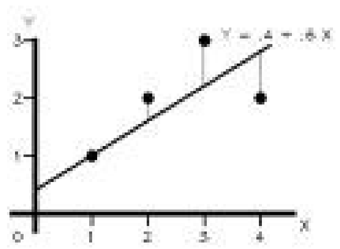
ALGEBRA LINEAL EN CONTEXTO
José Arturo Barreto G.,M.A
The University of Texas at Austin

Correo electrónico: joseartur obarreto@yahoo.com

Páginas Web: www.abaco.com.ve www.miprofe.com.ve www.abrakadabra.com.ve



Capítulo 6

Regresión Lineal

Al terminar este capítulo el estudiante estará en capacidad de resolver diversos problemas de aproximación lineal por mínimos cuadrados, utilizando la ecuación normal.

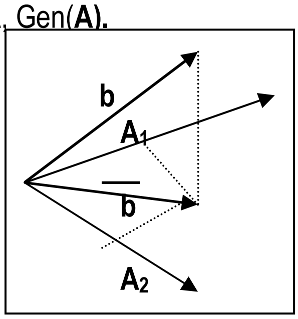
Preliminares

En el capítulo de vectores en R^n, vimos que cuando el sistema de ecuaciones lineales

Ax = **b** no tiene solución, podemos siempre resolver **Ax** = $\bar{\mathbf{b}}$,

El cual siempre tiene solución ya que $\bar{\mathbf{b}}$ es la "proyección" de **b** sobre el subespacio generado por las columnas de **A.**

En ese capítulo solucionamos el problema **Ax = b**, en donde A era una matriz de orden 2, con vectores columna (columnas), linealmente independientes, proyectando a **b** sobre el subespacio generado por las columnas de **A**, Gen(**A**).



Para resolver el problema procedimos así:

i) Calculamos una base $\mathbf{V_1}$, $\mathbf{V_2}$, de Gen (**A**).
ii) Calculamos la proyección $\bar{\mathbf{b}}$ de **b** en Gen(**A**).
iii) Procedimos a resolver **Ax** = $\bar{\mathbf{b}}$.

La solución **x** así obtenida se denominó **una solución por mínimos cuadrados de Ax = b.**

En tal capítulo, a título de ilustración, resolvimos, por mínimos cuadrados, el sistema de ecuaciones lineales simultáneas:

$$\begin{aligned} 2x + y &= 1 \\ 2x - y &= 2 \\ x + y &= 1 \end{aligned}$$

La matriz de los coeficientes del sistema de ecuaciones planteado es .

$$A = \begin{pmatrix} 2 & 1 \\ 2 & -1 \\ 1 & 1 \end{pmatrix}$$

Sus vectores columna son:

$$\mathbf{v_1} = \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix} \quad \mathbf{v_2} = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$$

Como $\mathbf{v_1}$ y $\mathbf{v_2}$, son linealmente independientes, anteriormente construimos a partir de ellos, un conjunto ortogonal de Gen(**A**), formado por:

$$\mathbf{u_1} = \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix} \quad \mathbf{u_2} = \begin{pmatrix} 7/9 \\ -11/9 \\ 8/9 \end{pmatrix}$$

El proceso de ortogonalización aplicado, exigía que los dos vectores no fueran colineales, esto es que fuesen linealmente independientes.

La condición de independencia lineal de **n** vectores $\mathbf{v_1}, \mathbf{v_2}, \dots, \mathbf{v_n}$, es que la ecuación

$$\lambda_1 \mathbf{v_1} + \lambda_2 \mathbf{v_2} + \dots + \lambda_n \mathbf{v_n} = \mathbf{0},$$

tenga como única solución:

$$\lambda_1 = \lambda_2 = \dots = \lambda_n = 0$$

o sea que la única solución sea la trivial $\boldsymbol{\lambda} = \mathbf{0}$.

Esta condición la satisfacen plenamente los vectores $\mathbf{v_1}$ y $\mathbf{v_2}$.

Por esta razón se pudo aplicar el proceso de ortogonalización de Gram-Schmidt. Si no, no se podría aplicar. Aún más $\mathbf{v_1}$ y $\mathbf{v_2}$ son una base de Gen(**A**), ya que además de ser linealmente independientes, por la propia definición de Gen(**A**), lo generan.

La situación se complica cuando las columnas de **A** no son linealmente independientes como veremos en el ejemplo siguiente.

Problema

Hallar el conjunto solución del sistema de ecuaciones lineales:

(*)

$$x + y + z = 1$$
$$2x + v$$

Utilizando el método de Gauss, procedemos así:

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 2 & 1 & 0 & 2 \\ -1 & 2 & 5 & -3 \end{array}\right) \equiv \begin{array}{c} F_2 - 2F_1 \\ F_3 + F_1 \end{array} \equiv \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & -1 & -2 & 1 \\ 0 & 3 & 6 & -2 \end{array}\right) \equiv \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & -1 & -2 & 1 \\ 0 & 0 & 0 & 1 \end{array}\right)$$

El sistema de ecuaciones equivalente :

$$\begin{aligned} x + y + z &= 1 \\ -y - 2z &= 1 \\ 0z &= 1 \end{aligned}$$

No tiene solución (es inconsistente).

Procederemos a hallar su solución por mínimos cuadrados:

Tomemos los vectores

$$\mathbf{A}_1 = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} \quad \mathbf{A}_2 = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} \quad \mathbf{A}_3 = \begin{bmatrix} 1 \\ 0 \\ 5 \end{bmatrix}$$

Veamos si estos vectores son linealmente independientes.

A partir de la ecuación: $\lambda_1 \mathbf{A}_1 + \lambda_2 \mathbf{A}_2 + \lambda_3 \mathbf{A}_3 = \mathbf{0},$

O equivalentemente:

$$\lambda_1 \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} + \lambda_2 \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} + \lambda_3 \begin{bmatrix} 1 \\ 0 \\ 5 \end{bmatrix} = \mathbf{0}$$

La cual equivale a:

$$\begin{aligned} \lambda_1 + \lambda_2 + \lambda_3 &= 0 \\ 2\lambda_1 + \lambda_2 &= 0 \\ -\lambda_1 + 2\lambda_2 + 5\lambda_3 &= 0 \end{aligned}$$

Equivalente a resolver el sistema homogéneo $\mathbf{A\lambda} = \mathbf{0}$, asociado al sistema no homogéneo planteado en (*).

