbf{y}$ cualquier vector en $\mathbb{R}^n$, y $\hat{\mathbf{y}}$ la proyección ortogonal de $\mathbf{y}$ sobre W. Entonces $\hat{\mathbf{y}}$ es el punto de W más cercano a $\mathbf{y}$, en el sentido que

$$\|\mathbf{y} - \hat{\mathbf{y}}\| < \|\mathbf{y} - \mathbf{v}\| \tag{3}$$

para todo $\mathbf{v}$ en W distinto de $\hat{\mathbf{y}}$.



Teorema de la descomposición ortogonal

Sea W un subespacio de $\mathbb{R}^n$. Entonces toda **y** en $\mathbb{R}^n$ puede escribirse únicamente es la forma y= $\hat{y}$ + z

Donde $\hat{y}$ está en W y z en $W^{\perp}$. De hecho, si{u1, … , u2} es cualquier base ortogonal de W, entonces

$$\hat{\mathbf{y}} = \frac{\mathbf{y}\cdot\mathbf{u}_1}{\mathbf{u}_1\cdot\mathbf{u}_1}\mathbf{u}_1 + \cdots + \frac{\mathbf{y}\cdot\mathbf{u}_p}{\mathbf{u}_p\cdot\mathbf{u}_p}\mathbf{u}_p$$

Y además z= y − $\hat{y}$

El vector $\hat{\mathbf{y}}$ de (1) es la **proyección ortogonal de y sobre** W y a menudo se escribe como $\text{proy}_W\,\mathbf{y}$. Vea la figura 2. Cuando W es un subespacio unidimensional, la fórmula para $\hat{\mathbf{y}}$ corresponde a la fórmula dada en la sección 6.2.

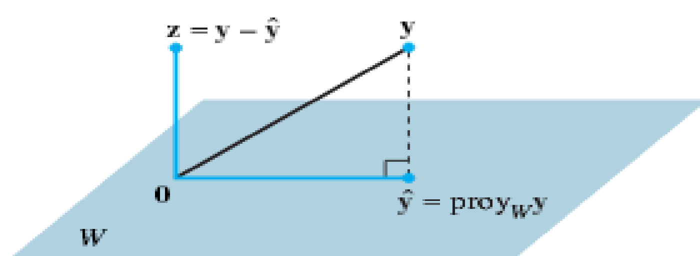


FIGURA 2 La proyección ortogonal de **y** sobre W.

Solución del problema general de mínimos cuadrados

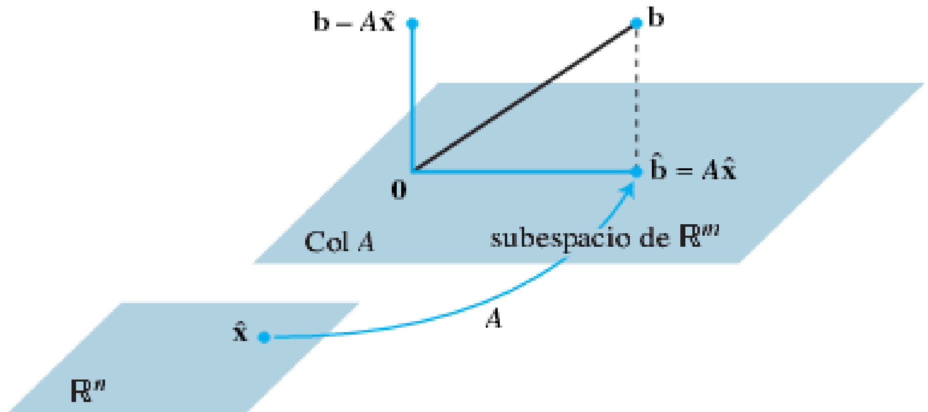


FIGURA 2 La solución por mínimos cuadrados $\hat{\mathbf{x}}$ está en $\mathbb{R}^n$.

