Ajuste de datos experimentales por medio de mínimos cuadrados
Parte I

Salvador Acha Daza*, Diana Acha Izquierdo*

Abstract

In the present work we try to show in simple terms the curve fitting approach to experimental data, using the concept of minimum squares. In many engineering and science applications, models are linear in the coefficients; which meas that an explicit solution to the problem can be obtained. When redundancy exists, this means a larger number of observations than coefficients to be determined, it can be established under appropriate statistical considerations a significance test for the parameters of the proposed model, this topic will be covered in part II.

Keywords*: minimum square, curve fitting.*

INTRODUCCIÓN

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. El presente artículo muestra de forma sencilla la solución matricial del problema, ilustrando la forma de llegar a una expresión cerrada de la solución. Con ejemplos sencillos se tiene una comprobación al aplicar en forma directa la solución presentada. En el apartado 1 se desarrollan las ecuaciones para el cálculo de los coeficientes, después en el apartado 2 se discuten dos ejemplos, en uno se muestran las transformaciones para lograr un modelo como el discutido con detalle en el apartado 1. Al final se presentan las conclusiones del trabajo y en apéndices se extiende el material para incluir pruebas de significación estadística lo cual debe complementar el análisis y agrega valor al modelo obtenido bajo el criterio de los mínimos cuadrados.

DESARROLLO DE ECUACIONES PARA EL CÁLCULO DE LOS COEFICIENTES

Con el fin 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, figura 1.

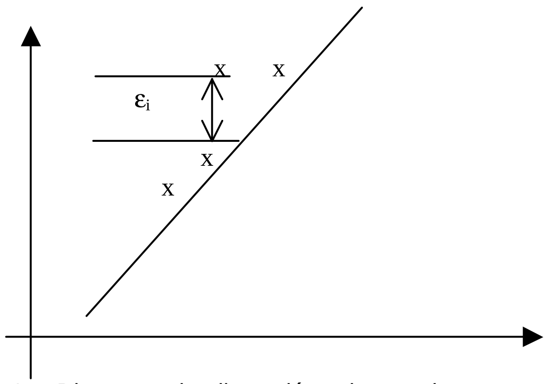

Fig. 1. Diagrama de dispersión, observaciones en el plano x-y.

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 un error o "desajuste" a la diferencia entre el valor observado y el valor teórico. Al sumar los errores al cuadrado, se tiene SS.

* Programa Doctoral en Ingeniería Eléctrica FIME-UANL.

$$SS = \varepsilon_1^2 + \varepsilon_2^2 + \ldots + \varepsilon_n^2 = \sum_{i=1}^{n} \varepsilon_i^2 \qquad (1)$$

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:

$$\varepsilon_i = y_i^{obs} - y_i^{teórica} = y_i^{obs} - (a_1 x_i + a_o) \qquad (2)$$

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

$$SS = (y_1 - a_1 x_1 - a_o)^2 + (y_2 - a_1 x_2 - a_o)^2 + \ldots (y_3 - a_1 x_3 - a_o)^2 + (y_4 - a_1 x_4 - a_o)^2 \qquad (3)$$

El gradiente de SS:

$$\nabla SS = \begin{bmatrix} \dfrac{\partial SS}{\partial a_1} \\ \dfrac{\partial SS}{\partial a_o} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \qquad (4)$$

y las componentes del gradiente (4) se obtienen como:

$$\frac{\partial SS}{\partial a_1} = -2(y_1 - a_1 x_1 - a_o) x_1 - 2(y_2 - a_1 x_2 - a_o) x_2 - \ldots 2(y_3 - a_1 x_3 - a_0) - 2(y_4 - a_1 x_4 - a_o) x_4$$

$$\frac{\partial SS}{\partial a_o} = -2(y_1 - a_1 x_1 - a_o)(1) - 2(y_2 - a_1 x_2 - a_o)(1) - \ldots 2(y_3 - a_1 x_3 - a_o)(1) - 2(y_4 - a_1 x_4 - a_o)(1)$$

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

$$[(y_1 - a_1 x_1 - a_2)\ (y_2 - a_1 x_2 - a_2)\ (y_3 - a_1 x_3 - a_2)\ (y_4 - a_1 x_4 - a_2)] \ldots \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = [0] \qquad (5.a)$$

$$[(y_1 - a_1 x_1 - a_o)\ (y_2 - a_1 x_2 - a_o)\ (y_3 - a_1 x_3 - a_o)\ (y_4 - a_1 x_4 - a_o)] \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = [0] \qquad (5.b)$$