El sistema homogéneo siempre tiene al menos una solución, la trivial: $\boldsymbol{\lambda} = \mathbf{0}$. ($\lambda_1 = \lambda_2 = \lambda_3 = 0$).
Si la única solución es la trivial, los vectores son linealmente independientes. Si hay al menos una solución, $\boldsymbol{\lambda} \neq \mathbf{0}$, los vectores son linealmente dependientes.

Resolvamos por Gauss, el sistema homogéneo:

$$\begin{aligned} \lambda_1 + \lambda_2 + \lambda_3 &= 0 \\ 2\lambda_1 + \lambda_2 &= 0 \\ -\lambda_1 + 2\lambda_2 + 5\lambda_3 &= 0 \end{aligned}$$

Este es exactamente el sistema homogéneo asociado planteado al iniciar el problema. Siguiendo los mismos pasos con los cuales se intentó resolver tal problema por Gauss, pero cambiando el parámetro por el vector **0,** llegamos a:

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & -1 & -2 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

El nuevo sistema a resolver es por lo tanto:

$$\begin{array}{rcl} \lambda_1 + \lambda_2 + \lambda_3 & = & 0 \\ -\lambda_2 - 2\lambda_3 & = & 0 \\ 0\,\lambda_3 & = & 0 \end{array}$$

Vemos que λ_3 puede tomar cualquier valor, no necesariamente 0. Si $\lambda_3 = 1$, concluimos que $\lambda_2 = -2$ y
$\lambda_1 = 1$ otros, infinitos, valores de $\boldsymbol{\lambda}$ son válidos.

En consecuencia, los vectores columna de **A** son linealmente dependientes y por ello no podemos calcular a partir de ellos, una base ortogonal de Gen(**A**).

Sin embargo los valores calculados $\lambda_3 = 1$, $\lambda_2 = -2$, $\lambda_3 = 1$, nos permiten afirmar que:

$$\mathbf{A}_1 - 2\,\mathbf{A}_2 + \mathbf{A}_3 = \mathbf{0},$$

Luego: $\mathbf{A}_1 = 2\,\mathbf{A}_2 - \mathbf{A}_3$, **(o** $\mathbf{A}_2 = \frac{1}{2}\mathbf{A}_1 + \frac{1}{2}\mathbf{A}_3$, etc.**),**

Como $\mathbf{A}_1$ se puede expresar como combinación lineal de $\mathbf{A}_2$ y $\mathbf{A}_3$, entonces $\mathbf{A}_1$ se puede extraer del conjunto $\{ \mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3 \}$ de generadores de Gen(**A**) y el conjunto restante $\{ \mathbf{A}_2, \mathbf{A}_3 \}$ todavía genera a Gen(**A**). Ya que
si $\mathbf{x} \,\varepsilon\,$ Gen(**A**), entonces $\mathbf{x} = \lambda_1 \mathbf{A}_1 + \lambda_2 \mathbf{A}_2 + \lambda_3 \mathbf{A}_3$, para alguna tripla $\lambda_1, \lambda_2, \lambda_3$,
por definición de Gen(**A**).

Sustituyendo a $\mathbf{A}_1$ por $2\,\mathbf{A}_2 - \mathbf{A}_3$, tenemos que:

$$\mathbf{x} = \lambda_1 (2\,\mathbf{A}_2 - \mathbf{A}_3) + \lambda_2 \mathbf{A}_2 + \lambda_3 \mathbf{A}_3 = (2\lambda_1 + \lambda_2)\,\mathbf{A}_2 + (\lambda_3 - \lambda_1)\,\mathbf{A}_3$$

es una combinación lineal de los vectores columna, restantes $\mathbf{A}_2$, $\mathbf{A}_3$

Luego $\mathbf{A}_2$ y $\mathbf{A}_3$, generan Gen (**A**).

Será el conjunto $\{ \mathbf{A}_2, \mathbf{A}_3 \}$ una base de Gen(**A**)., para así construir por el proceso de ortogonalización de Gram-Schmidt una base ortogonal de Gen(**A**) ?.

La respuesta es sí, ya que estos dos últimos vectores son linealmente independientes.

Planteemos la ecuación que nos permite verificar la independencia lineal:

$$\text{Sea } \lambda_2 \mathbf{A_2} + \lambda_3 \mathbf{A_3} = \mathbf{0},$$

Entonces:

$$\lambda_2 \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} + \lambda_3 \begin{pmatrix} 1 \\ 0 \\ 5 \end{pmatrix} = \mathbf{0}$$

Lo cual es equivalente a:

$$\lambda_2 + \lambda_3 =$$
$$0$$
$$\lambda_2 =$$

Por consiguiente, $\lambda_2 = \lambda_3 =$

En consecuencia $\mathbf{A_2}$ y $\mathbf{A_3}$, son linealmente independientes. Lo son también $\mathbf{A_1}$ y $\mathbf{A_3}$, por lo tanto la escogencia de una base de Gen(**A**), tiene muchas alternativas.

A partir de

$$\mathbf{v}_1 = \mathbf{A_2} = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} \mathbf{v}_2 = \mathbf{A}_3 = \begin{pmatrix} 1 \\ 0 \\ 5 \end{pmatrix}$$

Construiremos la base ortogonal de Gen(**A**), que nos permitirá proyectar el vector **b** en $\bar{\mathbf{b}}$ ε Gen(**A**).