La matr A^tA es invertible si y sólo si las columnas de A son linealmente independientes. En este caso, la ecuación A**x**=**b** tiene solamente una solución por mínimos cuadrados $\hat{\mathbf{x}}$, y está dada por :

$$\hat{\mathbf{x}} = (A^TA)^{-1}A^T\mathbf{b}.$$

ntre una solución por mínimos cuadrados del sistema inconsistente A**x**=b

$$A = \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix}, \qquad \mathbf{b} = \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix}$$

Solución Para usar (3), calcule:

$$A^TA = \begin{bmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 17 & 1 \\ 1 & 5 \end{bmatrix}$$

$$A^T\mathbf{b} = \begin{bmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix} = \begin{bmatrix} 19 \\ 11 \end{bmatrix}$$

Entonces la ecuación $A^TA\mathbf{x} = A^T\mathbf{b}$ se vuelve

$$\begin{bmatrix} 17 & 1 \\ 1 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 19 \\ 11 \end{bmatrix}$$

Pueden usarse operaciones por fila para resolver este sistema, pero como A^TA es invertible y 2×2, probablemente sea más fácil calcular

$$(A^TA)^{-1} = \frac{1}{84} \begin{bmatrix} 5 & -1 \\ -1 & 17 \end{bmatrix}$$

y luego resolver $A^TA\mathbf{x} = A^T\mathbf{b}$ como

$$\begin{aligned} \hat{\mathbf{x}} &= (A^TA)^{-1}A^T\mathbf{b} \\ &= \frac{1}{84} \begin{bmatrix} 5 & -1 \\ -1 & 17 \end{bmatrix} \begin{bmatrix} 19 \\ 11 \end{bmatrix} = \frac{1}{84} \begin{bmatrix} 84 \\ 168 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \end{aligned}$$

Criterios para una única solución:

Sea A una matríz de $m \times n$. Los siguientes enunciados son lógicamente equivalentes:

a) La ecuación $A\mathbf{x} = \mathbf{b}$ tiene una solución de mínimos cuadrados única para cada $\mathbf{b}$ en $\mathbb{R}^m$.

b) Las columnas de A son linealmente independientes.

c) La matriz A^TA es invertible.

Cuando estos enunciados son verdaderos, la solución $\hat{\mathbf{x}}$ de mínimos cuadrados está dada por

$$\hat{\mathbf{x}} = (A^TA)^{-1}A^T\mathbf{b} \tag{4}$$

Tenemos que probar que a)✉b), b)✉c) y c)✉a

Veamos un ejemplo:

Dado: $A = \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix}$ $\mathbf{b} = \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix}$ $\hat{\mathbf{x}} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ y , determine el error de mínin rac

en la solución por mínimos cuadrados de A**x=b**

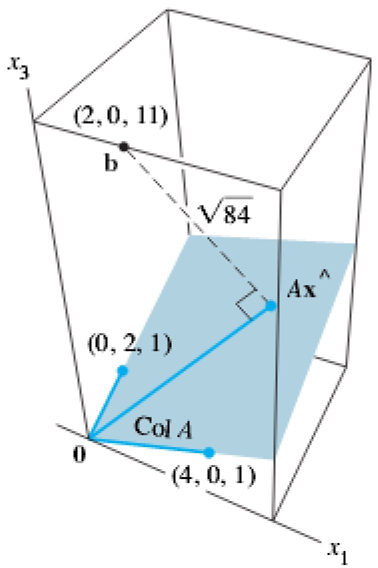


FIGURA 3

$$\mathbf{b} = \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix} \quad \text{y} \quad A\hat{\mathbf{x}} = \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \\ 3 \end{bmatrix}$$

De aquí que

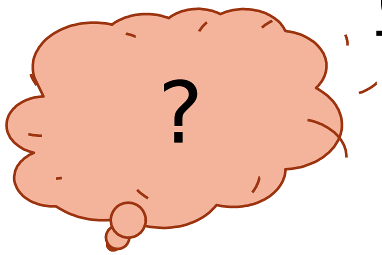
$$\mathbf{b} - A\hat{\mathbf{x}} = \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix} - \begin{bmatrix} 4 \\ 4 \\ 3 \end{bmatrix} = \begin{bmatrix} -2 \\ -4 \\ 8 \end{bmatrix}$$