Las formas anteriores (5.a) y (5.b) se pueden escribir en una sola expresión matricial:

$$[(y_1 - a_1 x_1 - a_o)\ (y_2 - a_1 x_2 - a_o)\ (y_3 - a_1 x_3 - a_o)\ (y_4 - a_1 x_4 - a_o)] \begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ x_3 & 1 \\ x_4 & 1 \end{bmatrix} = [0 \quad 0] \qquad (6)$$

y al transponer (6):

$$\begin{bmatrix} x_1 & x_2 & x_3 & x_4 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} y_1 - a_1 x_1 - a_o \\ y_2 - a_1 x_2 - a_o \\ y_3 - a_1 x_3 - a_o \\ y_4 - a_1 x_4 - a_o \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \qquad (7)$$

Se puede arreglar (7) y obtener una forma condensada de la solución para a_1 y a_o.

$$\begin{bmatrix} x_1 & x_2 & x_3 & x_4 \\ 1 & 1 & 1 & 1 \end{bmatrix} \left\{ \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} - \begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ x_3 & 1 \\ x_4 & 1 \end{bmatrix} \begin{bmatrix} a_1 \\ a_o \end{bmatrix} \right\} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{8}$$

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

$$A^t (y - A\,a) = 0 \tag{9}$$

$$A^t y - A^t A\,a = 0 \tag{10}$$

con solución:

$$a = \left(A^t\ A\right)^{-1} A^t\ y \tag{11}$$

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

EJEMPLOS

a) Para ilustrar la aplicación de (11) 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:

$$x = \begin{bmatrix} 1.0 \\ 1.8 \\ 2.0 \\ 3.0 \end{bmatrix}$$

$$y = \begin{bmatrix} 1.0 \\ 3.0 \\ 1.8 \\ 2.9 \end{bmatrix}$$

$$A = \begin{bmatrix} 1.0 & 1 \\ 1.8 & 1 \\ 2.0 & 1 \\ 3.0 & 1 \end{bmatrix}$$

$$A^t = \begin{bmatrix} 1.0 & 1.8 & 2.0 & 3.0 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

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

$$\left[A^t A\right]^{-1} = \begin{bmatrix} 0.49261 & -0.96059 \\ -0.96059 & 2.12315 \end{bmatrix}$$

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

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

$$\begin{bmatrix} a_1 \\ a_o \end{bmatrix} = \begin{bmatrix} 17.24 & 7.80 \\ 7.80 & 4.00 \end{bmatrix}^{-1} \begin{bmatrix} 18.7 \\ 8.7 \end{bmatrix}$$

$$\begin{bmatrix} a_1 \\ a_o \end{bmatrix} = \begin{bmatrix} 0.85468 \\ 0.50837 \end{bmatrix}$$

Con solución para los coeficientes a_1 y a_o, que resultan en la recta y = 0.85468 x + 0.50837. Una gráfica de la recta ajustada y de los valores observados se muestra en la figura 2.

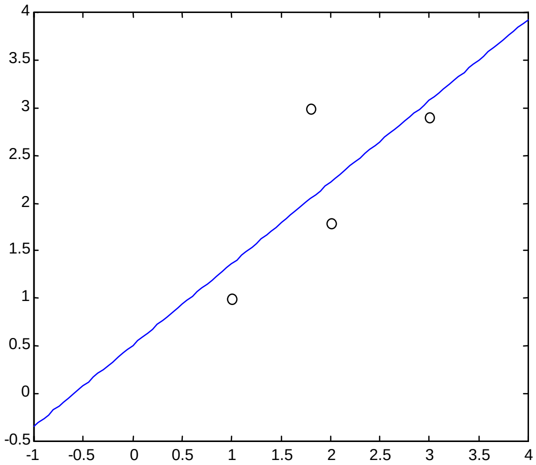

Fig. 2. Gráfica de función ajustada y = 0.85468 x + 0.50837 y las mediciones.

b) Modelo no-lineal, observaciones originales (x,y) y transformaciones logarítmicas, pag. 123, Ref. [1].