Sea $\mathbf{u}_1 = \mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}$

Luego,

$$\mathbf{u}_2 = \mathbf{v}_2 - (< \mathbf{v}_2, \mathbf{u}_1 > / < \mathbf{u}_1, \mathbf{u}_1 >) \, \mathbf{u}_1 =$$

$$\begin{pmatrix} 1 \\ 0 \\ 5 \end{pmatrix} - (11/6) \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} = \begin{pmatrix} -5/6 \\ -11/6 \\ 8/6 \end{pmatrix}$$

Ya que $< \mathbf{v}_2, \mathbf{u}_1 > = \mathbf{v}_2^T \mathbf{u}_1 = (1 \quad 0 \quad 5) \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} = 11$

$< \mathbf{u}_1, \mathbf{u}_1 > = \mathbf{u}_1^T \mathbf{u}_1 = (1 \quad 1 \quad 2) \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} = 6$

Luego, una base ortogonal de Gen(**A**) (compruebe que es ortogonal), es:

$$\mathbf{u}_1 = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} \quad \mathbf{u}_2 = \begin{pmatrix} -5/6 \\ -11/6 \\ 8/6 \end{pmatrix}$$

La proyección $\overline{\mathbf{b}}$ de **b** en Gen(**A**), está dada por:

$$\overline{\mathbf{b}} = (\langle \mathbf{b}, \mathbf{u}_1 \rangle / \langle \mathbf{u}_1, \mathbf{u}_1 \rangle)\, \mathbf{u}_1 + (\langle \mathbf{b}, \mathbf{u}_2 \rangle / \langle \mathbf{u}_2, \mathbf{u}_2 \rangle)\, \mathbf{u}_2 = \quad (-2/6) \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} + \left(-\,\overline{31/3)}\right) \begin{pmatrix} -5/6 \\ -11/6 \\ 8/6 \end{pmatrix}$$

$$= \begin{pmatrix} 8/7 \\ \\ 102/3 \end{pmatrix}$$

Ya que

$$\langle \mathbf{b}, \mathbf{u}_1 \rangle = \mathbf{b}^T \mathbf{u}_1 = (1 \quad 3 \quad -3) \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} = -2$$

$$\langle \mathbf{u}_1, \mathbf{u}_1 \rangle = 1^2 + 1^2 + 2^2 = 6$$

$$\langle \mathbf{b}, \mathbf{u}_2 \rangle = \mathbf{b}^T \mathbf{u}_2 = (1 \quad 3 \quad -3) \begin{pmatrix} -5/6 \\ -1/6 \\ 8/6 \end{pmatrix} = (-5/6) - (33/6) - (24/6) = -(62/6) = -(31/3)$$

$$\langle \mathbf{u}_2, \mathbf{u}_2 \rangle = \mathbf{u}_2^T \mathbf{u}_2 = (-5/6)^2 + (1/6)^2 + (8/6)^2 = 35/6$$

Procedemos a resolver $\quad \mathbf{Ax} = \overline{\mathbf{b}}$

En lugar de $\quad \mathbf{Ax} = \mathbf{b}$

Hallemos, por lo tanto las soluciones de:

$$\begin{array}{l} x + y + z = 8/7 \approx 1{,}14 \\ 2x + y = 102/35 \approx 2{,}91, \\ -x + 2y + 5z = -106/35 \approx -3{,}03 \end{array} \quad \text{en lugar de} \quad \begin{array}{l} x + y + z = 1 \\ 2x + y = 3 \\ -x + 2y + 5z = -3 \end{array}$$

Al observar la variación, diferencia o aproximación de , la cual podría medirse con la "norma"

$$|\mathbf{b} - \overline{\mathbf{b}}|$$

O distancia de **b** a $\overline{\mathbf{b}}$ (o de $\overline{\mathbf{b}}$ a **b**),

y si aceptamos una aproximación a la décima

$$1{,}1 \approx 1, \qquad 2{,}9 \approx 3, \qquad -3{,}03 \approx -3,$$

Podremos pensar que la solución por mínimos cuadrados va a modelar satisfactoriamente a los datos. Esto parece así, pero no podemos estar seguros hasta que no hagamos los cálculos y consideremos por diferentes criterios, así sean artesanales, que los datos "aproximan satisfactoriamente" a los resultados.

Resolveremos por Gauss, el sistema de ecuaciones lineales que aproxima al sistema de ecuaciones planteado, por mínimos cuadrados, así:

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 8/7 \\ 1 & 1 & 0 & 102/35 \\ -1 & 2 & 5 & -106/35 \end{array}\right) \underset{\approx}{F_2 - 2F_1} \left(\begin{array}{ccc|c} 1 & 1 & 1 & 8/7 \\ 0 & -1 & -2 & 22/35 \\ 0 & 3 & 6 & -66/35 \end{array}\right) \underset{F_3 + 3F_2}{\approx} \left(\begin{array}{ccc|c} 1 & 1 & 1 & 8/7 \\ 0 & -1 & -2 & 22/35 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Como puede observarse, las tres primeras columnas de las matrices aumentadas, o sea la matriz de los coeficientes es precisamente la matriz de los coeficientes del sistema incompatible planteado en este problema y sólo se ha cambiado el vector **b** por $\overline{\mathbf{b}}$

Concluimos que el conjunto de las soluciones por mínimos cuadrados, debe satisfacer

$$\begin{aligned} x + y + z &= 8/7 \\ y + 2z &= -22/35 \end{aligned}$$

Todas las soluciones son soluciones por mínimos cuadrados. Cuál escoger?. Habrán criterios adicionales en el problema real?.

Una solución podría corresponder a z= 0, de donde y = -22/35, x= 62/35.

Otra podría ser (escogiendo valores por simple comodidad para los cálculos) y = 0, de donde z= -11/35,
x= 51/35. Las posibilidades de solución son infinitas.

Todas cumplen por supuesto con la condición $\mathbf{Ax} = \overline{\mathbf{b}}$. Por lo tanto, todas minimizan la distancia a Gen(**A**).

Por lo general, al tener la matriz de los coeficientes de $\mathbf{Ax} = \mathbf{b}$, mas ecuaciones que incógnitas **,** sus columnas seran linealmente independientes. En este caso la solución es **única** puesto que la expresión lineal $\mathbf{Ax} = \overline{\mathbf{b}}$ será del tipo,

$$x_1 \mathbf{A_1} + x_2 \mathbf{A_2} + \ldots + x_n \mathbf{A_n} = \overline{\mathbf{b}}$$

la cual es única ya que $\{ \mathbf{A_1}, \mathbf{A_2}, \ldots, \mathbf{A_n} \}$ sería una base de Gen(**A**).