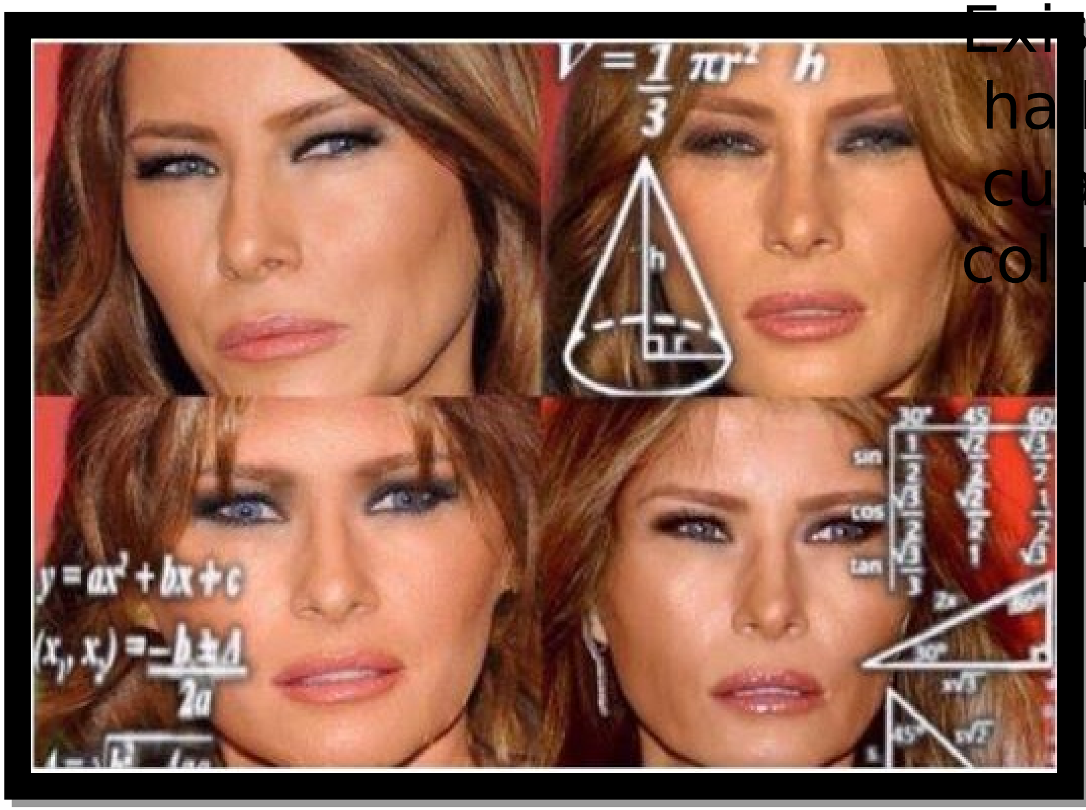
y

$$\|\mathbf{b} - A\hat{\mathbf{x}}\| = \sqrt{(-2)^2 + (-4)^2 + 8^2} = \sqrt{84}$$

El error de mínimos cuadrados es $\sqrt{84}$. Para cualquier $\mathbf{x}$ en $\mathbb{R}^2$, la distancia entre $\mathbf{b}$ y el vector $A\mathbf{x}$ es de al menos $\sqrt{84}$. Véase la figura 3. Observe que la solución por mínimos cuadrados $\hat{\mathbf{x}}$ no aparece en la figura.

Cálculo alternativo de solución de mínimos cuadrados





Existe una forma alternativa
hallar una solución de mínim
cuadrados de A**x**=**b** cuando
columnas de A son ortogona

Veamos un ejemplo:

EJEMPLO 4 Encuentre una solución por mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$ para

$$A = \begin{bmatrix} 1 & -6 \\ 1 & -2 \\ 1 & 1 \\ 1 & 7 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} -1 \\ 2 \\ 1 \\ 6 \end{bmatrix}$$

Solución Como las columnas $\mathbf{a}_1$ y $\mathbf{a}_2$ de A son ortogonales, la proyección ortogonal de $\mathbf{b}$ sobre Col A está dada por

$$\hat{\mathbf{b}} = \frac{\mathbf{b} \cdot \mathbf{a}_1}{\mathbf{a}_1 \cdot \mathbf{a}_1}\mathbf{a}_1 + \frac{\mathbf{b} \cdot \mathbf{a}_2}{\mathbf{a}_2 \cdot \mathbf{a}_2}\mathbf{a}_2 = \frac{8}{4}\mathbf{a}_1 + \frac{45}{90}\mathbf{a}_2 \tag{5}$$

