Calculo Numérico MA33A.-Primavera 2007
Prof.: Jaime Michelow
Aux.: Ignacio Trujillo-Hugo Ulloa

Guía de Mínimos cuadrados –Problemas Propuestos Control 3

Resumen

Los mínimos cuadrados es una herramienta muy utilizada desde su invención por Gauss y Legendre hacia el años 1800. En términos de algebra lineal se trata de resolver sistemas lineales $Ax = b$ con más ecuaciones que incógnitas, donde el término a resolver se entiende en el sentido de minimizar la norma euclidiana del vector residuo $Ax - b = r$.

Es frecuente el problema de tener un conjunto de mediciones y desear ajustar los datos observados a funciones que describan la relación entre la variable dependiente y la variable independiente. El problema original parece estar asociado al nombre de Gauss quien trató de ajustar curvas a datos experimentales obtenidos en observaciones astronómicas. En nuestros días el planteamiento permite establecer un modelo matemático basado en optimización, para dar solución al problema de calcular los coeficientes que satisfacen el criterio de mínimos cuadrados. En este resumen se muerta de forma sencilla la solución matricial del problema ilustrando la forma de llegar a una expresión cerrada de la solución.

Desarrollo de las ecuaciones para el cálculo de los coeficientes.

Con el objetivo de establecer la nomenclatura general se considera que se tiene un conjunto de n pares de observaciones (x_i, y_i). En un plano x-y se puede visualizar la distribución de observaciones usando un diagrama de dispersión.

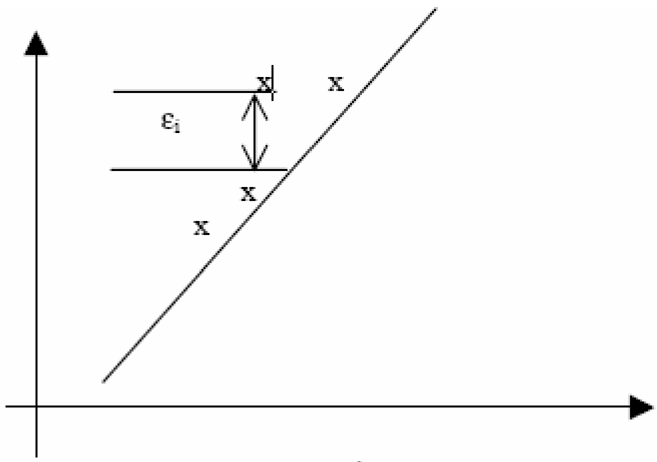

Fig. 1. Diagrama de dispersión, observaciones en el plano x-y.

Como se muestra en la figura 1, para cada observación se tiene un valor teórico expresado por la ecuación de la función. Así, se puede hablar de errores o desajuste a la diferencia entre el valor observado y el valor teórico. Al sumar los errores al cuadrado se tiene SS .

$$SS = e_1^2 + e_2^2 + e_3^2 + ... + e_n^2 = \sum_{i=1}^{n} e_i^2$$

En el caso particular cuando los datos pueden ser descritos por una recta, $y = a_1 x + a_o$ se requieren los coeficientes a_1 y a_o. Los mejores coeficientes serán aquellos que minimizan el cuadrado de los errores, dando lugar al método de mínimos cuadrados. El gradiente de SS igualado a cero será la solución buscada. De la figura 1:

$$e_i = y_i^{obs} - y_i^{teórica} = y_i^{obs} - a_1 x_i + a_o$$

Para el caso particular de $n = 4$ observaciones, suma SS del cuadrado de los errores:

$$SS = (y_1 - a_1 x_1 - a_o)^2 + (y_2 - a_1 x_2 - a_o)^2 + (y_3 - a_1 x_3 - a_o)^2 + (y_3 - a_1 x_3 - a_o)^2$$

El gradiente de SS:

$$\nabla SS = \begin{bmatrix} \dfrac{\partial SS}{\partial a_1} \\ \dfrac{\partial SS}{\partial a_0} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

y las componentes del gradiente se obtienen como:

$$\frac{\partial SS}{\partial a_1} = -2(y_1 - a_1 x_1 - a_o)x_1 - 2(y_2 - a_1 x_2 - a_o)x_2 - 2(y_3 - a_1 x_3 - a_o)x_3 - 2(y_1 - a_1 x_4 - a_o)x_4$$

$$\frac{\partial SS}{\partial a_o} = -2(y_1 - a_1 x_1 - a_o) - 2(y_2 - a_1 x_2 - a_o) - 2(y_3 - a_1 x_3 - a_o) - 2(y_1 - a_1 x_4 - a_o)$$