<u>Solución por mínimos cuadrados utilizando la ecuación normal</u>

Como $\mathbf{b} - \overline{\mathbf{b}}$ es ortogonal a Gen(**A**) = Gen $\{ \mathbf{A_1}, \mathbf{A_2}, \ldots, \mathbf{A_n} \}$, donde las $\mathbf{A_i}$ son las columnas de **A,** entonces:

$$< \mathbf{A}_1, \mathbf{b} - \overline{\mathbf{b}} > = < \mathbf{A}_2, \mathbf{b} - \overline{\mathbf{b}} > = \ldots = < \mathbf{A}_n, \mathbf{b} - \overline{\mathbf{b}} > = \mathbf{0}$$

En consecuencia, los productos matriciales siguientes son válidos:

$$\mathbf{A}_1^{T} (\mathbf{b} - \overline{\mathbf{b}}) = \mathbf{A}_1^{T} (\mathbf{b} - \overline{\mathbf{b}}) = \ldots = \mathbf{A}_n^{T} (\mathbf{b} - \overline{\mathbf{b}}) = \mathbf{0}$$

Pero estos productos son precisamente las filas de $\mathbf{A}^T (\mathbf{b} - \mathbf{b})$, de donde concluimos que:

$$\mathbf{A}^T (\mathbf{b} - \bar{\mathbf{b}}) = \mathbf{0}$$

Y en consecuencia

$$\mathbf{A}^T \mathbf{b} - \mathbf{A}^T \bar{\mathbf{b}} = \mathbf{0},$$

Lo cual es equivalente a:

$$(*)\ \mathbf{A}^T \mathbf{b} = \mathbf{A}^T \bar{\mathbf{b}}$$

Sabemos que

$\mathbf{Ax} = \mathbf{b}$ no tiene solución

Mas sin embargo, si pre-multiplicamos al sistema inconsistente $\mathbf{Ax} = \mathbf{b}$ por $\mathbf{A}^T$, obtenemos el sistema consistente (con solución. Por qué?):

$$\mathbf{A}^T \mathbf{Ax} = \mathbf{A}^T \mathbf{b},$$

El cual es equivalente por **(*)** a:

$$\mathbf{A}^T \mathbf{Ax} = \mathbf{A}^T \bar{\mathbf{b}},$$

Que por la misma razón anterior (Cuál?), siempre tiene solución. Además toda solución de

$$(**) \qquad \mathbf{Ax} = \bar{\mathbf{b}},$$

Es una solución de $\qquad (***) \quad \mathbf{A}^T \mathbf{Ax} = \mathbf{A}^T \mathbf{b},$

Aceptando que estos dos sistemas (**(**) y (***)** son equivalentes (Lo son?), en lugar de resolver la ecuación

$$\mathbf{Ax} = \mathbf{b},$$

Resolvemos la **ecuación normal**

$$\mathbf{A}^T \mathbf{Ax} = \mathbf{A}^T \mathbf{b},$$

Simplificando así nuestro método de solución por mínimos cuadrados.

Si además $\mathbf{A}^T \mathbf{A}$ es una matriz no-singular (inversible), la ecuación normal tendrá solución única. Este caso corresponde a matrices A **con columnas linealmente independientes**. Para una mayor comprensión del tema revise los apéndices correspondientes del proyecto "Algebra Lineal para Todos" en http://www.abaco.com.ve

Una breve mirada a las aplicaciones

Utilizando la ecuación normal, resolveremos por mínimos cuadrados un sistema de ecuaciones que resolvimos en el capítulo de vectores en R^n utilizando el proceso de ortogonalización de Gram-Schmidt.

Los resultados se podrán comparar ya que la matriz $\mathbf{A}^T\ \mathbf{A}$ es, en este ejemplo, de rango completo y por lo tanto, obtendremos una solución única que debe coincidir con la hallada en tal capítulo. No tendremos que pasar por lo tanto por el complicado proceso de ortogalización.

Resolveremos por mínimos cuadrados el sistema inconsistente:

$$\begin{aligned} 2x + y &= 1 \\ 2x - y &= 2 \\ x + y &= 1 \end{aligned}$$

Solución:

$$\mathbf{A} = \begin{pmatrix} 2 & 1 \\ 2 & -1 \\ 1 & 1 \end{pmatrix}$$

$$\mathbf{A}^T\ \mathbf{A} = \begin{pmatrix} 2 & 2 & 1 \\ 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 2 & -1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 9 & 1 \\ 1 & 3 \end{pmatrix} \qquad \mathbf{A}^T\ \mathbf{b} = \begin{pmatrix} 2 & 2 & 1 \\ 1 & -1 & 1 \end{pmatrix} \boxed{\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}} = \begin{pmatrix} 7 \\ 0 \end{pmatrix}$$

La ecuación normal $\mathbf{A}^T\ \mathbf{Ax} = \mathbf{A}^T\mathbf{b}$,

Es por lo tanto: $\mathbf{A}^T\ \mathbf{Ax} = \mathbf{A}^T\mathbf{b}$

,

$$\begin{pmatrix} 9 & 1 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 7 \\ 0 \end{pmatrix}$$

cuya solución única es x = 21/26, y = -7/26

Esta solución es exactamente la obtenida por el proceso de ortogonalización de Gram-Schmidt.

Al reemplazar los valores así hallados en el sistema original

$$\begin{aligned} 2x + y &= 1 \\ 2x - y &= 2 \\ x + y &= 1, \end{aligned}$$

no obtendremos, por supuesto, el vector **b** = $\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$

sino a su proyección $\bar{\mathbf{b}}$**,** en el subespacio generado por las columnas de **A.**

Sustituyendo, encontramos que $\bar{\mathbf{b}} = \begin{pmatrix} 35/26 \\ 49/26 \\ 14/26 \end{pmatrix}$

Coincidiendo con los resultados hallados anteriormente.

Ejemplo:

Los puntos P(1, ½), Q(2,1), R(3, 3/2), están situados sobre la recta cuya ecuación es

$$y = \tfrac{1}{2} x$$

Por dicha razón, <u>no es posible hallar una recta que pase por</u> P(1, ½), Q(2,1) y R(3,2), ya que la recta que pasa por P y Q es la citada anteriormente y el único punto de dicha recta de abscisa 3 es R(3,3/2) y nó R(3,2).