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

Dato	$A^t y$
1	158.68420
2	95.56481

Coeficientes solución a(i)

1	2.96514
2	5.07782

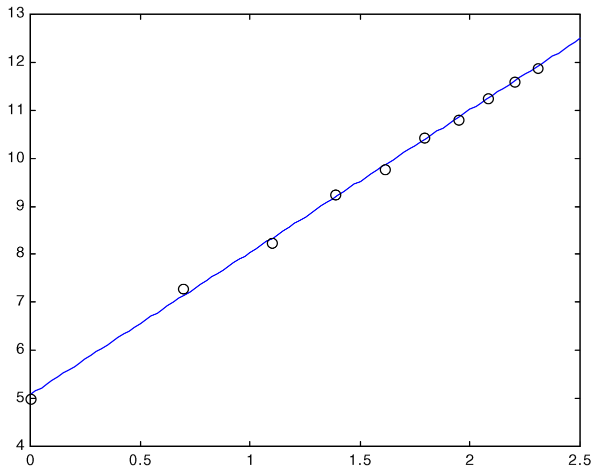

Fig. 3. Ajuste de datos transformados a una forma lineal Ln[y] = 2.96514 Ln[x] + 5.07782.

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 5 se muestran los datos originales y el efecto de la transformación logarítmica sobre los datos originales.

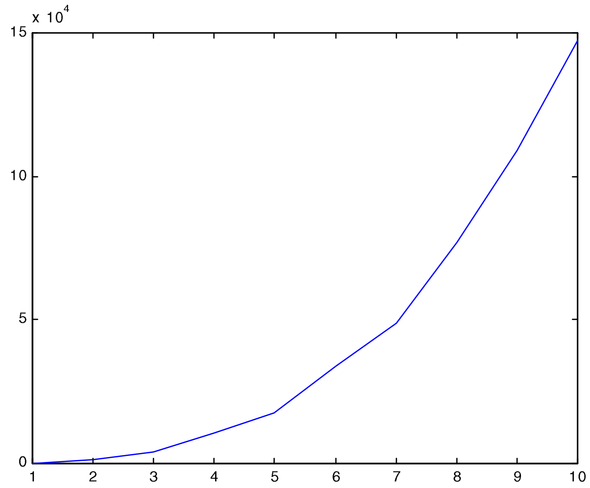

Fig. 4. Datos observados y vs. x, Ejemplo 2.

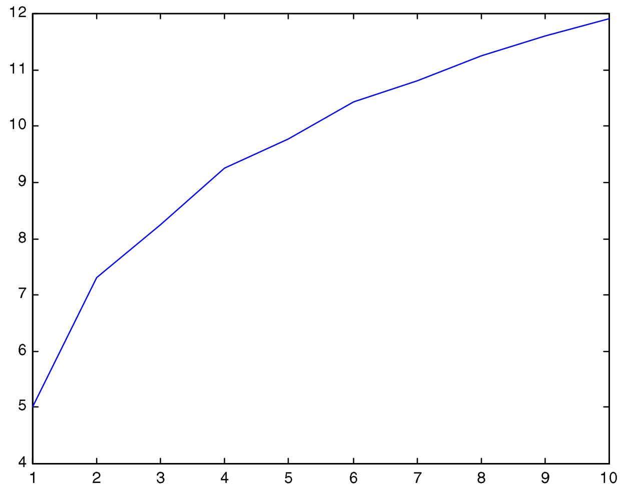

Fig. 5 Datos transformados Ln[y] vs. x, Ejemplo 2.

CONCLUSIONES

En el trabajo se han presentado las ideas fundamentales para el ajuste de datos experimentales a polinomios o formas no lineales, que forman un sistema lineal en los coeficientes. Se desarrolló de manera detallada el procedimiento para obtener las fórmulas matriciales generales las cuales permiten resolver este importante problema en la práctica de la ingeniería y de la ciencia. Al tener un mayor número de mediciones, se aumenta la REDUNDANCIA y se puede aplicar pruebas de hipótesis sobre los coeficientes y detección de datos erróneos. También se puede asignar un mayor peso o "credibilidad" que refleje la calidad de la medición; con este procedimiento se obtiene el resultado que se conoce como mínimos cuadrados ponderados.

REFERENCIAS

1. Stephen A. DeLurgio, Forecasting Principles and Applications, McGraw-Hill International Editions, 1998.
2. Masterton, Slowinski, Elementary Mathematical Preparation for General Chemistry, Saunders Golden Series, 1974.
3. V. P. Spiridonov, A. A. Lopatkin, Tratamiento Matemático de Datos Físico-Químicos, Editorial MIR, Moscú, 1973.