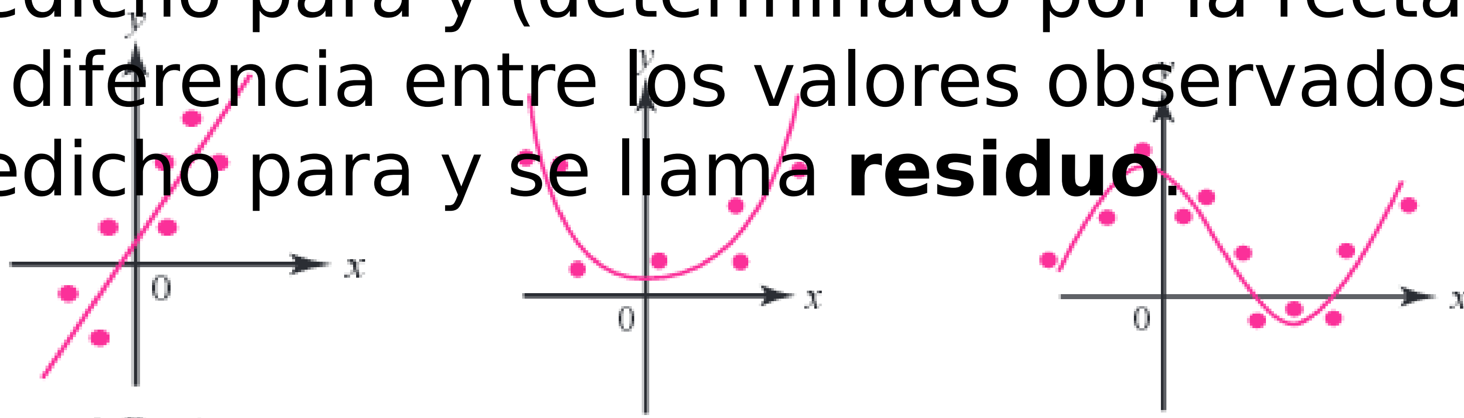
$$= \begin{bmatrix} 2 \\ 2 \\ 2 \\ 2 \end{bmatrix} + \begin{bmatrix} -3 \\ -1 \\ 1/2 \\ 7/2 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ 5/2 \\ 11/2 \end{bmatrix}$$

Ahora que se conoce $\hat{\mathbf{b}}$, puede resolverse $A\hat{\mathbf{x}} = \hat{\mathbf{b}}$. Pero esto es trivial, pues ya se sabe qué pesos colocar en las columnas de A para producir $\hat{\mathbf{b}}$. A partir de (5) resulta claro que

$$\hat{\mathbf{x}} = \begin{bmatrix} 8/4 \\ 45/90 \end{bmatrix} = \begin{bmatrix} 2 \\ 1/2 \end{bmatrix}$$

Aplicaciones a modelos lineales

Supongamos que B0 y B1 están fijos, y consideremos la recta y= B0+B1x de la figura. Para cada dato (x(j);y(j)) existe un punto (x(j);B0+B1x(j)) sobre la recta con la misma coordenada x. Por otro lado, y(j) es el valor observado de y, y B0+B1x(j) es el valor predicho para y (determinado por la recta). La diferencia entre los valores observados y predicho para y se llama **residuo**.