Por supuesto, que al colocar valores a las variables x e y, en una ecuación de la forma y = mx + b, a partir de la tabla:

x	y
1	½
2	1
3	2

Llegaremos al sistema inconsistente

$$\begin{aligned} m + b &= \tfrac{1}{2} \\ 2m + b &= 1 \\ 3m + b &= 2 \end{aligned}$$

La ecuación normal para tal sistema es: $A^T Ax = A^T b$,

$$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \end{pmatrix} \mathbf{x} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} \tfrac{1}{2} \\ 1 \\ 2 \end{pmatrix}$$

o de forma equivalente:

$$\begin{pmatrix} 14 & 6 \\ 6 & 3 \end{pmatrix} \begin{bmatrix} m \\ b \end{bmatrix} = \begin{pmatrix} 17/2 \\ 7/2 \end{pmatrix}$$

tal sistema tiene solución única $m = 3/4$, $b = -1/3$

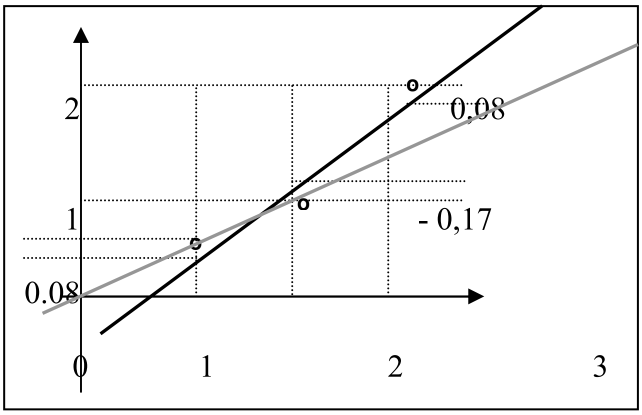
La ecuación de la recta sería: **(*)** $y = 3/4\ x - 1/3$

Sustituyamos valores para ver que tan lejos pasa de los puntos o que tan bien los ajusta

Sustituyendo los valores de **x,** primera columna de la tabla, a partir de (*), obtenemos la tercera columnaa valores aproximados de **b** (o **y**) por simple sustitución.

x	**y = b**	$\bar{\mathbf{b}}$	$\mathbf{b - \bar{b} = r}$
1	½ ≈ 0,5	5/12 ≈ 0,42	$0,08 = r_1$
2	1 ≈ 1	7/16 ≈ 1,17	$-0,17 = r_2$
3	2 ≈ 2	23/2 ≈ 1,92	$0,08 = r_3$

Dibujemos la recta obtenida y observemos qué indican las desviaciones o componentes del vector $\bar{\mathbf{b}} - \mathbf{b}$.



Hemos representado las desviaciones $r_1 = 0,08$, $r_2 = -0,17$, $r_3 = 0,08$

Una medida global de la desviación es la norma $|\mathbf{b} - \bar{\mathbf{b}}| = \sqrt{(r_1^2 + r_2^2 + r_3^2)} \approx \sqrt{0,0417} \approx 0.204$.

El método de los mínimos cuadrados minimiza la suma de los cuadrados de dichas desviaciones.

La recta calculada es la que mejor aproxima los datos siguiendo el método de los mínimos cuadrados. Haremos una comparación gráfica con la recta en tono gris, que pasa por los primeros dos puntos P(1, ½) y Q(2,1).

Las dos primeras desviaciones para la recta gris son $r_1 = r_2 = 0$, ya que la recta gris pasa por los dos primeros puntos. La única desviación diferente de 0 es , en este caso, $r_3 = 2 - 1.5 = 0,5$. (recuerde que la ecuación de la recta gris es: $y = ½\,x$. En este caso $\sqrt{(r_1^2 + r_2^2 + r_3^2)} = \sqrt{0,25} = 0,5$.

En cualquier caso, pese a que la recta $y = ½ X$, PASA EXACTAMENTE POR DOS DE LOS TRES PUNTOS, su desviación en el tercer punto es tal que por supuesto no minimiza la suma total de los cuadrados de las desviaciones. Esto se puede verificar estudiando otras rectas como la que pasa exactamente por P(1, 1/2) y R(3,2), o la que pasa por P(2,1) y R(3,2).

En consecuencia reafirmamos que la solución que optimiza la aproximación por mínimos cuadrados se halla resolviendo la ecuación normal.

El método de los mínimos cuadrados se puede aplicar tambien en el caso en que se desee ajustar una serie de datos por una función dada que es **combinación lineal** de otras funciones, siempre y cuando el problema pueda reducirse a una solución por mínimos cuadrados de un sistema de ecuaciones lineales simultáneas.

Ejemplo:

Halle el polinomio de segundo grado $p(x) = c + dx + ex^2$, que mejor ajusta, en el sentido de lkos mínimos cuadrados, los datos:

x	y
-1	2
0	1
1	6,01
2	16,99

O sea, que pasa por
P(-1,2)
Q(0, 1)
R(1, 6.01)
S(2,16.99)

Sustituyendo x e y en $p(x) = c + dx + ex^2$, obtenemos:

$$\begin{aligned} c - d + e &= 2 \\ c \quad\quad &= 1 \\ c + d + e &= 6{,}01 \\ c + 2d + 4e &= 16{,}99 \end{aligned}$$

La solución por mínimos cuadrado[...]aciones anterior, es una solución de la ecuación normal

$$\mathbf{A}^T \mathbf{A}\mathbf{x} = \mathbf{A}^T\mathbf{b},$$

Donde

$$\begin{pmatrix} 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{pmatrix}$$

La solución es :

$$\begin{pmatrix} 1{,}006 \\ 2{,}003 \\ 2{,}995 \end{pmatrix}$$

En consecuencia, el polinomio de segundo grado buscado es

$$p(x) = 1{,}006 + 2{,}003x + 2{,}995x^2 ,$$

La proyección $\bar{\mathbf{b}}$, de **b**, sobre el subespacio de las columnas de **A** es: $\bar{\mathbf{b}} = \begin{pmatrix} 1{,}998 \\ \\ 1{,}006 \end{pmatrix}$ Como **b**= $\begin{pmatrix} 2 \\ 1 \\ 6{,}01 \\ 16{,}9 \end{pmatrix}$

Las desviaciones r_i son:	$r_1 = 2 - 1{,}998 = 0{,}002$		
	$r_2 = 1 - 1{,}006 = -0{,}006$		
	$r_3 = 6{,}01 - 6{,}004 = 0{,}006$		
	$r_4 = 16{,}99 - 16{,}992 = -0{,}002$		

La solución en este caso parece bastante aceptable.