AL igualar a cero las componentes del gradiente y en forma matricial:

$$[(y_1 - a_1 x_1 - a_o), (y_2 - a_1 x_2 - a_o), (y_3 - a_1 x_3 - a_o), (y_3 - a_1 x_3 - a_o)]\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = [0]$$

$$[(y_1 - a_1 x_1 - a_o), (y_2 - a_1 x_2 - a_o), (y_3 - a_1 x_3 - a_o), (y_3 - a_1 x_3 - a_o)]\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = [0]$$

Las ecuaciones anteriores se pueden escribir en una misma ecuación.

$$[(y_1 - a_1 x_1 - a_o), (y_2 - a_1 x_2 - a_o), (y_3 - a_1 x_3 - a_o), (y_3 - a_1 x_3 - a_o)]\begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ x_3 & 1 \\ x_4 & 1 \end{bmatrix} = [0 \quad 0]$$

Y al transponer esta ultima expresión

$$\begin{bmatrix} x_1 & x_2 & x_3 & x_4 \\ 1 & 1 & 1 & 1 \end{bmatrix}\begin{bmatrix} (y_1 - a_1 x_1 - a_o) \\ (y_2 - a_1 x_2 - a_o) \\ (y_3 - a_1 x_3 - a_o) \\ (y_3 - a_1 x_3 - a_o) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Se puede arreglar la ultima ecuación y obtener una forma condensada de la solución para a_1 y a_o.

$$\begin{bmatrix} x_1 & x_2 & x_3 & x_4 \\ 1 & 1 & 1 & 1 \end{bmatrix} \left\{ \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} - \begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ x_3 & 1 \\ x_4 & 1 \end{bmatrix} \begin{bmatrix} a_1 \\ a_0 \end{bmatrix} \right\} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Llamando A a la matriz que tiene la información de la medición en variables independiente x, como A^T a la matriz que presenta en forma de reglones las columnas de A y al vector de información de medición en la variable dependiente y, se tiene:

$$A^T(y - Aa) = 0$$
$$A^T y - A^T Aa = 0$$

Con solución:

$$a = \left(A^T Aa\right)^{-1} Ay \qquad (1)$$

La dimensión de A depende de n número de observaciones y del número de coeficientes a determinar.

Ejemplos

1.- Para ilustrar la aplicación de **(1)** se propone los siguientes pares de valores medidos, para la variable independiente x y el correspondiente valor de la variable dependiente y **.**

x_i	y_i
1.0	1.0
1.8	3.0
2.0	1.8
3.0	2.9

Los arreglos de información para n=4:

$$x = \begin{bmatrix} 1.0 \\ 1.8 \\ 2.0 \\ 3.0 \end{bmatrix} \qquad y = \begin{bmatrix} 1.0 \\ 3.0 \\ 1.8 \\ 2.9 \end{bmatrix} \qquad A = \begin{bmatrix} 1.0 & 1 \\ 1.8 & 1 \\ 2.0 & 1 \\ 3.0 & 1 \end{bmatrix} \qquad A^T = \begin{bmatrix} 1.0 & 1.8 & 2.0 & 3.0 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

$$A^T A = \begin{bmatrix} 1.0 & 1.8 & 2.0 & 3.0 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1.0 & 1 \\ 1.8 & 1 \\ 2.0 & 1 \\ 3.0 & 1 \end{bmatrix} = \begin{bmatrix} 17.24 & 7.80 \\ 7.80 & 4.00 \end{bmatrix}$$

$$\left(A^T A\right)^{-1} = \begin{bmatrix} 0.4926 & -0.9606 \\ -0.9606 & 2.1232 \end{bmatrix}$$

$$A^T y = \begin{bmatrix} 1.0 & 1.8 & 2.0 & 3.0 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1.0 \\ 3.0 \\ 1.8 \\ 2.9 \end{bmatrix} = \begin{bmatrix} 18.7 \\ 8.7 \end{bmatrix}$$

$$\begin{bmatrix} a_1 \\ a_0 \end{bmatrix} = \left[A^T A\right]^{-1} A^T y$$

$$\begin{bmatrix} a_1 \\ a_0 \end{bmatrix} = \begin{bmatrix} 0.4926 & -0.9606 \\ -0.9606 & 2.1232 \end{bmatrix} \begin{bmatrix} 18.7 \\ 8.7 \end{bmatrix} = \begin{bmatrix} 0.8547 \\ 0.5084 \end{bmatrix}$$