a) Recta b) Cuadrática c) Cúbica

Aproximación por una recta

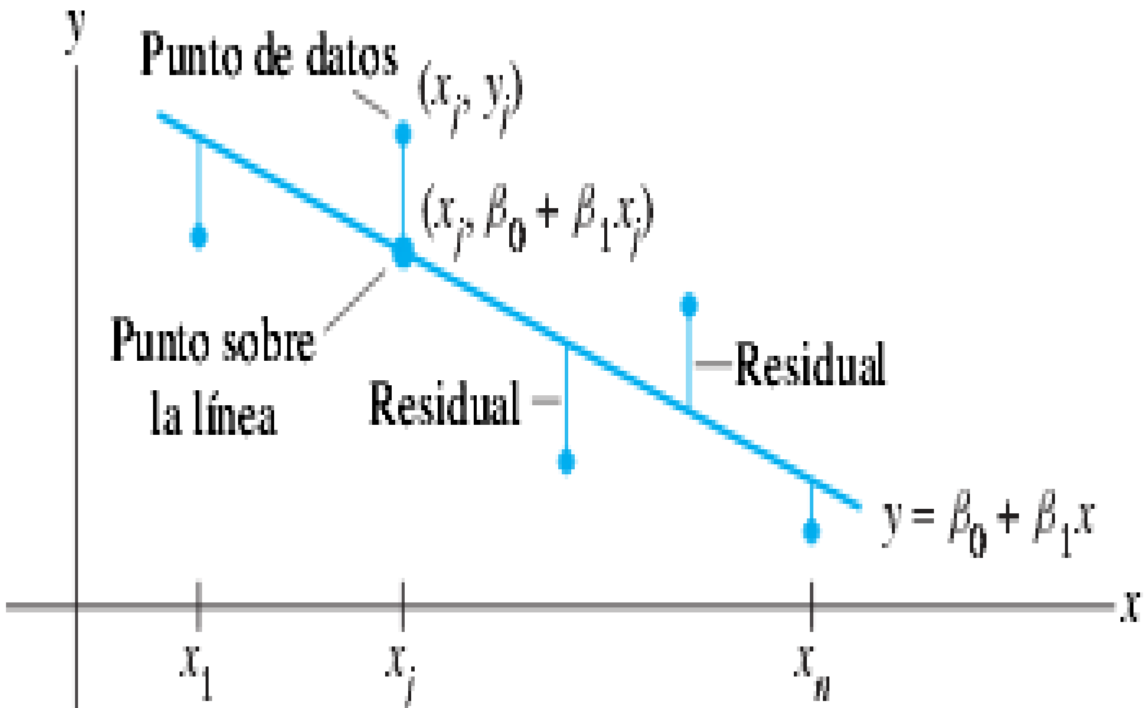


FIGURA 1 Ajuste de una línea a datos experimentales.

Si los puntos de datos estuvieran sobre la línea, los parámetros β_0 y β_1 satisfarían las ecuaciones

Valor de y pronosticado		Valor de y observado
$\beta_0 + \beta_1 x_1$	$=$	y_1
$\beta_0 + \beta_1 x_2$	$=$	y_2
$\vdots$		$\vdots$
$\beta_0 + \beta_1 x_n$	$=$	y_n

Este sistema puede escribirse como

$$X\boldsymbol{\beta} = \mathbf{y}, \quad \text{donde } X = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix}, \quad \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \qquad (1)$$

EJEMPLO 1 Encuentre la ecuación $y = \beta_0 + \beta_1 x$ de la línea de mínimos cuadrados que se ajuste mejor a los puntos de datos (2, 1), (5, 2), (7, 3), (8, 3).

Solución Utilice las coordenadas x de los datos para estructurar la matriz X en (1) y las coordenadas y para estructurar el vector **y**:

$$X = \begin{bmatrix} 1 & 2 \\ 1 & 5 \\ 1 & 7 \\ 1 & 8 \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 3 \end{bmatrix}$$

Para la solución por mínimos cuadrados de $X\boldsymbol{\beta} = \mathbf{y}$, obtenga las ecuaciones normales (con la notación nueva):

Es decir, calcule

$$X^T X = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 5 & 7 & 8 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 5 \\ 1 & 7 \\ 1 & 8 \end{bmatrix} = \begin{bmatrix} 4 & 22 \\ 22 & 142 \end{bmatrix}$$

$$X^T \mathbf{y} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 5 & 7 & 8 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \\ 3 \end{bmatrix} = \begin{bmatrix} 9 \\ 57 \end{bmatrix}$$

Las ecuaciones normales son

$$\begin{bmatrix} 4 & 22 \\ 22 & 142 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} 9 \\ 57 \end{bmatrix}$$

De donde

$$\begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} 4 & 22 \\ 22 & 142 \end{bmatrix}^{-1} \begin{bmatrix} 9 \\ 57 \end{bmatrix} = \frac{1}{84} \begin{bmatrix} 142 & -22 \\ -22 & 4 \end{bmatrix} \begin{bmatrix} 9 \\ 57 \end{bmatrix} = \frac{1}{84} \begin{bmatrix} 24 \\ 30 \end{bmatrix} = \begin{bmatrix} 2/7 \\ 5/14 \end{bmatrix}$$

Entonces la línea de mínimos cuadrados tiene la ecuación

$$y = \frac{2}{7} + \frac{5}{14}x$$

Vea la figura 2.

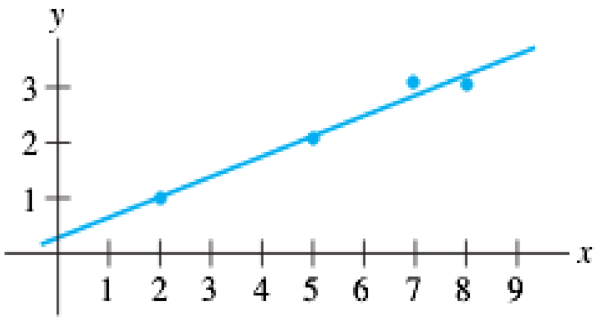


FIGURA 2 La línea de mínimos cuadrados $y = \frac{2}{7} + \frac{5}{14}x$.