Una medida de cuanto ajusta la solución por mínimos cuadrados a los m datos, podría ser

$$S = (r_1^2 + r_2^2 + \dots + r_m^2) / m$$

En este caso sería $s = 0{,}00008 / 4 = 2 \times 10^{-5}$.

Otra medida podría ser $|\mathbf{r}| / m$, donde **r** es el vector cuyas componentes son las desviaciones r_i.

Los mismos datos ajustados por una recta deben proporcionar resultados menos halagadores.

Por supuesto que la medida de la "bondad" de los ajustes depende también del tamaño de los datos ya que una desviación "pequeña" para datos con valores"pequeños" puede no ser suficientemente halagadora, como lo sería una desviación razonable para datos con valores "grandes". Existen por lo tanto diferentes propuestas de fórmulas que "miden" la bondad del ajuste.

Planteamos a continuación el siguiente ejemplo "curioso":

Asuma que se quiere medir la temperatura en un lugar tomando "n" pruebas, así:

Prueba i	Temperat
Prueba 1	T_1
Prueba 2	T_2
Prueba 3	T_3
.	.
.	.
Prueba m	T_m

No es de esperar que todas las temperaturas sean iguales auncuando hayan sido tomadas a la mosma hora por diferentes razones: inexactituid del instrumento de medida, errores "menores" del operario, factores climáticos u otras razones.

Nuestra temperatura resultante t , debe satisfacer el sistema de ecuaciones simultáneas:

$$\begin{gathered} t = T_1 \\ t = T_2 \\ t = T_3 \\ \vdots \\ t = T_m \end{gathered}$$

Si las temperaturas T_i no son todas iguales, este sistema será inconsistente.

Como su forma matricial es

$$\begin{pmatrix}1\\1\\1\\.\\.\\1\end{pmatrix} t) = \begin{pmatrix}T_1\\T_2\\T_3\\.\\.\\T_m\end{pmatrix}$$ 0 sea **A t = T,**

En donde

$$\mathbf{A} = \begin{pmatrix}1\\1\\1\\.\\.\\1\end{pmatrix}, \mathbf{t} = (t) \text{ y } \mathbf{T} = \begin{pmatrix}T_1\\T_2\\T_3\\.\\.\\T_m\end{pmatrix}$$

Tenemos que :

$$\mathbf{A}^T \mathbf{A}\, \mathbf{t} = \underbrace{(1 \quad 1 \quad 1 \ldots 1)}_{m\ 1\text{'s}} \begin{pmatrix}1\\1\\1\\.\\.\\1\end{pmatrix} t) = (m\,t) \text{ y } \mathbf{A}^T\mathbf{T} = \overbrace{(1 \quad 1 \quad 1 \ldots 1)}^{m\ 1\text{'s.}} \begin{pmatrix}T_1\\T_2\\T_3\\.\\.\\T_m\end{pmatrix} + T_2 + T_3 + \ldots + T_m$$

Por lo tanto

$$\boxed{m\,t = T_1 + T_2 + T_3 + \ldots + T_m}$$

Concluimos que la solución por mínimos cuadrados es:

$$T = (T_1 + T_2 + T_3 + \ldots + T_m) / m$$

Conclusión: La respuesta que optimiza estas mediciones en el sentido de los mínimos cuadrados es el promedio o media aritmética. Esta es la primera medida de **tendencia central**

Advertencia: Los dos métodos presentados en este texto para resolver sistemas de ecuaciones lineales por mínimos cuadrados satisfacen la necesidad de comprender los conceptos teóricos que han sido el centro de los desarrollos actuales. Sin embargo desde el punto de vista computacional son inestables, o sea que son susceptibles de propagar errores por redondeo propios de los computadores, por lo tanto la solución de dichos problemas por tales métodos, utilizando computadores o calculadoras no es en general conveniente.
De aquí la importancia de explorar estos temas mas a fondo en cursos de métodos numéricos, en donde se plantearán otras técnicas. Al respecto puede consultar en el futuro la página web: www.abaco.com.ve

Dos ejemplos aplicados a aerolíneas

Los datos de los siguientes dos ejemplos son tomados del libro "Estadìstica Aplicada a los negocios y a la Economía". Autor: Wexler. Ed. McGraw Hill.

La compañía HOP SCOCHT AIRLANES desea hallar una relaciòn lineal entre sus gastos de publicidad en cada uno de 15 meses y el número de pasajeros que movilizó mensualmente. A partir de la siguiente tabla se efectuaron los càlculos de **regresión**.

Mes de observación	Gastos de publicidad en US$ 1.000's	Pasajeros movilizados En 1.000's
1	10	45
2	12	17
3	8	13
4	17	23
5	10	16
6	15	21
7	10	14
8	14	20
9	19	24
10	10	17
11	11	16
12	13	18
13	16	23
14	10	15
15	12	16

Asumiendo que la relación funcional entre las variables P : Número de pasajeros movilizados

G : gastos en publicidad, es del tipo lineal. $P = mG + b$

y reemplazando en la ecuaciòn anterior los datos recibidos, trataremos de calcular por regresión lineal, los coeficientes m y b, obteniendo el siguiente sistema de ecuaciones simultáneas en las incognitas m y b.