Con solución para los coeficientes a_1 y a_0, que resultan en la recta $y = 0.8547x + 0.5084$. Una grafica de la recta ajustada y de los valores observados se muestra en la figura 2.

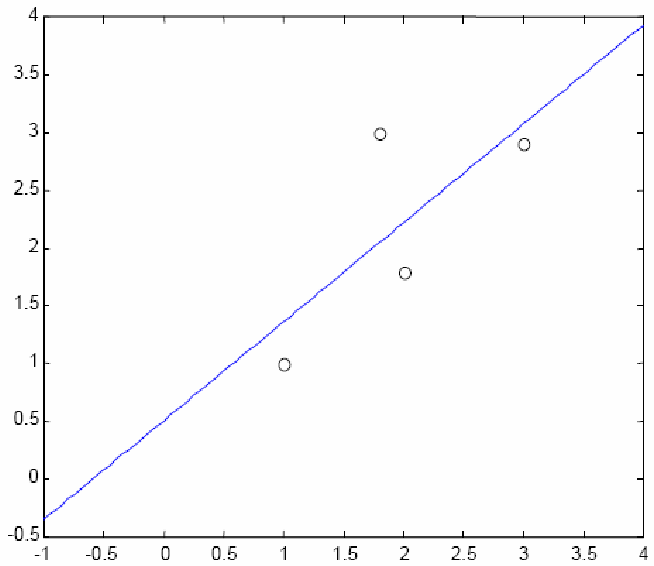

Fig. 2. Gráfica de función ajustada y = 0.85468 x + 0.50837 y las mediciones.

2.- Modelo no-lineal, observaciones originales (x, y) y transformaciones logarítmicas.

Dato	x(i)	Ln[x(i)]	y(i)	Ln[y(i)]
1	1.00000	0.00000	148.00	4.99721
2	2.00000	0.69315	1487.00	7.30452
3	3.00000	1.09861	3807.00	8.24460
4	4.00000	1.38629	10498.00	9.25894
5	5.00000	1.60944	17551.00	9.77287
6	6.00000	1.79176	34057.00	10.43579
7	7.00000	1.94591	48905.00	10.79763
8	8.00000	2.07944	76987.00	11.25139
9	9.00000	2.19722	109193.00	11.60087
10	10.00000	2.30259	147413.00	11.90099

Matriz de información A usando Ln (x)

0.00000	1.00000
0.69315	1.00000
1.09861	1.00000
1.38629	1.00000
1.60944	1.00000
1.79176	1.00000
1.94591	1.00000
2.07944	1.00000
2.19722	1.00000
2.30259	1.00000

La matriz A transpuesta

0.00000	0.69315	1.09861	1.38629	1.60944	1.79176	1.94591	2.07944	2.19722	2.30259
1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000

$$A^T y = \begin{bmatrix} 158.6842 \\ 95.5648 \end{bmatrix}$$

$$\begin{bmatrix} a_1 \\ a_0 \end{bmatrix} = \left[A^T A\right]^{-1} A^T y = \begin{bmatrix} 2.9651 \\ 5.0778 \end{bmatrix}$$