Modelo Lineal general

En algunas aplicaciones es necesario ajustar puntos de datos a algo diferente de una línea recta, ya sea órbitas de planetas o cometas, curvas de luz, etc. En este caso la ecuación matricial sigue siendo XB=y pero la forma específica de X cambia de un problema a otro. Por lo general, los especialistas en estadística introducen un vector res como e=y-XB y se escribe

$$\mathbf{y} = X\boldsymbol{\beta} + \boldsymbol{\epsilon}$$

Cualquier ecuación de esta forma es un **modelo lineal**

Una vez que X e y están determinadas, el objetivo es minimizar la longitud de e, lo que equivale a encontrar una solución de mínimos cuadrados de XB=y. En cada caso, la solución de mínimos cuadrados B’ es una solución de las ecuaciones normales:

$$X^T X\boldsymbol{\beta} = X^T\mathbf{y}$$

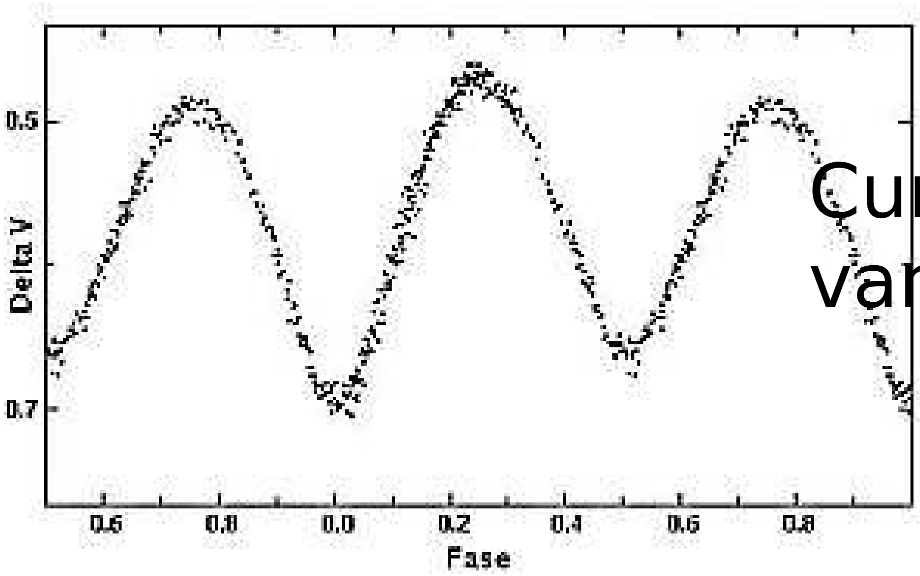
Ajuste de otras curvas con mínimos cuadrados

Cuando los puntos de datos (x1;y1),…,(xn;yn) en una gráfica no se encuentran cerca de una recta, tal vez resulte pertinente postular alguna otra relación funcional entre x e y.

Los datos se van a ajustar mediante curvas que tienen la forma general

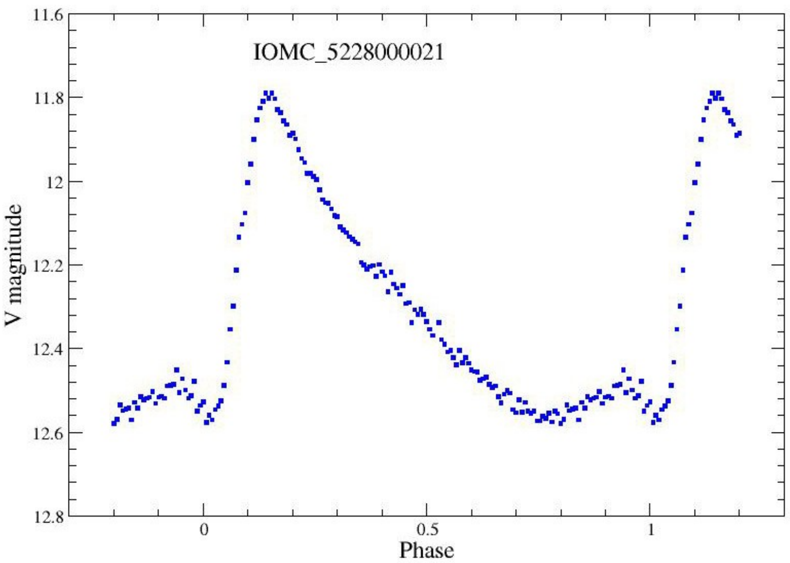
$$y = \beta_0 f_0(x) + \beta_1 f_1(x) + \cdots + \beta_k f_k(x)$$