Obtenemos el sistema de 15 ecuaciones simultáneas con dos incógnitas:

El sistema de ecuaciones lineales

$$
\begin{array}{rcl}
10m + b &=& 15 \\
12m + b &=& 17 \\
8m + b &=& 13 \\
17m + b &=& 23 \\
10m + b &=& 16 \\
15m + b &=& 21 \\
10m + b &=& 14 \\
14m + b &=& 20 \\
19m + b &=& 24 \\
10m + b &=& 17 \\
11m + b &=& 16 \\
13m + b &=& 18 \\
16m + b &=& 23 \\
10m + b &=& 15 \\
12m + b &=& 16 \\
(*) & & (**)
\end{array}
$$

(*) Gastos Publicidad **(**)** Pasajeros movilizados

Es claramente inconsistente

Planteemos el sistema matricial **Ax = b**

En donde

$$
A = \begin{pmatrix} 10 & 1 \\ 12 & 1 \\ 8 & 1 \\ 17 & 1 \\ 10 & 1 \\ 15 & 1 \\ 10 & 1 \\ 14 & 1 \\ 19 & 1 \\ 10 & 1 \\ 11 & 1 \\ 13 & 1 \\ 16 & 1 \\ 10 & 1 \\ 12 & 1 \end{pmatrix}
\qquad
b = \begin{pmatrix} 15 \\ 17 \\ 13 \\ 23 \\ 16 \\ 21 \\ 14 \\ 20 \\ 24 \\ 17 \\ 16 \\ 18 \\ 23 \\ 15 \\ 16 \end{pmatrix}
\qquad b =
$$

Meses

Publicidad en US$ 1.000's

Pasajeros En 1.000s

Revisando cuidadosamente las filas y las columnas que intervienen en la ecuación normal:

$$\mathbf{A^T\, A\, x = A^T\, b}$$

concluimos que:

$$A^T A = \begin{pmatrix} \Sigma G_i^2 & \Sigma G_i \\ \Sigma G_i & \# \end{pmatrix} = \begin{pmatrix} 2469 & 187 \\ 187 & 15 \end{pmatrix} \qquad A^T b = \begin{pmatrix} \Sigma G_i \\ P_i \end{pmatrix} = \begin{pmatrix} 3.490 \\ 268 \end{pmatrix}$$,

Debemos resolver el sistema de ecuaciones simultáneas:

$$\begin{aligned} 2469\,m + 187\,b &= 3490 \\ 187m + 15\,b &= 268 \end{aligned}$$

Cuya soluciòn aproximada està dada por m = 1,08 y b= 4,39.

Por lo tanto: P =

es la ecuaciòn que relaciona por regresiòn lineal, al número de pasajeros transportados mensualmente, con los gastos mensuales en publicidad.

Esto nos permitirìa predecir "aproximadamente", el número de pasajeros que se podrían transportar a partir de los gastos mensuales en publicidad.

Este método fue utilizado por el astrónomo Matemático Carl Gauss para predecir la posición del desaparecido asteroide "Ceres" en 1801.

Si a los datos anteriores les anexamos la columna de ingreso nacional mensual (en billones de dólares) así:

Mes de observación	Gastos de publicidad en US$ 1.000's	Pasajeros movilizados En 1.000's	Ingreso nacional (en billones de dólares)
1	10	15	2.40
2	12	17	2.72
3	8	13	2.08
4	17	23	3.68
5	10	16	2.56
6	15	21	3.76
7	10	14	2.24
8	14	20	3.20
9	19	24	3.84
10	10	17	2.72
11	11	16	2.07
12	13	18	2.33
13	16	23	2.98
14	10	15	1.94
15	12	16	2.17

Al plantear la relación lineal entre las variables

P: Pasajeros, G: Gastos publicidad/mes, I: Ingreso Nacional/mes

Al plantear tambien, la relación funcional lineal:: $P = m_1 G + m_2 I + b$

Y al sustituir los datos, llegamos al sistema de 15 ecuaciones lineales con tres incognitas

$$\begin{array}{ccc} G_i & I_i & P_i \\ \downarrow & \downarrow & \downarrow \end{array}$$

$$\begin{aligned} 10\, m_1 + 2.40\, m_2 + b &= 15 \\ 12\, m_1 + 2.72\, m_2 + b &= 17 \\ 8\, m_1 + 2.08\, m_2 + b &= 13 \\ &\vdots \\ 12\, m_1 + 2.17\, m_2 + b &= 16 \end{aligned}$$

Planteando la ecuación normal tenemos que:

$$\begin{pmatrix} 10 & 12 & 8 & \dots & 16 & 10 & 12 \\ 2.40 & 2.72 & 2.08 & \dots & 2.98 & 1.94 & 2.17 \\ 1 & 1 & 1 & \dots & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 10 & 2.40 & 1 \\ 12 & 2.72 & 1 \\ 8 & 2.08 & 1 \\ & . & \\ & . & \\ & . & \\ 16 & 2.98 & 1 \\ 10 & 1.94 & 1 \\ 12 & 2.17 & 1 \end{pmatrix} = \begin{pmatrix} 2469 & 531,38 & 187 \\ 531,38 & 116,19 & 40,69 \\ 187 & 40,69 & 15 \end{pmatrix}$$

=

Los cálculos se efectuaron a partir de valores tomados de las tablas siguientes:

Mes de observación	Gastos de publicidad en US$ 1.000's	Pasajeros movilizados En 1.000's	Ingreso nacional (en billones de dólares)	G_iP_i
				150
1	10	15	2.40	204
2	12	17	2.72	104
3	8	13	2.08	391
4	17	23	3.68	160
5	10	16	2.56	315
6	15	21	3.76	140
7	10	14	2.24	280
8	14	20	3.20	456
9	19	24	3.84	170
10	10	17	2.72	176
11	11	16	2.07	234
12	13	18	2.33	368
13	16	23	2.98	150
14	10	15	1.94	192
15	12	16	2.17	
	$\Sigma G_i = 187$	$\Sigma P_i = 268$	$\Sigma I_i = 40,69$	$\Sigma G_iP_i = 3490$

G_iI_i	I_iP_i	I_i^2	G_i^2
24	36	5,76	100
32	46,24	7,40	144
16,64	27,04	4,33	64
62,56	84,64	13,54	289
25,6	40,96	6,55	100
56,4	78,96	14,14	225
22,4	31,36	5,02	100
44,8	64	10,24	196
72,96	92,16	14,75	361
27,20	46,24	7,40	100
22,77	33,12	4,28	121
30,29	41,94	5,43	169
47,68	68,54	8,88	256
19,40	29,1	3,76	100
26,04	34,72	4,71	144
$\Sigma G_iI_i=$ 531,38	$\Sigma I_iP_i =$ 755,02	$\Sigma I_i^2 =$ 116,19	$\Sigma G_i^2 = 2469$

Aᵀ **P** donde **P** es la matriz de los pasajeros transportados/mes, es:

$$\begin{pmatrix} 10 & 12 & 8 & \ldots & 16 & 10 & 12 \\ 2{,}40 & 2{,}72 & 2{,}08 & \ldots & 2{,}98 & 1{,}94 & 2{,}17 \\ 1 & 1 & 1 & \ldots & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 15 \\ 17 \\ 13 \\ . \\ . \\ . \\ 16 \end{pmatrix} = \begin{pmatrix} \Sigma G_iP_i \\ \Sigma I_i P_i \\ \Sigma P_i \end{pmatrix} = \begin{pmatrix} 3490 \\ 755{,}02 \\ 268 \end{pmatrix}$$

La solución a este problema según Wexler es:

$$P = 0{,}84\ P + 1.44\ I + 3{,}53$$

Note como un problema que involucraba matrices de 15 filas, se ha reducido por la ecuación normal a un problema con matriz de coeficientes, cuadrada de orden 3, el número de columnas de la matriz original.