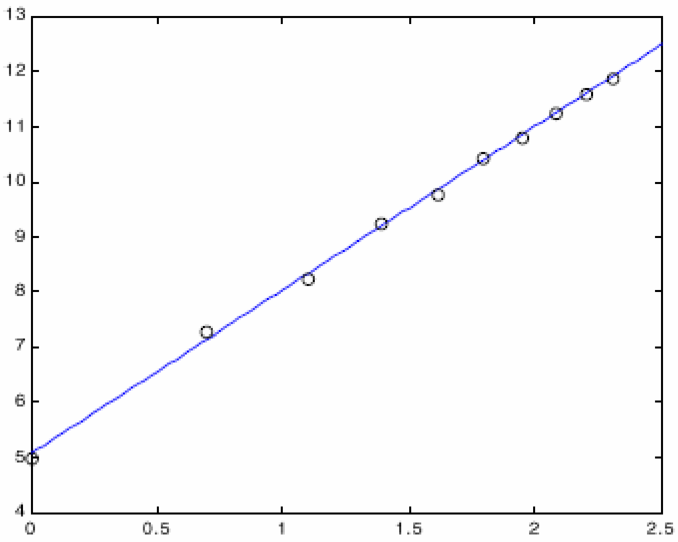

Fig. 3. Ajuste de datos transformados a una forma lineal
Ln[y] = 2.96514 Ln[x] + 5.07782.

+

Para este segundo ejemplo se observa que una transformación logarítmica en ambas variables permite tener un problema lineal de dos coeficientes. El ajuste que se logra se puede observar en la figura 3. en las figuras 4 y 5se muestran los datos originales y el efecto de la transformación logarítmica sobre los datos originales.

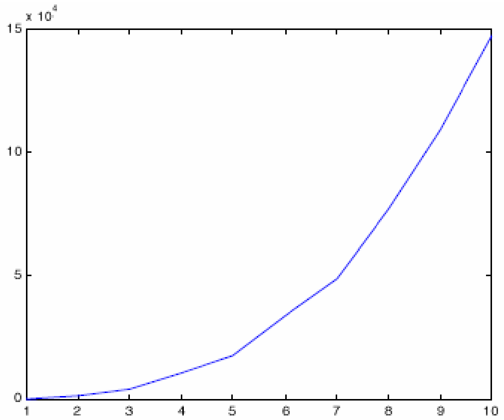

Fig. 4. Datos observados y vs. x, Ejemplo 2.

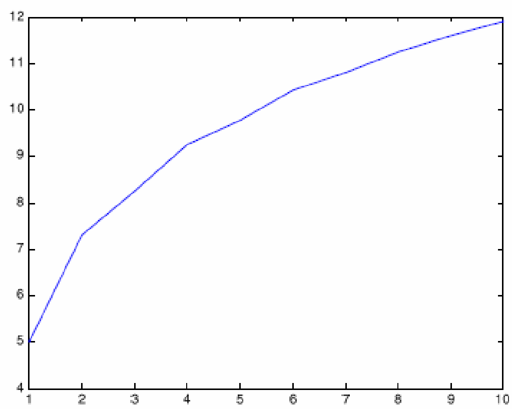

Fig. 5 Datos transformados Ln[y] vs. x, Ejemplo 2.

Problemas Propuestos.

P1.- El nivel del agua en el mar del Norte está principalmente determinado por el llamado índice de marea que tiene la forma

$$H(t) = h_0 + a_1 \sin\left(\frac{pt}{6}\right) + a_2 \cos\left(\frac{pt}{6}\right)$$

Con t medido en horas. Ajuste los valores de los parámetros si han tomado las siguientes mediciones

t	0	2	4	6	8	10
$H(t)$	1.0	1.6	1.4	0.6	0.2	0.8

P2.- Para conocer la relación entre la velocidad de caída de un paracaidista y la fuerza de fricción hacia arriaba, se han efectuado las siguientes mediciones

v	1.0	2.0	3.0	4.0	5.0
f	5	15.3	29.3	46.4	66.3

Donde v se mide en centímetros por segundo y rozamiento f en 10^6dinas. Dibuje los puntos de la tabla y obtenga la aproximación de mínimos cuadrados.

P3.- En 1601 el astrónomo alemán J. Kepler formulo su tercera ley del movimiento planetario: $T = cd^{3/2}$ donde "d" es la distancia de un planeta al sol medida en millones de kilómetros, "T" es el periodo orbital y "c" es una constante. Los datos observados para los cuatro planetas Mercurio, Venus, Tierra y Marte son:

d_i	58	108	155	228
T_i	88	225	365	687

Ajuste el valor "c" para estos datos en el sentido de mínimos cuadrados.

Referencias:

[1] Stephen A. DeLurgio, Forecasting Principles and Applicatións, McGraw-Hill Edición intenacional, 1998.