Donde f0,…,fk son funciones conocidas y B0,…,Bk son parámetros que se deben determinar. Para un valor particular de x, la ecuación da un valor predicho o “ajustado” de y. La diferencia entre el valor observado y el valor predicho es el residuo. Los parámetros B0,…,Bk se deben determinar para minimizar la suma de los cuadrados de los residuos.



Curva de luz de una estre
variable V1363 Ori

Cambio en el brillo de
una estrella Cefeida



Regresión múltiple

Supongamos que en un experimento tenemos dos variables independientes **u** y **v** y una variable dependiente **y**. Una ecuación para predecir **y** a partir de **u** y **v** tiene la forma

$$y = \beta_0 + \beta_1 u + \beta_2 v \quad (4)$$

Una ecuación predictiva más general podría tener la forma

$$y = \beta_0 + \beta_1 u + \beta_2 v + \beta_3 u^2 + \beta_4 uv + \beta_5 v^2$$

Ambas ecuaciones conducen a un modelo lineal porque son lineales en los parámetros desconocidos (aún cuando **u** y **v** están multiplicados). En general, un modelo lineal surgirá siempre que **y** se prediga mediante una e

$$y = \beta_0 f_0(u,v) + \beta_1 f_1(u,v) + \cdots + \beta_k f_k(u,v)$$

EJEMPLO 4 En geografía, se estructuran modelos locales de terreno a partir de datos $(u_1, v_1, y_1), \ldots, (u_n, v_n, y_n)$, donde u_j, v_j y y_j son latitud, longitud y altitud, respectivamente. Describa el modelo lineal basado en (4) que proporciona un ajuste por mínimos cuadrados a datos de este tipo. La solución es el *plano de mínimos cuadrados*. Vea la figura 6.

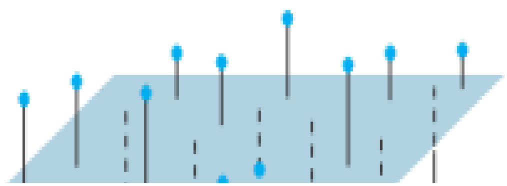


FIGURA 6 Un plano de mínimos cuadrados.

Ejercicio de aplicación

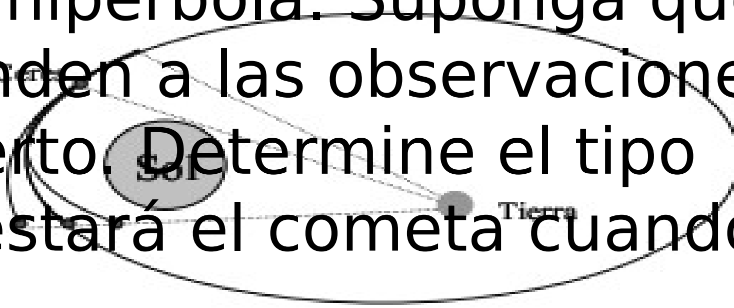
De acuerdo con la segunda ley de Kepler, un cometa debería tener una órbita elíptica, parabólica o hiperbólica (ignorando las atracciones gravitacionales entre los θ planetas). En convenientes coordenadas polares, la posici... cometa satisface la ecuación de la forma:

$$r = \beta - e(r \cdot \cos \vartheta)$$

β

Donde es una constante y e es la excentricidad de la órbita, con 0<e<1 para una elipse, e=1 para una parábola y e>1 para una hipérbola. Suponga que θ los siguientes datos corresponden a las observaciones c... cubierto. Determine el tipo c... nde estará el cometa cuando =4.6 radianes.

ϑ	.88	1.10	1.42	1.77	2.14
r	3.00	2.30	1.65	1.25	1.01




gracias
Ed Hardy

iografía

ebra Lineal y sus Aplicaciones. David C L

ebra Lineal. Stanley Grossman

ernet:

onografias.com

ikipedia