Ejercicios

1)Halle la ecuación de la recta que mejor ajusta en el sentido de los mínimos cuadrados a los puntos P(1,1), Q(4,2), R(2,3).

a) Halle la suma de las desviaciones al cuadrado.

b) Calcular $|\ \mathbf{b} - \bar{\mathbf{b}}\ |$ c) Calcule $|\ \mathbf{b} - \bar{\mathbf{b}}\ | / |\ \mathbf{b}\ |$ (medida de **error relativo**)

d) Compare las medidas $\sqrt{(\Sigma r_i^2)}$ y $(\sqrt{(r_i^2)}) / |\mathbf{b}|$, de las desviaciones de los datos respecto a

i) La recta que pasa por P(1,1), Q(4,2).
(R. / Ecuación de dicha recta, la cual debe ser calculada: y = 2x - 1).
ii) La recta que pasa por P(1,1), R(2,3).
iii)La recta solución por mínimos cuadrados hallada anteriormente.

e) Compare los resultados de d) y comente respecto a que la solución por mínimos cuadrados es la que mejor ajusta los datos en este sentido.

2) Halle la ecuación de la recta y = cx + d que mejor ajusta los puntos P(1,2), Q(3,5), R(4,8) en el sentido de los mínimos cuadrados. Calcule medidas de desviación similares a las del problema 1d). R) y = 27/1

3) Demuestre que en la ecuación normal que halla por la técnica de los mínimos cuadrados, la recta
Y = cx + d que mejor ajusta los datos

a	b
a_1	b_1
a_2	b_2
.	.
.	.
a_m	b_m

participan las matrices:

$$\begin{pmatrix} \Sigma a_i^2 & \Sigma a_i \\ \Sigma a_i & m \end{pmatrix} \qquad \begin{pmatrix} \Sigma a_i b_i \\ \Sigma b_i \end{pmatrix}$$

4) Determinar una ecuación de la forma P = c + d E, que relacione por la técnica de los mínimos cuadrados, la edad promedio E de un grupo de personas con su presión sanguínea P, teniendo en cuenta los siguientes datos

Punto medio de las edades	Presión sanguínea media
35	114
45	124
55	143
65	158
75	166

R/ $P = 65{,}1 + 1{,}38\, E$. Desviaciones $\approx r_i$

0,6 ; -3,2 ; 2 ; 3,2 ; -2,6

5) En cada uno de los problemas a), b), c), d), halle las rectas que mejor ajustan los datos en cada una de las tablas, en el sentido de los mínimos cuadrados. Calcule las desviaciones r_i .

a)

a_i	b_i
1	3
2	5
3	6
4	9

R/ $y = 1{,}9\, x + 1$

b)

a i	b i
0	1,2
1	-3,1
1	-7
2	-11

R/ $y = -4{,}05\, x + 1{,}1$

c)

a i	b i
1	2
2	3
3	4
4	5,1

R/ $y = 1{,}03\, x + 0{,}95$

d)

a i	b i
1	2
2	3
3	4
4	5,01
5	6

R/ $y = 1{,}001\, x + 0{,}99$

6) Las siguientes son mediciones de alturas de plantas de frijol de soya de un campo, efectuando una selección al azar cada semana.

Edad en semanas (E)	1	2	3	4	5	6	7
Altura en Cmts. (A)	5	13	16	23	33	38	40

Calcule la ecuación de la recta $A = dE + c$ que mejor ajusta las variables, utilizando la técnica de los mínimos cuadrados. Calcule las desviaciones r_i . R/ $6{,}143\, E - 0{,}572$

7) La siguiente tabla relaciona los diámetros (x) con los pesos (y) de bulbos de cebollas.

x	y
35,1	24,3
41,4	42,0
51,0	63,4
60,0	96,2

Halle utilizando la técnica de los mínimos cuadrados, los valores de c y d, en la relación $y = c\, x^d$

Utilizando la ecuación equivalente: $\log y = \log c + d \log x$.Calcule las desviaciones.

8) La siguiente tabla relaciona los censos de población (y) de una ciudad, con el tiempo transcurrido en décadas, a partir de 1860

Año (y)	Tiempo (x)	Población
1860	0	731
1880	2	2666
1900	4	17700
1920	6	74631
1940	8	203341

Halle por el método de los mínimos cuadrados los valores de c y d, que relacionan a las variables en la ecuacióny = c d ^{x} , utilizando la ecuación equivalente Log y = log c + (log d) x

Calcule las desviaciones r_i a partir de y = c d ^{x} .

9) La siguiente tabla relaciona el peso en gramos (y) de fríjoles verdes recogidos en un campo a medida que el tiempo (x) en días, transcurre a partir de 0.

x	y
0	27,4
4	39,3
7	46,2
10	47,8
13	44,5
18	24,5

a) Halle utilizando la técnica de los mínimos cuadrados los valores de c y d que relacionan a las variables en la ecuación y = cx + d. R/ y = -0,0609 x + 38,81

Calcule las desviaciones r_i .

b) Calcule utilizando la técnica de los mínimos cuadrados los valores de e , f , g que relacionan a las variables en la ecuación:

y = e + f x + g x ^{2}. Calcule las desviaciones r_i .

Cual de las dos relaciones encontradas, ajusta con mejor exactitud los datos de la tabla, según su criterio ?

R/ y = 26,3327 + 4,7828 x - 0,2690 x ^{2}.